The University of Texas at Arlington

MODELING OF HYDROGEN SULFIDE GENERATION AND RATES OF MICROBIALLY INDUCED CONCRETE CORROSION IN MANHOLE SHAFTS

By

SUNAKSHI HADA

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in

Civil Engineering

THE UNIVERSITY OF TEXAS AT ARLINGTON
August 2021

ACKNOWLEDGEMENTS

Doing a Ph.D. was harder than I had imagined; however, it was more rewarding than I can express. None of this would have been possible without the grace and blessings of the Almighty God, to whom I express my gratitude for providing me the patience, wisdom, strength, and opportunity to complete this research successfully.

Nothing can be achieved alone; the support and strength of people around us drive us to continue fighting and strive for better. Fortunately, I am blessed to have a supportive circle of family, friends, and mentors, who made this Ph.D. journey successful. To all the individuals who provided me guidance, support, and encouragement, I would like to say thank you.

Dr. Melanie L. Sattler is one of the finest mentors, always encouraging and supportive of my career aspirations. I would like to thank her for her invaluable supervision, and guidance throughout my Ph.D. journey. Brainstorming sessions with her helped me become a better researcher by developing critical thinking and deep diving to look for answers. She always encouraged me to do better with her constructive feedback and is the greatest source of inspiration for me. She not only guided me to be a better researcher but also gave me a chance to develop leadership skills by giving me an opportunity to lead a project with a great group of individuals.

I would like to express my gratitude to my committee member Dr. Victoria Chen, who helped me define the multiple linear regression process, with her insightful feedback. I really appreciate her timely response and inputs to propel my analysis. Without her guidance it would have been difficult to complete the analysis. She helped me to identify minute details that go in a model building process and improving the model.

I would like to thank Dr. Srinivas Prabakar for devoting his time to be on my committee and providing his valuable feedback which greatly helped me improve the analysis, and illuminating several key issues that this work now addresses.

I would like to thank my committee member Dr. Arpita Bhatt for helping me with field studies and guidance on data collection. Her constant support during the data collection phase of the project was so crucial, without which the project would not have been successful.

I am profoundly grateful to my team members, who helped me with field and lab data collection. The field and lab work required to continue irrespective of the summer or winter as we were trying to capture effect of seasons as well, and during this two and a half years of rigorous data collection, I got a chance to lead an awesome team of 20 people (now some of the closest friends) who provided their constant support during the data collection. I would like to express my gratitude to Dr. Niloofar Parsaeifard, Dr. Sunita Baniya, Dr. Gomathy Radhakrishna Iyer, Dr. Hoda Rahimi, Dr. Ketan Shah, Aiswarya Acharath Mohanakrishnan, Mithila Chakraborty, Bria Natasha Wooten, Pablo Choquis Rosales, Tyler Terashima, Samir Nathu, Jinsae Kim, Mohammad Rafatuzzama, Harshika Jha, Ankita Sinha, Mateo Galvez, Ellen Dinh, Opeyemi Emmanuel Adelegan, Hussain Ali, Sorakrich Techapaphawit. Also, a special thanks to my friends Dr. Naomi Gevaerd De Souza and Dr. Akshay Chandrashekar Parenky for their constant support.

I would like to express my gratitude to the City of Arlington, for providing the funding for the project, and also for their constant support with manhole information and access which helped us to collect data for 366 manholes in the City of Arlington.

All this would not be possible without the strength and support of my family, who encouraged me to dream big and work to fulfill it. It was my parents Neera Hada and Balkishan Hada, who dreamt to see me complete a Ph.D., and I would like to say a big thank you to them for all the sacrifices they made for me to achieve this. My deepest love and gratitude to my two pillars of strength, my brother Manav and my husband Rajesh Kumar, who are always there for me as an inspiration, guidance, and my biggest cheerleaders.

Dedicated to my parents Neera and Balkishan Hada, for their love, encouragement, and sacrifices.

ABSTRACT

MODELING OF HYDROGEN SULFIDE GENERATION AND RATES OF MICROBIALLY INDUCED CONCRETE CORROSION IN MANHOLE SHAFTS

Sunakshi Hada, Ph.D.

The University of Texas at Arlington, 2021

Supervising Professor: Dr. Melanie L. Sattler

Manhole shafts can undergo infrastructure failure due to microbially induced concrete corrosion (MICC). Deterioration of the sewer systems by MICC process occurs due to the volatilization of hydrogen sulfide into the gas phase. The volatized hydrogen sulfide gas reacts with oxygen to form sulfuric acid, resulting in concrete corrosion. If optimum conditions are present, 24 mm of concrete loss can occur in 31 months, and the cost of fixing can range from $10-$30 per square foot. Also, if the manholes are not properly maintained, they can cause serious safety hazards, such as manhole lids falling through the structure, streets giving way, or sewers flowing beyond their capacity.

The project aimed to identify and predict factors contributing to corrosion in the unsubmerged sections of the manholes. Wastewater characteristics and sewer system designs affect sulfide generation, volatilization, and corrosion of concrete manholes. Since various factors contribute towards the MICC process, the initial phase of the project involved data collection of liquid and gas phase parameters and measuring depth of corrosion. The second stage of the project involved developing models to predict gas-phase hydrogen sulfide concentrations and corrosion rates.

Based on their physical design, 366 manholes in the City of Arlington were selected for the study. Since physical design of the manhole determines the rates of corrosion, the various design criteria considered for manhole selection were presence of drop, pipe size changes, type of flows, presence of bends, and number of inlets. Field studies were conducted with 48 hours of continuous data collection of gas-phase parameters, including temperature, hydrogen sulfide concentrations, relative humidity, oxygen concentration, and liquid phase parameters including temperature, dissolved oxygen, pH, sulfide, sulfate, and biochemical oxygen demand. The collected data was then used to build predictive models using multiple linear regression. A preliminary model was developed to address concerns with normality or nonconstant variance and carried out necessary transformations. Possible interactions were identified based on literature and added to the transformed model. Best subset and stepwise method were adopted to

identify best possible model using Cp value, variance inflation factor, Adjusted R^2, AIC, BIC, and p-value. Detailed analysis utilizing model diagnostics (modified Levene's test, normality test, and outlier test), and model validation using the test set was conducted to identify the best possible model.

Plots of data collected showed manholes with hydraulic jump or downstream of a lift station recorded highest H_2S concentration in comparison to the other manhole types, and the lowest was observed in manholes with subcritical flow. In addition, for manholes studied for seasonal variation, the concentration of hydrogen sulfide was found to be maximum during the summer, when the temperature was elevated, and less during fall and winter.

The multiple linear regression model for hydrogen sulfide had predictor variables (significant at $\alpha = 0.1$) of velocity, presence of hydraulic jump, precipitation, pH, manhole design with inlet smaller in diameter than outlet pipe, straight flow, location of manhole upstream to a high hydrogen sulfide manhole, and the interaction of sulfate with dissolved oxygen and supercritical flow with dissolved oxygen ($R^2=0.36$). High velocities, hydraulic jump, and supercritical flow will generate turbulence, which transfers hydrogen sulfide from the liquid to gas phase. Precipitation lowers of ambient temperatures, which reduces hydrogen sulfide volatilization according to Henry's law.

The multiple linear regression model for corrosion had predictor variables (significant at $\alpha = 0.1$) of manhole age, presence of right angle, presence of multiple inlets, BOD concentration and interaction between manhole age with gas phase temperature ($R^2=0.35$). The corrosion model association with the age of the manhole is important: concrete characteristics change with age, thus effecting hydrogen sulfide diffusion into the concrete.

TABLE OF CONTENTS

Acknowledgements iii

Abstract VI

Chapter 1 1

1.1 Introduction 1

1.2 Manholes 2

1.3 Problems With Manholes And Economic Costs 3

1.4 Corrosion 4

1.5 Factors Affecting Corrosion 5

1.6 Need For The Study 6

1.7 Project Goal And Objectives 7

1.8 Hypotheses 7

1.9 Manuscript Organization 8

Chapter 2 9

2.1 Literature Review Organization 9

2.2 Hydrogen Sulfide 9

2.2.1 Mechanism Of Sulfide Generation 10

2.2.2 Mechanism Of Microbially Induced Concrete Corrosion 12

2.3 Factors Affecting The Micc Process 16

2.3.1 Microbial Community 16

2.3.2 Wastewater And Sewer Airspace Characteristics 16

2.3.3 Sewer System Characteristics 21

2.3.4 Interrelationship Among Various Factors 26

2.4 Methods Available To Mitigate Corrosion 28

2.4.1 Corrosion Resistant Manhole Materials 28

2.4.2 Lining The Manhole Walls 29
2.4.3 Use Of Chemical Agents 29
2.4.4 Design Changes In Sewer Systems 31
2.5 Existing Models For Predicting Sulfide And Corrosion In Sewers 33
2.6 Concerns With Currently Available Models 40
2.7 Data Analysis And Predictive Modeling 40
2.8 Goal And Objectives Of The Study 46
Chapter 3 48
3.1 Experimental Design For Data Collection 48
3.1.1 350 Manholes Selection 48
3.1.2 Data Collection 51
3.1.3 Parameter Selection 52
3.1.4 Field Instrument And Lab Protocols 55
3.1.4.1 Weather Conditions 55
3.1.4.2 Gas Phase Instruments 55
3.1.4.2.1 Odalog Sl 1000 Or Odalog Sl 50 (App-Tek International) 55
3.1.4.2.2 Kestrel® Drop™ D2 57
3.1.4.2.3 Toxirae Pro 58
3.1.4.3 Liquid Phase Instruments 59
3.1.4.3.1 Aqua Troll 600 Multiparameter Sonde 59
3.1.4.3.2 Hanna Multiparameter (Ph/Ec/Do) Probe Hi7698194 With Meter Hi98194 61
3.1.4.3.3 Isco 6712 Full-Size Portable Sampler 63
3.1.4.4 Lab Analysis Instruments And Procedure 64
3.1.4.4.1 Lab Sample Preparation 64
3.1.4.4.2 Sample Preservation 65
3.1.4.4.3 Sulfide – Epa Method 9034 Titrimetric Procedure 66
3.1.4.4.3.1 Iodometric Titration 67
3.1.4.4.3.2 Standardization Of Iodine 67

3.1.4.4.4 Sulfate – Turbidimetric Method 10227 68
3.1.4.4.4.1 Accuracy Check Of Spectrophotometer 69
3.1.4.4.5 Biological Oxygen Demand – Epa Standard Method 5210b 70
3.1.4.4.5.1 Bod 5210b Procedure 70
3.1.4.4.5.2 Bod Ldo Probe – Model Lbod10101 72
3.1.4.4.5.3 Hqd Portable Meter 72
3.1.4.5 Depth Of Corrosion Measurement 73
3.1.5 Qa/Qc For The Instruments 76
3.1.5.1 Gas Phase Instruments Qa/Qc 76
3.1.5.2 Liquid Phase Instruments Qa/Qc 77
3.2 Data Analysis 79
3.2.1 Multiple Linear Regression 80
Chapter 4 83
4.1 Data Collection Results 83
4.2 Variables And Data Selection For Hydrogen Sulfide And Corrosion Rate Model-Building 85
4.2.1 Variables Not Included In Model-Building 88
4.3 Hydrogen Sulfide Generation And Factors 90
4.3.1 Variation Of Hydrogen Sulfide With Depth 91
4.3.2 Examples Illustrating Expected Effect Of Manhole Design On Hydrogen Sulfide 92
4.3.2.1 Example Effect Of Drop On Gas-Phase Hydrogen Sulfide 93
4.3.2.2 Example Effect Of Pipe Size Change On Gas-Phase Hydrogen Sulfide 94
4.3.2.3 Example Effect Of Angle Of Flow/Multiple Inlets On Gas-Phase Hydrogen Sulfide: 96
4.3.2.4 Example Effect Of Presence Of Guide On Gas-Phase Hydrogen Sulfide 98
4.3.2.5 Example Effect Of Flow Type/Turbulence On Gas-Phase Hydrogen Sulfide 98
4.3.2.6 Example Effect Of Location Of Manhole On Gas-Phase Hydrogen Sulfide 100
4.3.2.7 Summary Of Average Gas-Phase Hydrogen Sulfide By Manhole Category 103
4.3.3 Examples Illustrating Effects Of Wastewater Characteristics On Hydrogen Sulfide 104
4.3.3.1 Example Effects Of Temperature On Hydrogen Sulfide 106

4.3.3.2 Example Effects Of Ph On Hydrogen Sulfide 108
4.3.3.2 Example Effects Of Dissolved Oxygen On Hydrogen Sulfide 109
4.3.4 Effects Of Sulfide, Sulfate, And Bod On Hydrogen Sulfide 110
4.3.5 Effect Of Seasons On Hydrogen Sulfide 112
4.4 Effect Of Manhole Design On Sulfide, Sulfate, Bod, And Dissolved Oxygen, And Correlation Among The Parameters 116
4.5 Gas- And_Liquid-Phase Data Summary For Complete Manhole Data Sets 118
4.6 Corrosion Rate And Factors 122
4.6.1 Corrosion Rate Determination 122
4.6.2 Corrosion Rate Factors 125
4.6.3 Effect Of Manhole Design On Corrosion Rate 126
4.6.4 Effect Of Gas And Liquid Parameters On Corrosion Rate 130
Chapter 5 135
5.1 Model Development Process 135
5.2 Hydrogen Sulfide Generation And Volatilization Model 136
5.2.1 Preliminary Analysis 138
5.2.1.1 Correlation Matrix 138
5.2.1.2 Fitting The Preliminary Model 139
5.2.1.3 Preliminary Model – Assumptions Check 142
5.2.2 Transformed Model 143
5.2.2.1 Transformed Model – Assumptions Check 148
5.2.3 Selection Of Significant Variables 153
5.2.4 Selection Of Interaction Terms 156
5.2.5 Model Search 160
5.2.6 Model Selection 166
5.2.6.1 M1 Model 166
5.2.6.1.1 M1 Model Assumptions 168
5.2.6.1.2 M1 Model Validation 174

5.2.6.2 M2 Model 174
5.2.6.2.1 M2 Model Assumptions 176
5.2.6.2.2 M2 Model Validation 184
5.2.6.3 M1 And M2 Model Comparison 184
5.2.7 Model Interpretation 186
5.3 CORROSION MODEL 189
5.3.1 Preliminary Analysis 191
5.3.1.1 Correlation Matrix 191
5.3.1.2 Fitting The Preliminary Model 192
5.3.1.3 Preliminary Model – Assumptions Check 195
5.3.2 Transformed Model 195
5.3.2.1 Transformed Model – Assumptions Check 202
5.3.3 Selection Of Significant Variables 207
5.3.4 Selection Of Interaction Terms 209
5.3.5 Model Search 212
5.3.6 Model Selection 216
5.3.6.1 M1 Model 216
5.3.6.1.1 M1 Model Assumptions 217
5.3.6.1.2 M1 Model Validation 223
5.3.6.2 M2 Model 224
5.3.6.2.1 M2 Model Assumptions 225
5.3.6.2.2 M2 Model Validation 234
5.3.6.3 M1 And M2 Model Comparison 234
5.3.7 Model Interpretation 236
Chapter 6 240
6.1 FINDINGS AND CONCLUSIONS 240
6.1.1 Data Characterization 240
6.1.2 Model For Hydrogen Sulfide Generation And Volatilization 241

6.1.3 Model For Corrosion Rate 243
6.1.4 Overall Importance Of Study 244
6.2 RECOMMENDATIONS FOR FUTURE RESEARCH 245
6.2.1 Consider Additional Factors To Explain Variability In Hydrogen Sulfide Concentrations And Corrosion Rates 245
6.2.2 Improve The Accuracy Of Corrosion Rate Measurements 246
6.2.3 Additional Field And Lab Data Collection 247
6.2.3.1 Manhole Design With Limited Number Of Manholes 247
6.2.3.2. Velocity Data 247
6.3.2.3 Precipitation Data 247
6.2.3.3 Deeper Manholes 248
References 249
Appendix A 254
Appendix B 257
Biographical Information 273

LIST OF FIGURES

Fig 1.1 Manhole Components 2
Fig 2.1 Process of sulfide generation and utilization in the presence of sufficient oxygen 11
Fig 2.2 Process of sulfide generation and release into wastewater in absence of sufficient oxygen 11
Fig 2.3 Mechanism of microbially induced sulfide corrosion 15
Fig 2.4 Effect of MICC on concrete pH 15
Fig 2.5 Effect of velocity on hydrogen sulfide generation 22
Fig 2.6 Shows daily and weekly flowrates variations in a sewer system 23
Fig 2.7 Effect of turbulence on hydrogen sulfide generation 25
Fig 2.8 Shows nomograph A and B 36
Fig 2.9 Flow and Slope relationship 37
Fig 3.1 ISCO and gas and liquid phase instruments installation in field for 48 hours of data collection 51
Fig 3.2 Collected wastewater samples for lab analysis 52
Fig 3.3 OdaLog SL 50 (App-Tek International) in the manhole and OdaLog with the IrDA 57
Fig 3.4 Kestrel Drop for Temperature and Relative humidity measurement 58
Fig 3.5 ToxiRAE Pro Oxygen sensor 59
Fig 3.6 Aqua TROLL 600 Multiparameter Sonde 61
Fig 3.7 Hanna Multiparameter probe HI7698194 with the Meter HI98194 62
Fig 3.8 ISCO 6712 Full-Size Portable Sampler with 24 bottles of 500 ml volume and Adjustable Distributer Arm . 64
Fig 3.9 Prepared Preservative for sulfide sample preservation 66
Fig 3.10 Vacuum filtration of the sample and iodometric method for sulfide determination 68
Fig 3.11 HACH 10227 method for sulfate determination 69
Fig 3.12 BOD reagent preparation and 0th Day and 5th Day BOD determination 73
Fig 3.13 Corrosion rate determination: Lower 20% most corroded 74
Fig 3.14 CleverScan being used design determination 75
Fig 3.15 CleverScan being used in the field for manhole diameter measurement and the point cloud 76
Fig 4.1 Average oxygen concentration for 31 manholes, along with hydrogen sulfide concentrations 89

Fig 4.2 Average relative humidity for 118 manholes, along with hydrogen sulfide concentrations 90
Fig 4.3 Histogram for average gas phase hydrogen sulfide measured in 146 manholes 91
Fig 4.4 Depth effect on temperature and average hydrogen sulfide concentration..... 92
Fig 4.5 Hydrogen sulfide concentration for example manholes with drops of different heights..... 94
Fig 4.6 Hydrogen sulfide concentration for varied pipe size changes in manholes..... 95
Fig 4.7a Hydrogen sulfide concentration for varied bends and flow direction..... 97
Fig 4.7b Hydrogen sulfide concentration for varied bends and flow direction 97
Fig 4.8 Hydrogen sulfide concentration for guided/non-guided manholes..... 98
Fig 4.9 Hydrogen sulfide concentration for varied flow types in manholes..... 100
Fig 4.10a Average hydrogen sulfide distribution for 146 manholes..... 101
Fig 4.10b Maximum hydrogen sulfide distribution for 146 manholes 101
Fig 4.11 Hydrogen sulfide concentration depending on location of manholes..... 102
Fig 4.12 Average hydrogen sulfide concentration depending on manhole design 103
Fig 4.13 Effects of gas and liquid-phase parameters on hydrogen sulfide for manhole with supercritical flow (D14MH0192)..... 105
Fig 4.14 Effects of gas and liquid-phase parameters on hydrogen sulfide for manhole with trends for manhole with non-guided channel (E07MH0111)..... 105
Fig 4.15 Effect of gas and liquid-phase parameters on hydrogen sulfide for manhole near lift station (B12MH0026) 106
Fig 4.16a and b Average hydrogen sulfide in ppm vs average gas phase temperature in °F..... 107
Fig 4.16c and d Average hydrogen sulfide in ppm vs average liquid phase temperature in °F..... 108
Fig 4.17a and b Average hydrogen sulfide in ppm vs average liquid phase pH..... 109
Fig 4.18a and b Average hydrogen sulfide in ppm vs average liquid DO in mg/l..... 110
Fig 4.18c Average dissolved oxygen for various manhole categories..... 110
Fig 4.19a and b Average hydrogen sulfide vs. Average sulfide 111
Fig 4.19c and d Average hydrogen sulfide vs. Average sulfate..... 112
Fig 4.19e and f Average hydrogen sulfide vs. Average BOD 112
Fig 4.20 Average hydrogen sulfide generated every season..... 114

Fig 4.21 Average hydrogen sulfide generated every season 115
Fig 4.22 Average sulfide, sulfate, and BOD by manhole category 116
Fig 4.23 Average sulfate vs. average sulfide 118
Fig 4.24 Average sulfate vs. average DO 118
Fig 4.25 Average BOD vs. average sulfate 118
Fig 4.26 Average BOD vs. average sulfide 118
Fig 4.27 Manhole point cloud showing portion used for corrosion rate determination* 122
Fig 4.28 Rejecting data 10% less than the maximum dx and dy from top of manhole shape file 123
Fig 4.29a F04MH0026: Gas phase temp, H_2S vs depth 124
Fig 4.29b F04MH0026: Gas phase temp, H_2S vs depth% 124
Fig 4.29c G12MH0226: Gas phase temp, H_2S vs depth 125
Fig 4.29d G12MH0226: Gas phase temp, H_2S vs depth% 125
Fig 4.29e H09MH0411: Gas phase temp, H_2S vs depth 125
Fig 4.29f H09MH0411: Gas phase temp, H_2S vs depth% 125
Fig 4.30 Histograms for average corrosion rate along dx and dy in mm/y 126
Fig 4.31a Effect of manhole design on average corrosion rate along dx 127
Fig 4.31b Effect of manhole design on average corrosion rate along dy 128
Fig 4.31c Average hydrogen sulfide and corrosion rates for various designs 128
Fig 4.32 Corrosion rate vs gas and liquid phase parameters along the dx and dy axis 133
Fig 5.1 Residuals plots for preliminary hydrogen sulfide model 143
Fig 5.2 Residuals plots for SQRT(Y) transformation of hydrogen sulfide model 144
Fig 5.3 Residuals plots for $LOG_{10}(0.01+Y)$ transformation of hydrogen sulfide model 145
Fig 5.4 Predictor vs. Residuals plots for $LOG_{10}(0.01+Y)$ transformation, hydrogen sulfide model 150
Fig 5.5 Partial regression plots for considered interactions, transformed hydrogen sulfide model 159
Fig 5.6 Normality plot, residual plot and time plot for M1 for hydrogen sulfide 169
Fig 5.7 Normality plot, residual plot and time plot for M2 model for hydrogen sulfide 177
Fig 5.8 Residuals plots for the preliminary model 195
Fig 5.9 Residuals vs predictor variables 196

Fig 5.10 Residuals vs transformed predictor variables 198

Fig 5.11 Residuals plots for $LOG_{10}(0.01+Y)$ transformation 199

Fig 5.12 Partial regression plots for considered interactions 211

Fig 5.13 Normality plot, Residual plot and time plot 218

Fig 5.14 Normality plot, Residual plot and time plot 227

LIST OF TABLES

Table 1.1 Inspection frequency for manhole 4

Table 2.1 Solubility of H_2S at different temperatures 18

Table 2.2 pH effect on H_2S and solubility 19

Table 2.3 Minimum slope requirements 24

Table 2.4 Interrelationship between the various factors 26

Table 2.5 Cost of lining the manholes 29

Table 3.1 Manhole selection 50

Table 3.2 Parameters measured and method of measurement 53

Table 3.3 Sensor summary for Odalog SL 50 and Odalog SL 1000[2] 56

Table 3.4 ToxiRAE Pro specifications[4] 58

Table 3.5 Sensor summary for Aqua TROLL[5] 60

Table 3.6 Sensor Summary for Hanna[6] 62

Table 3.7 ISCO 6712 specifications[7] 64

Table 3.8 Sample preparation 65

Table 3.9 Soluble and insoluble sulfide determination procedure[9,13] 67

Table 3.10 TNTplus 864 and TNTplus 865 procedure[17,18] 69

Table 3.11 Minimum sample volume requirements for BOD[20] 71

Table 3.12 Gas phase instruments QA/QC 76

Table 3.13 Liquid phase instruments QA/QC 77

Table 3.14 Predictor and response variables for multiple linear regression 80

Table 4.1 Manholes sampled by season 83

Table 4.2 Target and Actual Manholes Sampled in Each Design Category 84

Table 4.3 Dataset selection for hydrogen sulfide model-building 86

Table 4.4 Manhole data sets for developing hydrogen sulfide and corrosion rate models, by design category 87

Table 4.5 Average oxygen and relative humidity values for manholes for which they were measured 90

Table 4.6 Distribution of manholes based on seasons 113

Table 4.7 Average gas and liquid phase parameters based on seasons 113
Table 4.8 Gas and liquid temperatures for 146 manholes based on seasons, °F 115
Table 4.9 Gas and Liquid Parameter Summary for 146 manholes 120
Table 5.1 Predictor variables in the hydrogen sulfide model 137
Table 5.2 Coefficient estimates for the predictor variables – preliminary hydrogen sulfide model 139
Table 5.3 ANOVA for preliminary hydrogen sulfide model 141
Table 5.4 Coefficient estimates for the predictor variables – transformed hydrogen sulfide model 145
Table 5.5 ANOVA for transformed hydrogen sulfide model 147
Table 5.6.1 F-Test for equality of variances, transformed hydrogen sulfide model 149
Table 5.6.2 T-Test for error variances, transformed hydrogen sulfide model 149
Table 5.7 X- Outliers, transformed hydrogen sulfide model 151
Table 5.8 Output of stepwise model, transformed hydrogen sulfide model 154
Table 5.9 Selected predictor variables, transformed hydrogen sulfide model 155
Table 5.10 Selected interactions, transformed hydrogen sulfide model 157
Table 5.11 Coefficients for the model with interactions, transformed hydrogen sulfide model 159
Table 5.12 Output of stepwise model – model search, transformed hydrogen sulfide 161
Table 5.13 Goodness-of-fit metrics of candidate models for transformed hydrogen sulfide, selected using best subset method* 162
Table 5.14 Candidate models for transformed hydrogen sulfide model, selected using best subset method 163
Table 5.15 Coefficients, summary, and ANOVA for M1 model for hydrogen sulfide 167
Table 5.16.1 F-Test for equality of variances, M1 hydrogen sulfide model 170
Table 5.16.2 T-Test for error variances, M1 hydrogen sulfide model 170
Table 5.17 X- Outliers for M1 model for hydrogen sulfide 171
Table 5.18 Coefficients, summary, and ANOVA for M2 model for hydrogen sulfide 175
Table 5.19.1 F-Test for equality of variances for the M2 model for hydrogen sulfide 178
Table 5.19.2 T-Test for error variances for M2 model for hydrogen sulfide 178
Table 5.20 X- Outliers for M2 model for hydrogen sulfide 181
Table 5.21 M1 and M2 hydrogen sulfide model comparison 185

Table 5.22 Predictor variables in the corrosion model 189
Table 5.23 Estimates of coefficients for the predictor variables – preliminary corrosion rate model 192
Table 5.24 ANOVA for preliminary corrosion rate model 193
Table 5.25 Estimates of coefficients for the predictor variables – transformed corrosion rate model 199
Table 5.26 ANOVA for transformed corrosion rate model 201
Table 5.27.1 F-Test for equality of variances, transformed corrosion rate model 203
Table 5.27.2 T-Test for error variances, transformed corrosion model 203
Table 5.28 X- Outliers, transformed corrosion rate model 205
Table 5.29 Output of stepwise model, transformed corrosion rate model 207
Table 5.30 Selected predictor variables, transformed corrosion model 208
Table 5.31 Selected interactions, transformed corrosion model 210
Table 5.32 Coefficients for the model with interactions, transformed corrosion rate model 211
Table 5.33 Output of stepwise model – model search, transformed corrosion rate model 213
Table 5.34 Goodness-of-fit metrics for candidate models for transformed corrosion rate model, selected using best subset method* 214
Table 5.35 Candidate models for transformed corrosion model, selected using best subset method 214
Table 5.36 Coefficients, summary, and ANOVA for M1 model 217
Table 5.37.1 F-Test for equality of variances, M1 corrosion rate model 219
Table 5.37.1 T-Test for error variances, M1 corrosion rate model 219
Table 5.38 X- Outlier for M1 model for corrosion rate 221
Table 5.39 Coefficients, summary, and ANOVA for M2 model for corrosion rate 224
Table 5.40.1 F-Test for equality of variances for the M2 model for corrosion rate 228
Table 5.40.2 T-Test for error variances for M2 model for corrosion rate 228
Table 5.41 X-outliers for M2 model for corrosion model 230
Table 5.42 DFBETAS for M2 corrosion model 233
Table 5.43 Model comparison 235

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

According to EPA[38], there are approximately 16,000 sewer systems, and 12 million manhole structures in US serving 190 million people. However, manholes suffer from microbially induced concrete corrosion (MICC) due to the release of hydrogen sulfide gas and its oxidation to sulfuric acid which then corrodes the sewer systems. The problem of MICC became a bigger issue when government authorities started imposing limits on the nature of wastewater released into sewers, which resulted in low concentrations of toxic metals in the sewer system and accelerated growth of bacterial and fungal species[41]. The MICC problem is associated with increased urbanization which promotes use of hot water and the discharge of household detergents containing sulfur[45].

Corrosion can reduce the 50 to 100-year life of the infrastructure to less than 10 years[40]. In has been observed that 31 months of exposure of new concrete to sewer environments results in 8.5 mm of corrosion product build-up and almost 24 mm of concrete loss.[42] Concrete which has already undergone severe corrosion showed 12 mm of corrosion product build-up and 28.4 mm of concrete loss in 31 months[42].

Extensive corrosion can result in infrastructure damage, sometimes leading to catastrophic failure such as street collapses or sewer blockage[40]. MICC not only causes damage to infrastructure, but also results in economic losses, and poses a safety hazard. In the US, $36 billion is spent annually in managing sewer corrosion, including both pipes and manhole shafts[38]. Globally MICC processes cost billions of dollars in repairs and maintenance. In Germany estimated repair for all private and public sewage system is $100 billion[53]; in the UK the replacement costs for sewer mains is estimated to be £104 billion; and in Belgium the cost of sewer corrosion has been estimated at £4 million per year[45]. In Australia, AUS $40 million is spent annually on a rehabilitation program focusing on pipe corrosion[45]. Thus, there arises a need to identify ways to reduce damages and economic losses incurred due to microbially induced concrete corrosion in manholes.

1.2 MANHOLES

The name “manhole” was adopted because it basically defined a hole which allowed access to humans for maintenance and to conduct repair in sewers[35]. However, with developments in infrastructure, manholes connecting to the sewer pipes are building blocks of a wastewater collection and transportation system[26, 35]; also they provide access to the sewer systems which carry the wastewater[27, 28, 35]. The access points help the operators to conduct inspections, make modifications, and carry out cleaning and maintenance[27, 28, 36]. A manhole is made up of a vertical circular pipe called a chamber or the ring, which leads to the sewer systems, thus providing access for inspections,[27] They can be of varying sizes and depths depending on the requirement[27]. The Texas Commission on Environmental Quality (TCEQ) requires that the inside diameter of the manhole must be greater or equal to 48in[31]. Placement of the manholes is usually 0.5 m away from the curb lines and is determined on the following basis[27, 28, 30, 31]:

a. If there is a change in the flow of the wastewater,
b. If there is a change in the gradient of the utilities, or
c. If a particular area requires future access for inspections.

Various components of the manhole include cover, chamber, base, channel, stub and rocker pipes (Fig. 1.1) [27, 28, 29].

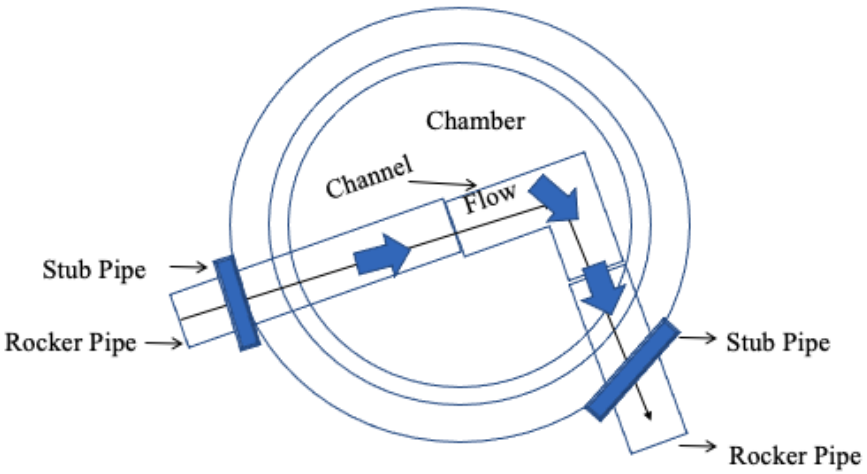


Fig 1.1 Manhole Components[29]

Manholes are usually classified based on their depth[28, 36]. Shallow manholes have a depth of 75 to 90 cm and are constructed usually in areas of less traffic[28, 36]. Manholes with a depth of 90 cm to 150 cm are called normal manholes, and those with depth greater than 150 cm are called deep manholes[28, 36].

1.3 PROBLEMS WITH MANHOLES AND ECONOMIC COSTS

Manholes face corrosion problems similar to sewers; in addition, they experience easy damage because they are not protected by several feet of earth like the sewer lines[38]. Manholes are excessively exposed to traffic loadings, and to surface climatic and environmental impacts[38]. The common problems faced by manholes are listed below:

a. Explosion or severe corrosion due to discharge of uncontrolled industrial wastes[30].
b. Odors and corrosion due to hydrogen sulfide gas[30].
c. Corrosion resulting in collapse of the sewers and manholes[30].
d. Corroded manholes are a source of inflow and infiltration, thus causing increased wastewater flow, and a need for bigger pipes and pumps/lift stations, thus adding to costs[30]. Infiltration can also be caused due to the fracturing of the upper portion of the manhole due to frost in cold climates[38].
e. Corroded manholes may leak, thus allowing soil to enter from the surrounding ground, resulting in surface settlement adjacent to manholes[38].

Of all the concerns, manhole corrosion results in the most extensive damage. Excessive and long-term exposure of wastewater systems to hydrogen sulfide (H_2S) causes deterioration, reduces life of the infrastructure and has economic implications[33, 37]. Also, H_2S is toxic and a leading cause of death for personnel working in sanitary sewers[33]. As of 2007, there were about 20 million manholes in the United States; thus, it is important to understand how to maintain, rehabilitate, restore, and replace these manholes for them to be economically cost effective[35]. Regular inspections play a role in improving their life and reducing the maintenance costs[35]. However, regular inspections require worker access into the manhole, which could be dangerous due to the toxic environment in the manhole[35]. Hughes suggests the frequency for manhole inspections in Table 1.1.[35]

Table 1.1 Inspection frequency for manhole[35]

Manhole Condition	Inspection Frequency (years)
General	5-10, 15 maximum
Corrosion or other maintenance problems	1-2
Designated critical sewers	1-3
Creek and stream locations	1-2
New or rehabilitated manholes[A]	1-2

[A]One year, then adjust appropriately: For a 10-25 year repair life, inspect every 2-5 years; For a >50 year repair life, inspect every 10-15 years.

As is clear from the above table, inspection for corrosion problems has to be done every 1 or 2 years; this would lead to workers been exposed to animals, pests, mechanical and electrical hazards, blood-borne pathogens, falling of objects, toxic manhole or sewer air, entrapment and engulfment while working[35]. To prevent danger to human life, proper protective clothing, gloves, eye protection, and air purifying respirators need to be used, and this all adds to the economic cost of maintaining a manhole. In addition, to rehabilitate the corroded manholes, liners such as sealant, epoxies, coating, chemical grout, and poly-urea based linings are used, which further adds to the cost[36]. Since inspections and then the process of maintaining and replacing manholes requires millions of dollars and poses danger to the workers as well, it becomes important to evaluate the conditions in the manhole and prioritize manholes in need of repairs without human access, thus improving the life of the manholes and also reducing maintenance and replacement costs.

1.4 CORROSION

Corrosion occurs by the release of the hydrogen sulfide from the wastewater surface to the atmosphere of the sewer or manhole, where it is oxidized to form sulfuric acid[33]. The sulfuric acid attaches to the walls of the sewer systems, leading to corrosion[33] in a process called microbially induced concrete corrosion (MICC). The process of oxidation of sulfur to sulfuric acid is usually carried out by the *Thiobacillus concretivorus* bacteria[33]. In the manhole shafts, the bacterial layers can proliferate rapidly and form a corrosive slime layer which can quickly

weaken and decompose the structures; under some conditions total structural degradation can occur in less than five years[35]. Some studies have shown that the rate of corrosion is not uniform but accelerates over time[35]. The corrosion of the sewer systems results in equipment failure and in adverse cases street collapses, thus necessitating premature replacement or rehabilitation and adding to expenditures[37].

1.5 FACTORS AFFECTING CORROSION

The corrosion process is controlled by various factors. The factors, which can be classified into 3 broad categories, work together to carry out the MICC process. The 3 major categories of factors involved are:

a. Microbial community

i. Aerobic bacteria: Aerobic bacteria oxidize sulfide in wastewater thus controlling hydrogen sulfide concentrations; hence limiting hydrogen sulfide available for volatilization to cause corrosion. Thus, reduced aerobes activity would result in higher sulfide concentration in wastewater[37].

ii. Anerobic bacteria: Anaerobic bacteria use chemically bound oxygen for respiration. When anaerobic bacteria use oxygen that is chemically bound in sulfate, they produce hydrogen sulfide and other reduced sulfides[37].

iii. Concrete-corroding bacteria: *Thiobacillus* and other acid tolerant bacteria on concrete walls oxidize hydrogen sulfide to sulfuric acid, which causes corrosion[37].

b. Wastewater and sewer airspace characteristics

i. Sulfur compounds: High concentrations of sulfur compounds increase sulfide generation under appropriate conditions[34].

ii. Temperature: High wastewater temperature increases volatilization of hydrogen sulfide, as well as increases the rates of microbial and chemical reactions[35].

iii. Suspended solids: The presence of suspended solids would increase the biochemical oxygen demand (BOD), and lower oxygen concentrations[33].

iv. Biological oxygen demand: High BOD results in oxygen consumption, thus creating anaerobic conditions[35].

v. pH: Acidity and alkalinity of the wastewater control the concentration and solubility of the hydrogen sulfide[33].

vi. Dissolved oxygen: High dissolved oxygen concentration of wastewater oxidizes sulfide to thiosulfate, thus reducing sulfide concentrations, and also inhibits anaerobic conditions in the wastewater[33].

vii. Humidity: High humidity results in faster biological growth[39].

c. Sewer system characteristics

i. Flow velocity: Low flow velocity of the wastewater prevents wastewater aeration, thus supporting the development of anaerobic conditions[33].

ii. Slope: Lesser sewer slope will result in lower flow velocities which inhibit wastewater aeration[33].

iii. Turbulence: Low turbulence can prevent aeration of wastewater, but high turbulence causes volatilization of hydrogen sulfide[33].

iv. Pump stations: Stagnant water in pump stations is prone to anaerobic conditions[33].

v. Organic acids: The presence of organic solids results in faster oxygen consumption[33].

vi. Grit and debris: Presence of grit and debris results in slowing the wastewater flow velocity[34], thus preventing wastewater aeration. Grit can trap organic solids which can exert an oxygen demand[34].

1.6 NEED FOR THE STUDY

As is clear from the above discussion, various factors work together to bring about corrosion in manholes. Though various equations and models are available to reflect conditions in a sewer pipe, limited focus has been given to manholes. Also, sewer pipe models cannot be employed to predict corrosion rates in manholes, as sewer design is different from that of manholes. In addition, manholes are points of intersection for various sewer pipes, resulting in hydrogen sulfide release from wastewater due to turbulence, thus accelerating corrosion process[33]. A sewer pipe is protected by several feet of earth soil, but manholes are very near to ground surface[38]. Since manholes are closer to the ground surface they allow easy diffusion of the oxygen into the manhole airspace which can result in increased corrosion rates, but such a scenario is avoidable in a sewer pipes since they are buried deep into the earth. Also, the nearness of the manholes to the ground surface could mean that environmental changes would greatly impact

corrosion rates in manholes in comparison to sewer pipes. A direct correlation of the sewer corrosion models is thus not advisable for manholes. Also, most of the studies primarily focus on mirroring field conditions in the lab or measure a limited number of field subjects to predict corrosion. The factors controlling corrosion, however, are very difficult to simulate in a lab to mirror field conditions precisely, as they depend on environmental conditions, manholes age, sewer designs, daily flow rates, wastewater quality, location of the manhole in the city, source of the wastewater, and presence of lift stations. Thus, there arises a need to conduct research which will encompass all these factors to determine the rate of hydrogen sulfide generation, and the rate of corrosion in a manhole.

1.7 PROJECT GOAL AND OBJECTIVES

The primary goal of the project is to develop a predictive model to determine hydrogen sulfide generation and corrosion rates in a unsubmerged sections of the manhole depending on design and manhole conditions.

The specific objectives of the projects are:

a. To measure liquid phase parameters (temperature, dissolved oxygen, pH, sulfide, sulfate, and biochemical oxygen demand, velocity), and gas phase parameters (temperature, relative humidity, oxygen, hydrogen sulfide), that are responsible for regulating hydrogen sulfide concentrations for 48 hours, as well the amount of corrosion for 350 manholes in the City of Arlington.
b. To build predictive models for hydrogen sulfide generation and rate of corrosion in manholes, using the manhole physical design and measured liquid and gas phase parameters.

1.8 HYPOTHESES

Our hypothesis is that manhole design, wastewater characteristics, manhole gas phase conditions, and weather play a major role in determining the rate of corrosion. So, manholes having drops, turbulent flow, or bends, should exhibit higher rates of corrosion and should record higher hydrogen sulfide concentrations, in comparison to manholes with no drops, subcritical flow, and no bends. Also, higher wastewater temperatures or low pH should also result in higher gas phase hydrogen sulfide concentrations, and increased rate of corrosion in manholes. In addition, manholes would be expected to register higher hydrogen sulfide concentrations in summer in comparison to winter, and thus greater rates of corrosion. The predictive model will help to determine the rate of corrosion and hydrogen

sulfide concentrations as a function of manhole design, wastewater characteristics, manhole gas phase conditions, and weather.

1.9 MANUSCRIPT ORGANIZATION

Subsequent chapters of the manuscript are organized as follows:

a. Chapter 2 describes previous studies that were conducted to evaluate corrosion in sewer systems and build predictive models. The chapter will also discuss the process of corrosion, and factors affecting the process in detail.
b. Chapter 3 discusses the methodology, and the process adopted to carry out field and lab studies. The chapter will also discuss the process of data collection and instrumentation used for field studies.
c. Chapter 4 discusses the relationship between the hydrogen sulfide generation, and corrosion rates in relation to manhole design, and gas and liquid phase parameters. In addition, impacts of seasons on hydrogen sulfide generation and corrosion is also evaluated.
d. Chapter 5 discusses the development of the predictive model for hydrogen sulfide generation and corrosion rate using multiple linear regression.
e. Chapter 6 discusses the summary of results and any future recommendations.

CHAPTER 2

LITERATURE REVIEW

2.1 LITERATURE REVIEW ORGANIZATION

This chapter provides background on the research done to understand microbially induced corrosion process (MICC), and currently available methods to measure and mitigate MICC. The chapter will conclude describing the goals and objectives of the proposed research. Specific sections of the literature review are:

a. Generation of hydrogen sulfide gas in sewer systems and mechanism of microbially induced concrete corrosion.
b. Factors controlling hydrogen sulfide generation and MICC process.
c. Methods currently adopted to mitigate corrosion.
d. Empirical models available to predict corrosion rates in sewer systems.
e. Data analysis and predictive models developed to predict corrosion in sewer systems.
f. Goals and objectives of the proposed research.

2.2 HYDROGEN SULFIDE

Corrosion in manholes is primarily caused due to hydrogen sulfide. Hydrogen sulfide gas is a colorless, flammable, poisonous gas with an odor of rotten eggs[33, 37]. Hydrogen sulfide gas is corrosive to various metals such as iron, steel, and copper, and thus is the leading cause of corrosion in collection systems made of cast or ductile iron[33, 37, 39]. In sewer systems hydrogen sulfide gas results in microbial induced concrete corrosion and is shown to reduce the life of the sewer pipes from 50-100 years to only 10-20 years.[48] The cost of replacement and rehabilitation is about $14 billion per year in the US and is increasing with years as the infrastructure ages[52]. In addition to causing damage to the infrastructure through the process of MICC, hydrogen sulfide gas is toxic to humans and can be smelled by humans in concentrations as low as 0.21 ppb (parts per billion)[33]. Even small concentration of hydrogen sulfide from 20 to 150 ppm (parts per million) has shown to cause eye irritation and irritation of the upper respiratory tract[33], and concentrations greater than 230 to 1000 ppm result in unconsciousness and subsequent death

if proper medical help is not provided[33]. To understand the process of corrosion, we first need to understand how hydrogen sulfide gas in produced in the collection systems.

2.2.1 MECHANISM OF SULFIDE GENERATION

Bacteria carry out oxidation-reduction reactions to produce new cells and to generate energy for metabolic functions[34]. Bacteria prefer to use molecular oxygen as an electron acceptor for these oxidation reduction reactions, but in the cases when molecular oxygen is limiting, bacteria have evolved to use low-energy electron acceptors such as nitrate, sulfate and many more[34]. This process of reducing sulfate to sulfide in the presence of organic substrates and using molecular hydrogen as an electron donor to generate energy is called dissimilatory sulfate reduction[34].

Sulfide generation requires an absence of oxygen[33, 37, 39] because the bacteria *Desulfovibrio*[37] required for dissimilatory sulfate reduction are strict anaerobes[34]; as oxygen is usually present in the headspace of the sewers, sulfide generation occurs in the slime layers which form in the submerged part of the sewer pipes[33, 34, 39]. These slime layers consist of 3 sections: aerobic zone, anaerobic zone, and inert anaerobic zone[33, 34]. The aerobic zone consists of the bacteria which utilize oxygen to oxidize sulfide, the anaerobic zone consists of the sulfate-reducing bacteria which produce sulfide, and the inert anaerobic zone contain bacteria which can generate sulfide but due to lack of nutrient supply are inactive[33]. Also, if sulfate is unable to reach the anaerobic zone, the anerobic zone also contains fermentative bacteria and methanogens which can generate energy by producing hydrogen and simple organic acids which then act as substrates for the sulfate-reducing bacteria[34]; thus sulfate reducing bacteria can survive even in the absence of sulfate. The sulfide generated in the anaerobic zone will be utilized by the bacteria in the aerobic zone in the presence of oxygen, thus ensuring that no sulfide is seen in the wastewater stream[33, 34]. However, this process is highly dependent on the presence of oxygen for the aerobic zone to function[33]. So, if oxygen is absent, then the aerobic zone cannot oxidize the sulfide and we will notice the sulfide in the wastewater stream[33]. Thus, the absence of oxygen will result in sulfide being released into the wastewater, thus initiating the process of corrosion. Fig 2.1 shows the process occurring in the sewers when sufficient oxygen is present and Fig 2.2 when sufficient oxygen is absent.

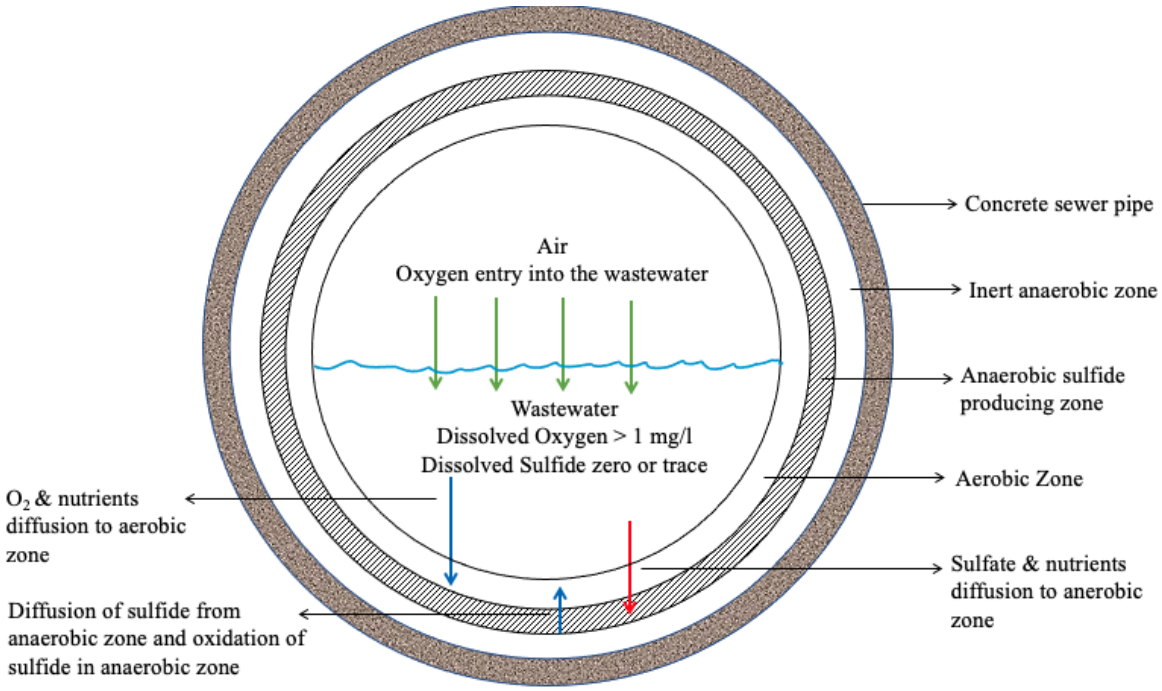


Fig 2.1 Process of sulfide generation and utilization in the presence of sufficient oxygen[33, 34, 37]

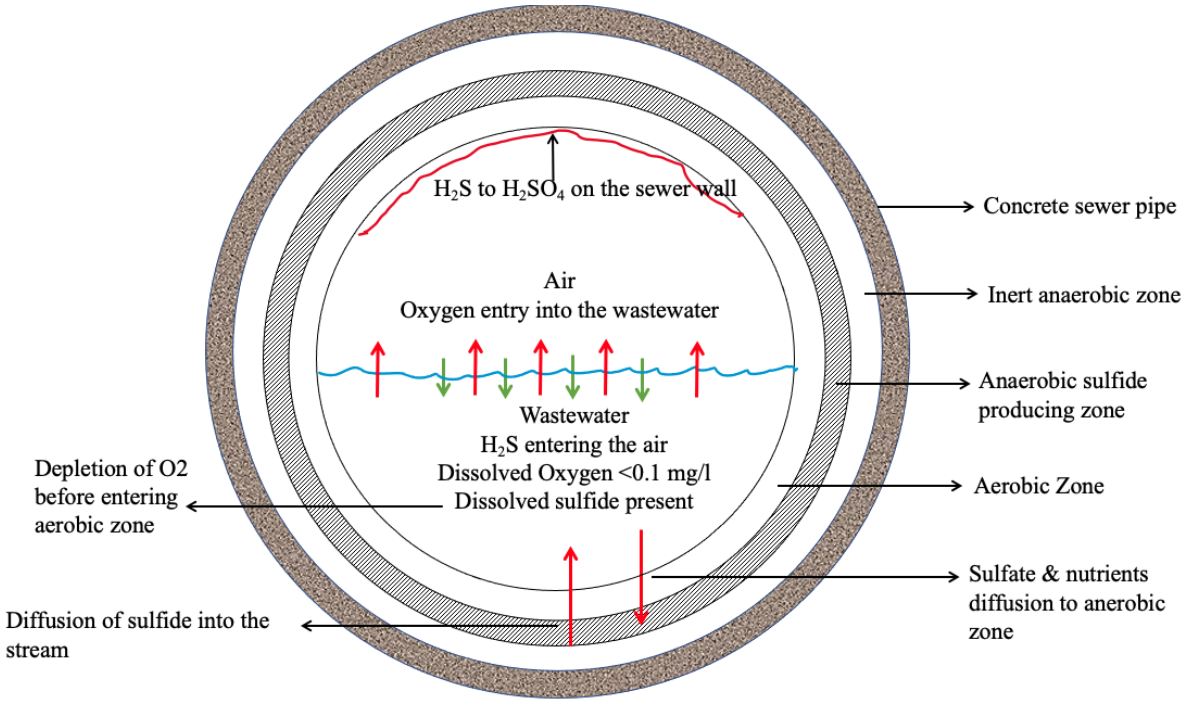


Fig 2.2 Process of sulfide generation and release into wastewater in absence of sufficient oxygen[33, 34, 37]

2.2.2 MECHANISM OF MICROBIALLY INDUCED CONCRETE CORROSION

After sulfide generation in the wastewater, sulfide volatilizes to the headspace in the sewer or manhole, where it oxidizes to form sulfuric acid on concrete walls, thus beginning the process of microbially induced concrete corrosion (MICC). The major reactions involved in sulfide generation and sulfuric acid production causing MICC are described below[37, 39, 41,50]:

Stage I: Abiotic neutralization of concrete surface:

a. pH of concrete around 12-13 and does not support microbial activity[37, 39, 41].

b. Sulfate reducing bacteria (SRB) in wastewater reduce sulfate[41,50].

$$SO_4^{2-} \xrightarrow{reduced} S^{2-}$$ [37, 39, 41] ……Eq. 2.1

$$Organic\ matter \xrightarrow{reduced} CO_2$$ [41] ……Eq. 2.2

c. Depending on the pH, sulfide stays in equilibrium with HS^- and H_2S[37, 39].

$$S^{2-} + 2H^+ \Longrightarrow H_2S \text{ or } S^{2-} + H^+ \Longrightarrow HS^-$$ (Depending on pH) [37, 39] ……Eq. 2.3

d. H_2S and CO_2 at the liquid-gas interface volatize to the sewer headspace, diffuse through the headspace atmosphere, and dissolve into the condensate film on the walls[37, 39, 41].

$$H_2S \Longrightarrow HS^- + H^+$$ [37, 39] ……Eq. 2.4

$CO_2 + H_2O \Rightarrow H_2CO_3$ *and* $H^+ + HCO_3^-$ *and* $2H^+ + CO_3^{2-}$ (in various concentrations) [41] ……Eq. 2.5

e. The weak acids react with calcium hydroxide in concrete and lower the pH to 9[41,50].

f. The duration of the stage I is a few months to years[41] .

Stage II: Colonization by neutrophilic bacteria:

a. After pH falls to 9, neutrophilic sulfur-oxidizing bacteria (NSOM) like *Thiobacillus* will colonize the concrete[41].

b. NSOM oxidize H_2S to H_2SO_4[41,50].

H_2S + moisture on non-submerged surface $\xrightarrow{oxidized\ by\ aerobic\ Thiobacillus\ sp.} H_2SO_4$ [37, 39, 41]……Eq. 2.6

c. Sulfuric acid reacts with concrete to reduce the pH further[41] .

Stage III: Colonization by acidophilic bacteria:

a. After pH falls to 4, acidophilic sulfur oxidizing microorganisms (ASOM) will start to colonize[41] .

b. ASOM oxidize H_2S, thiosulfate, and elemental sulfur to H_2SO_4[41].

c. pH drops to 1-2[37, 39, 41,50].

Stage IV: Loss of concrete mass:

a. H_2SO_4 reacts with silicate and carbonate in concrete to form gypsum[41] .

$$H_2SO_4 + CaO.SiO_2.2H_2O \implies CaSO_4 + Si(OH)_4 + H_2O$$ [41,50]……Eq. 2.7

$$H_2SO_4 + CaCO_3 \implies CaSO_4 + H_2CO_3$$ [41,50]……Eq. 2.8

$$H_2SO_4 + Ca(OH)_2 \implies CaSO_4 + 2H_2O$$ [41,50]……Eq. 2.9

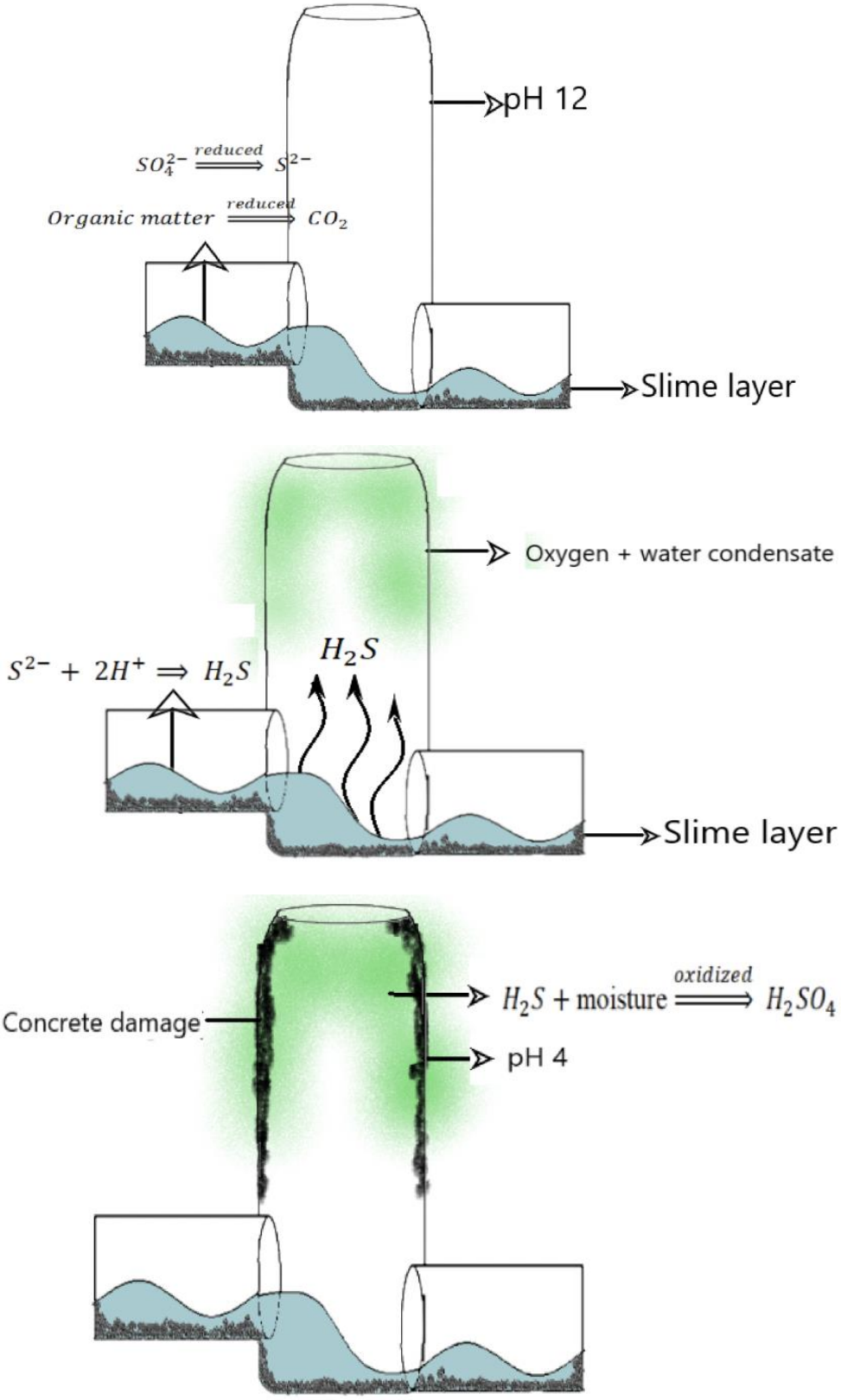
b. Gypsum increases volume by 124%, thus weakening the structure[41].

c. Gypsum can react with tricalcium aluminates to form Ettringite (calcium sulfoaluminate), with expansion volume of 227% to 700%[41].

$$CaSO_4 + 3CaO.Al_2O_3.6H_2O + 25H_2O \implies 3CaO.Al_2O_3.3CaSO_4.31H_2O$$ [41,50]……Eq. 2.9

d. Increase in volume results in concrete expansion and loss of concrete aggregates, thus exposing fresh concrete to be damaged by sulfuric acid[37, 39, 41].

The process of corrosion in sewer systems is illustrated in Fig 2.3. As clear from the above discussion, the microbially induced corrosion results in reducing the pH of the concrete in sewers from 12 to 1. The effect of MICC on pH of the concrete is shown in Fig 2.4.


pH 12
$SO_4^{2-} \xrightarrow{reduced} S^{2-}$
$Organic\ matter \xrightarrow{reduced} CO_2$
Slime layer
Oxygen + water condensate
$S^{2-} + 2H^+ \Rightarrow H_2S$
H_2S
Slime layer
Concrete damage
$H_2S + \text{moisture} \xrightarrow{oxidized} H_2SO_4$
pH 4

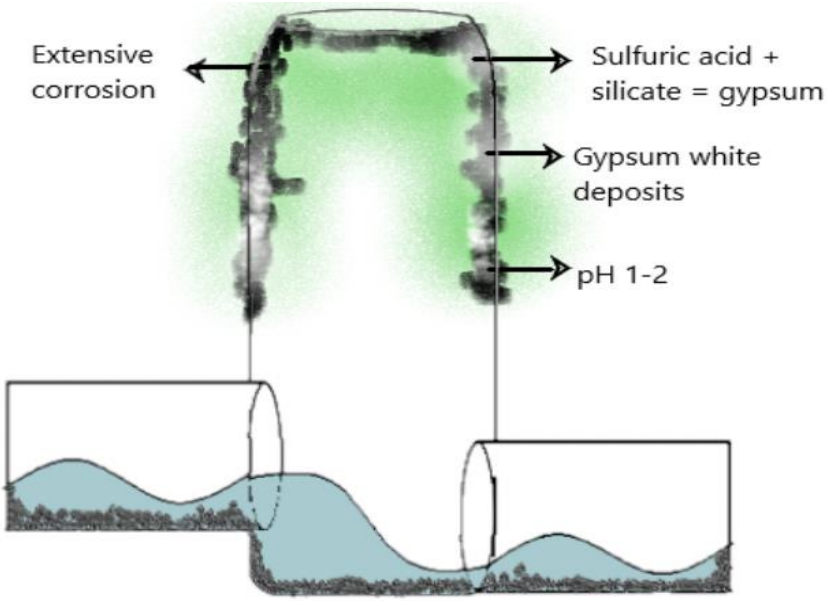


Fig 2.3 Mechanism of microbially induced sulfide corrosion[39]

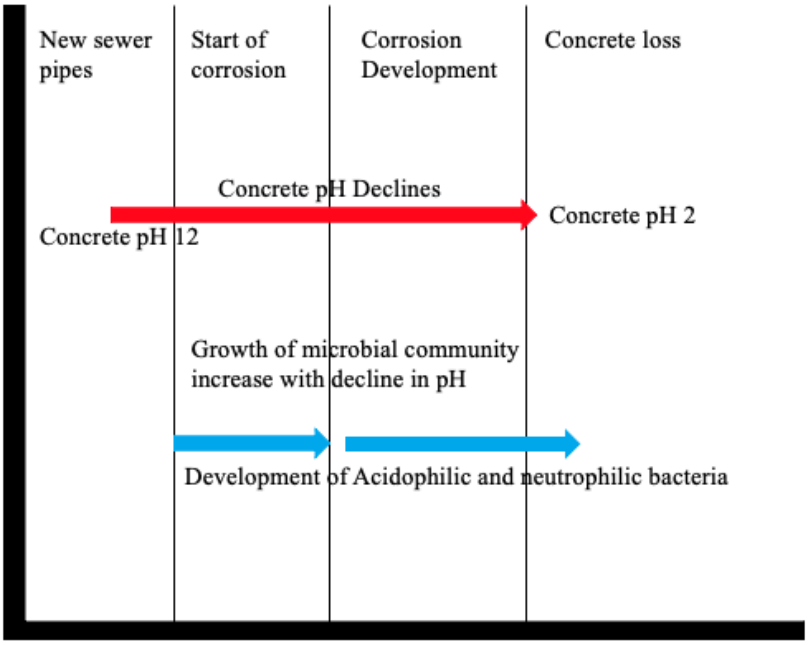


Fig 2.4 Effect of MICC on concrete pH[41]

2.3 FACTORS AFFECTING THE MICC PROCESS

The corrosion process is controlled by various factors. The factors can be classified into 3 broad categories which work together to carry out the MICC process. The 3 major categories of factors involved are:

a. Microbial community
b. Wastewater and sewer airspace characteristics
c. Sewer system characteristics

2.3.1 MICROBIAL COMMUNITY

The microbial communities in the sewer systems differ greatly between the submerged sections of the sewer systems and the unsubmerged sections[37]. Submerged sections are usually populated with aerobes, anaerobes, acid-producing bacteria, and sulfate-reducing bacteria (SRB), whereas the unsubmerged sections, particularly corroded sections, mostly contain aerobes, organic acid producing bacteria, and sulfate reducing bacteria[37].
Submerged anaerobic bacteria and SRB produce hydrogen sulfide from the sulfur compounds present in the wastewater[41]. Increased activity of the anerobic bacteria in the submerged sections results in higher sulfide generation rates, whereas increased activity of the aerobic bacteria results in less sulfide generation.

The unsubmerged sections contain acidophilic sulfur oxidizing bacteria (ASOM) [41] and neutrophilic sulfur oxidizing bacteria (NSOM) [41] such as *Thiobacillus* genus bacteria[37]. Sections where extensive corrosion has resulted in exposure of reinforced steel show the presence of *T. ferrooxidans*, which are iron and sulfur oxidizing bacteria[37]. In addition, acid-tolerant fungi and yeasts are also found in unsubmerged sections[37]. The NSOM and ASOM bacteria are responsible for the oxidation of hydrogen sulfide gas on concrete walls to sulfuric acid, thus causing manhole corrosion[37]. One study reported that *T. ferrooxidans*, and *Acidithiobacillus thiooxidans* comprise 12% - 42% of the total active bacteria present in the concrete samples[46].

2.3.2 WASTEWATER AND SEWER AIRSPACE CHARACTERISTICS

Various factors play a role in determining the amount of hydrogen sulfide generated in the wastewater and rate of concrete corrosion, including concentration of the sulfur compounds, temperature, suspended solids particle size distribution, biological oxygen demand, acidity and alkalinity, dissolved oxygen[33], humidity[39,43].

a. **Presence and concentration of sulfur compounds:** Domestic wastewater usually consists of a substantial amount of sulfate ions, and sulfur compounds which are released into the sewer systems[33, 34]. The sulfate concentrations in municipal wastewater typically lie between 30 mg/L to 250 mg/L[34]. Under anaerobic conditions, these sulfates can be converted to sulfide, which results in corrosion[33, 34]. Also, high concentrations of sulfate facilitate transfer of the sulfate into the slime layer for anaerobic oxidation, thus increasing the potential for sulfide generation[34]. According to EPA, study by County Sanitation Districts of Los Angeles County (CSDLAC) found that even when dissolved sulfide levels are less than 0.1 mg/l, there is still a high concentration of hydrogen sulfide in the sewer atmosphere; thus, corrosion can continue even in low sulfide concentrations[39].

b. **Temperature:** Wastewater is usually higher in temperature in comparison to drinking water[33], and with the increase use of hot water in residential areas, the temperature of wastewater plays an important role in MICC[41]. Temperatures play a role in determining the solubility of the H_2S[33, 35]. As the temperature increases, the solubility of the H_2S decreases because the Henry's law constant decreases with an increase in temperature[33, 34]. Thus, with increase in temperature more H_2S is released to the air space[37]. Table 2.1 shows the effect of temperature on solubility and Henry's law constant for hydrogen sulfide. Also, the solubility of the oxygen reduces with the increase in temperature; thus, less oxygen is available in the wastewater, resulting in more anaerobic conditions in the sewage[34, 37]. Also, temperature effects the concentration of the H_2S levels in the sewers, with H_2S concentration following diurnal cycle with concentrations changing by up to 100 ppm over 24 h period[42]. H_2S levels have also been shown to vary seasonally, with highest during the summer/autumnal months with a daily average > 100 ppm, reducing to ~ 40 ppm during winter[42].

Table 2.1 Solubility of H_2S at different temperatures[33]

Temperature °C	Solubility of H_2S, mg/L	$K_{H2S} \times 10^5$ atm^{-1}
0	7100	373.1
10	5160	272.5
15	4475	236.5
20	3925	207.0
25	3470	183.4
30	3090	164.2
40	2520	134.2
60	1810	97.1

Higher temperatures also increase microbial growth, thus increasing sulfide generation[37]. It has been reported that the rate of sulfide production increases 7% for every Celsius degree increase up to 40°C[39]. The rate of chemical reactions also increases with temperature. In addition, uptake of H_2S by concrete has shown to increase with a rise in the temperature[43]. The H_2S uptake rate rose by 17% for new concrete and 26% for old corroded concrete when the temperature was increased from 25°C to 30°C[43]. The increase in H_2S uptake rate by concrete is due to increased diffusion of H_2S in air and increased chemical and biological sulfide oxidation rates[43]. In addition, temperature increase from 25°C to 30°C has shown to increase the sulfide oxidation rate in the wastewater biofilm by 15%[43], resulting in increased corrosion.

c. **Suspended solids particle size distribution and BOD:** The amount of suspended solids (particle size 0.1 µ to 1 µ) is important because organic matter might be associated with it, and this would change the rate of BOD exertion[33, 34]. This becomes important because a high fraction of soluble BOD has been shown to increase the rate of H_2S generation[33]. If the wastewater has a higher BOD, then it would mean that the waste in the wastewater would consume the available oxygen faster, thus maintaining anerobic conditions and facilitating the production of sulfide and its easy access into the wastewater stream[33, 34, 35, 37].

d. **Acidity and Alkalinity:** The pH determines the chemical reaction and equilibria[33, 34], thus determining the concentration of H_2S and also its solubility in the wastewater[33]. As the pH of the wastewater decreases, so does the solubility of the H_2S[37]; however, the concentration of the H_2S increases because ionization of hydrogen sulfide to HS^- depends on the pH of the solution[33, 34, 39], as is clear from Table 2.2[33, 39]. Thus, at low pH the H_2S can easily volatize into the gas phase from the liquid phase due to the greater proportion of the unionized ions in the wastewater[34, 37], and be oxidized into sulfuric acid by the bacteria belonging to the *Thiobacillus* genus[33] thus increasing the rate of corrosion. The pH of the wastewater is usually around 6 to 7, so the percentage of H_2S and HS^- would tend to be 90.1 and 9.9[33], and 47.7 and 52.3[33] respectively, and such high concentrations of H_2S could increase the risk of corrosion in the manholes thus affecting the integrity of the structures.

Table 2.2 pH effect on H_2S and solubility[33]

pH	H_2S%	HS^-%	Solubility, mg/L
4	99.9	0.1	3470
5	98.9	1.1	3510
6	90.1	9.9	3840
7	47.7	52.3	7270
7.5	22.5	77.5	15,400
8	8.3	91.7	41,800
8.5	2.80	97.20	124,000
9	0.89	99.11	390,000

Even the pH of the concrete surfaces effects the rate of microbial growth and thus effects the rate of corrosion[42, 51]. Thus, concrete pH changes indicate the corrosiveness of the sewer atmosphere and the relative activity of the bacteria[42]. In a study, it was noted that lower pH of older concrete showed faster hydrogen sulfide uptake rates in comparison to new concrete which had higher pH. Also, lower pH of

concrete would support the growth of ASOM, which will faster oxidize the hydrogen sulfide gas to produce sulfuric acid [(43)]. A study conducted by Wells (2014)[(42)] over the course of 31 months found that the pH of new concrete pieces declined from 10.1 to 2.6, whereas old concrete pieces declined from 8.2 to 2.2 when exposed to sewer conditions[(42)], thus meaning that older concrete would exhibit higher sulfide uptake rates and faster corrosion.

e. **Dissolved Oxygen:** Bacteria belonging to the genera *Thiothrix* and *Beggiotoa* can facilitate the reaction of dissolved oxygen with H_2S to form thiosulfate, as shown in equation 2.10 below, thus reducing sulfide concentrations in wastewater[(33)].

$$2HS^- + 2O_2 \rightarrow S_2O_3^{2-} + H_2O \ldots\ldots \text{Eq. 2.10}$$

Also, oxygen concentration will determine how thick is the aerobic zone[(33)]. To prevent the access of the sulfide to the wastewater, the oxygen concentration should generally be 0.1 to 1.0 mg/l[(33, 34)]. In addition, temperature, concentration of organic nutrients, and oxygen together will determine how deep the oxygen can penetrate, because without oxygen the aerobic zone in the slime layer cannot oxidize sulfide[(33)] and prevent its access into the wastewater stream. Also, wastewater with a higher oxygen concentration will take longer to be depleted of oxygen, thus reducing the sulfide generation by inhibiting the creation of anaerobic conditions[(33, 34, 35, 37)].

f. **Humidity:** Humidity will favor the growth rate of both the aerobes and anaerobes present in the sewer system[(39)]. Thus, if a sewer system is devoid of oxygen, humidity will accelerate the growth of anaerobic bacteria, causing an increase in hydrogen sulfide generation, but in the presence of oxygen it will favor growth of aerobic bacteria, thus reducing the hydrogen sulfide generation. Also, moist air increases of hydrogen sulfide uptake rate by concrete through physical adsorption of H_2S and its chemical oxidation in presence of moisture[(43)]. In addition, relative humidity was shown to decrease the pH of the concrete surface when the relative humidity increased from 90% to 100% which is attributed to increased biological sulfide oxidation at higher relative humidity[(47)]. In laboratory study using pre-corroded sewer samples placed in gas phase of the wastewater containing chamber it was observed that, for 100% relative humidity,

the surface pH of the concrete recorded was 1-2 units lower than the concrete exposed to 90% relative humidity[47], thus increasing concrete corrosion by 2 mm in 45 months[47]. In the same study, it was found that the increase of relative humidity from 90% to 100% was also associated with increased sulfate concentrations on the concrete surface, thus indicating increased biological oxidation of hydrogen sulfide gas on the concrete surface[47].

2.3.3 SEWER SYSTEM CHARACTERISTICS

There are many conditions that could affect the 3 zones in the slime layer, thus determining the sulfide access into the wastewater, and its volatilization into the sewer headspace. The sewer system characteristics that are responsible for MICC process are outlined below:

a. **Flow velocities:** Higher flow velocity of the wastewater will incorporate oxygen into the wastewater through aeration; thus, the aerobic zone of the slime layer will able to oxidize the sulfide produced by the anaerobic zone, thereby reducing the sulfide access to the wastewater[34, 35, 37]. High velocities will also cause the oxygen-consuming organic solids to move along with the wastewater flow; thus, less organic solids will accumulate near the pipe surface and deplete oxygen from the wastewater[33]. High velocities keeps the slime layers thin, around 0.01 in (0.25 mm), thus increasing oxygen diffusion into the aerobic zone and keeping the aerobic zone thick[33, 37]. Also, a thin slime layer will produce less sulfide; this small amount of sulfide can be consumed by the thick aerobic zone easily[33, 37]. In contrast, wastewater which is moving really slowly may cause the oxygen to be depleted in slow-flowing area and result in sulfide generation and access into the wastewater stream, even when the bulk water has a good oxygen concentration[33, 37]. Though high velocities may be useful in preventing the access of sulfide in the wastewater, they may cause sulfide to be released to the headspace of the sewer pipes and manhole due to turbulence if sulfide is already present in the wastewater; it can then be easily oxidized to sulfuric acid to cause corrosion[33]. Fig 2.5 shows the effect of flow velocity on the generation and release of hydrogen sulfide gas.

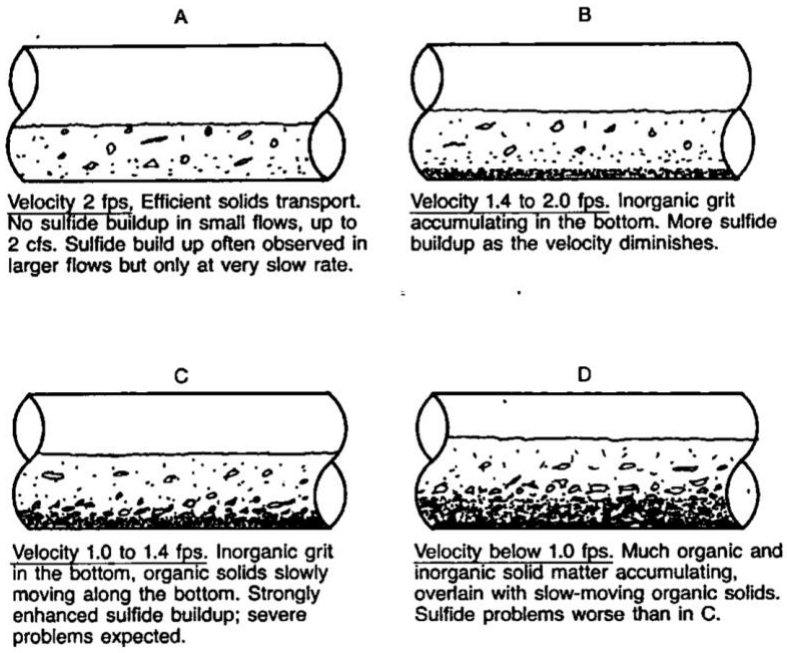


Fig 2.5 Effect of velocity on hydrogen sulfide generation[33]

In addition, the flow velocities in the sewer systems can be affected by the volume of wastewater flowing. It has been observed that wastewater shows diurnal patterns of flow with early morning showing low flows and the flow increasing from 6 am, with the first peak been observed at noon and second between 7 pm and 9 pm[33]. Daily and weekly variations of flow rates are shown in the Fig 2.6 below[33]. Wastewater flow rates in communities with resorts or colleges can also vary seasonally based number of tourists or college students, respectively[33].

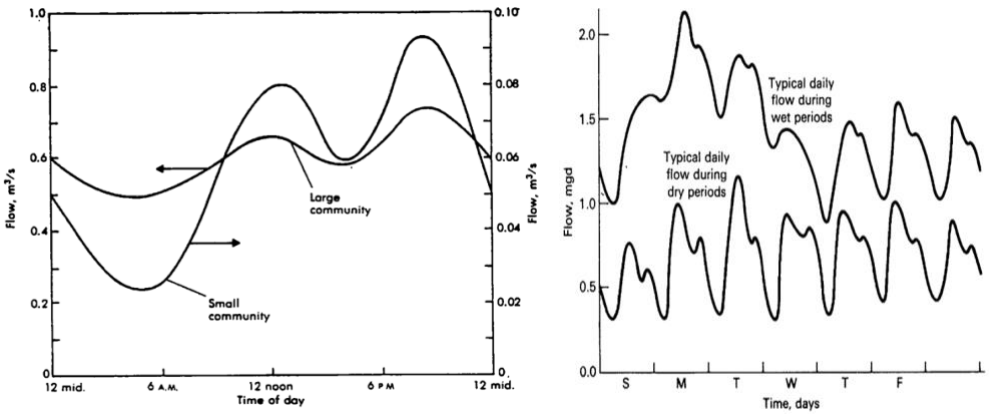


Fig 2.6 Shows daily and weekly flowrates variations in a sewer system[33]

b. **Sewer Slope:** Sewer slope affects the flow velocities, which in turn affects the sulfide production as discussed above. A lesser sewer slope would result in lower flow velocity, thus causing more sulfide generation[33, 37]. To overcome this concern the agencies, recommend particular level of slope for a particular pipe size[33]. However, this can sometimes result in a deep sewer line, thus increasing construction costs, and the need for a lift station[33]. To avoid such a situation, companies sometimes build larger diameter sewer pipes, as the larger the pipe diameter, the lesser the slope needed, but this then results in inadequate flow velocities, which in turn results in higher sulfide generation[33]. Even though flatter larger-sized pipes may generate more sulfide, they may also face less corrosion because they will have less turbulence[33]. However, when the flatter pipe connects to the steeper downstream pipe, any sulfide produced is released due to turbulence, and the downstream manhole and sewer pipe sections show excessive corrosion[33, 37]. Manholes are usually present at the junction of such pipes and experience excessive hydrogen sulfide gas released in such pipe connections; this results in excessive corrosion being observed in such manholes. So, for the construction of slope there is a trade-off between cost and the need to minimize sulfide generation. The table below shows the minimum slope requirements according to TCEQ[31].

Table 2.3 Minimum slope requirements[33]

Pipe Size, inches	Minimum Slope, %	Maximum Slope, %
6	0.50	12.35
8	0.335	8.40
10	0.25	6.23
12	0.20	4.88
15	0.15	3.62
18	0.115	2.83
21	0.095	2.30
24	0.08	1.93
27	0.07	1.65
30	0.06	1.43
33	0.05	1.26
36	0.045	1.12
39	0.04	1.01
>39	Slope determined by Manning's formula to maintain velocity >2ft/s and <10ft/s when flowing full	

c. **Turbulence:** Turbulence can be both beneficial and harmful for the collections systems[33]. Turbulence can result in oxygenation of the wastewater stream, which can then oxidize the sulfide to sulfate, and thus reduce the corrosion probability[33, 34, 37]. But on the other hand, turbulence can cause severe localized sulfide release into the headspace, which can then be oxidized to sulfuric acid and cause corrosion[33] as evidenced by research conducted by EPA[37]. The presence of drops, slopes causing critical flow or hydraulic jumps can all result in a large amount of sulfide release to the headspace[33], and corrosion rates could be 5 times the average rates, as evidenced by an EPA study conducted on a Western trunk sewer in

City of Lakeland in Florida[37]. Fig 2.7 below shows the various forms of turbulence in a sewer pipe and manhole[33].

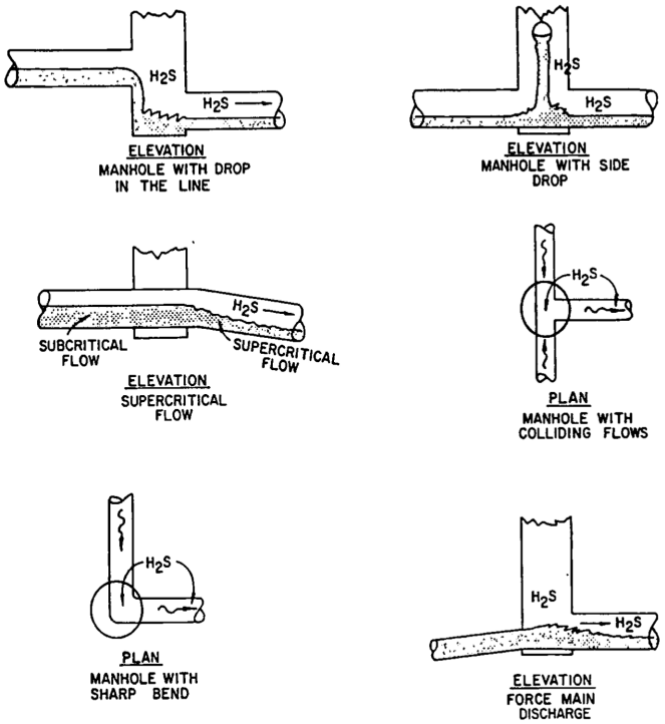


Fig 2.7 Effect of turbulence on hydrogen sulfide generation[33]

d. **Pump Stations:** Pump stations can be a source of sulfide for the downstream collection system. The sulfide gas coming from the influent wastewater tends to collect in pump stations[33, 39]. Also, sulfide may be generated in the pump stations due to the detention of wastewater and solids settling out, resulting in anaerobic conditions with relatively low oxygen to consume the produced sulfide[33, 37]. This accumulated sulfide is then discharged downstream, which can cause excessive corrosion[33].

e. **Organic solids:** The presence of organic solids on the floor of the pipe will result in higher sulfide released into the wastewater stream even when the concentration of oxygen is high[33]. Thus, it is required that gravity sewers maintain a minimum velocity of 0.6 m/s (2 ft/s) to prevent solids deposition from occurring and thus reduce sulfide generation[34].

f. **Grit and debris:** Grit is heavier than organic solids and tends to not move until there is a high flow velocity, so accumulation of grit results in slowing the flow of wastewater, thus causing sulfide generation[33, 34, 37]. Grit can trap organic solids which can exert an oxygen demand thus enabling anaerobic conditions in sewer[34].

2.3.4 INTERRELATIONSHIP AMONG VARIOUS FACTORS

The factors responsible for corrosion discussed above work together to result in the generation of hydrogen sulfide and corrosion of sewer systems. Table 2.4 below shows the interrelationship among the wastewater characteristics, microbial community, and sewer system characteristics.

Table 2.4 Interrelationship between the various factors[40]

Stage	Factor	Reasoning
Stage I - Oxygen Depletion	Velocity of sewage, Sewer slope	Low velocity and lesser sewer slope reduce turbulence and thus reduces aeration
	Detention time	Long detention time results in oxygen depletion
	Temperature	High temperature reduces oxygen solubility and increases bacterial growth rates and chemical reaction rates
	Grit and debris	High grit and debris slow wastewater flow, and reduce aeration
	Dissolved oxygen	High dissolved oxygen will prevent oxygen depletion
	BOD, organic acids	High BOD and organic acids will increase oxygen depletion
After oxygen depletion		

Stage II – Rate of sulfide generation	Organic materials and nutrients	High concentrations of organic materials and nutrients increase bacterial growth and metabolism
	Sulfur compounds	High concentration of sulfur compounds will increase sulfide generation
	Temperature	High temperatures increase bacterial growth
	Toxic materials like metals	Reduces bacterial activity thus reducing sulfide generation
After generation of hydrogen sulfide gas in liquid phase		
Stage III – Release of hydrogen sulfide to sewer atmosphere	pH	pH below 7 increases H_2S concentration and increases its release into the sewer headspace
	Turbulence	High turbulence promotes H_2S release into the sewer headspace
	Temperature	High temperature reduces H_2S solubility in the wastewater and increases its release into the headspace
	Metals	Insoluble metallic sulfides may precipitate in presence of metals, thus reducing H_2S release
After release of hydrogen sulfide gas to sewer headspace		
Stage IV – Concrete or metal pipe corrosion	Moisture, humidity	Increases microbial activity, thus accelerating sulfuric acid production
	Temperature	High temperatures increase bacterial growth
	Alkalinity	High alkalinity of the sewer pipes provides resistance to corrosion attack

As is clear from the above discussion, various factors affect the hydrogen sulfide and corrosion rates in a manhole; thus, to predict hydrogen sulfide concentrations and corrosion rates inside a manhole accurately, it is important to

consider all these factors. To incorporate the factors accurately, a pilot lab study will likely not be sufficient; thus an extensive field study is required, which will be carried out in this research by sampling 350 manholes in the City of Arlington.

2.4 METHODS AVAILABLE TO MITIGATE CORROSION

Since repair and rehabilitation of corroded manholes is expensive, various methods are adopted to mitigate corrosion. The most commonly used methods are:

a. Using corrosion-resistant manhole materials.
b. Lining the manhole walls to prevent corrosion.
c. Preventing hydrogen sulfide generation using chemical agents.
d. Design changes to sewer systems.

2.4.1 CORROSION RESISTANT MANHOLE MATERIALS

Until the 1930s manholes were built commonly with bricks in a difficult and time-consuming process[35]. According to TCEQ, the manhole must now be monolithic and should be able to withstand weight[31]. The most commonly used material today is precast concrete, providing 100 years of manhole life[28, 36], and ensuring quality and quick installation[28]. However, these manholes are prone to corrosion. With advances in technology and the understanding of the corrosion process, now many materials are available in the market for manhole construction[35]. The 2 main types of material currently used for corrosion protection are:

a. **Plastic Manholes:** Made of a single piece of polyethylene, [28]plastic manholes are very resistant to corrosion[35, 36] and need less rehabilitation and maintenance[28]. In addition, they are sustainable and eco-friendly, by reducing contamination to soil as they are corrosion resistant thus resistant to leakage[28, 36]. However, they are inadequately designed to prevent excessive deflections due to ground loadings[38].
b. **Fiberglass Manholes:** They are only 1/10th in weight in comparison to concrete manholes and are thus easy to install[28, 36]. In addition, they are durable and eco-friendly[28, 36].

However, these materials are not adopted by many cities. Also, currently there are millions of manholes in US; it would be impossible to replace all of them with plastic or fiberglass manholes.

2.4.2 LINING THE MANHOLE WALLS

Many kinds of liners are used to rehabilitate manholes. The most commonly used liners are polyurea-based linings, lining sprays, stretchable systems, two component systems, rehabilitation liners, life extension liners, and sealant[36]. Applying sealant is a common method of rehabilitation and prevents inflow from the manhole chimney[36].Epoxies can be used for fixing corroded and damaged manholes[36]. Another method involves using chemical grout which acts as a barrier to water and stabilizes the soil, thus preventing infiltration into the manhole[36]. The liner system is currently adopted in many cities but used only in manholes needing extensive rehabilitation or prone to excessive rates of corrosion, as it is expensive. Thus, identifying manholes needing immediate care, and then lining them would be less expensive, in comparison to inspecting every manhole in the city for determining the conditions. The City of Arlington has spent thousands of dollars in lining the damaged manholes. The table 2.5 presents the expenditure faced by City of Arlington in lining the manholes. The cost in below table does not include the cost of project mobilization and trench safety which would further add to expense.

Table 2.5 Cost of lining the manholes

Material	Number of Projects	Year	Unit Price	Cost
Raven Epoxy Lining	2	2015 and 2019	$451	$32,169
Spraylock Epoxy Lining	2	2018 and 2019	$477	$32,672
Conshield	1	2015	$343	$30,009

2.4.3 USE OF CHEMICAL AGENTS

The important points to consider for reducing corrosion in sewers are:

a. Maintain dissolved oxygen greater than 0.5 mg/L[39].
b. Keep dissolved sulfide less than 0.1 to 0.3 mg/L[39].

c. Maintain hydrogen sulfide in the air less than 3 to 5 ppm[39].

d. Increase pipe crown pH to 4.0 or higher[39].

Thus, many chemical methods are adopted to increase the concentration of the dissolved oxygen, reduce the concentration of dissolved sulfide, or increase the pH of the pipe. Some of the usual methods are listed below[33]:

a. Oxidizing agents: Various oxidizing agents can be used to oxidize sulfide in wastewater[34]. Commonly used oxidants are oxygen, hydrogen peroxide, potassium permanganate, and ozone[33, 37]. An oxygen level of 1 mg/L is sufficient to prevent sulfide build-up in the wastewater[33]. So, oxygen is added into force mains, lift stations, inverted siphons, U-tubes, pressurized sidestreams, and/or areas of hydraulic turbulence, either as air or pure oxygen[33]. However, inappropriate oxygen application can result in stripping of H_2S gas from wastewater, thus increasing corrosion potential[33]. Hydrogen peroxide reacts with sulfide to produce elemental sulfur in acidic pH and sulfate in alkaline pH and can act as a supplement oxygen supply, thus reducing sulfide concentration[33], but is expensive[37]. Potassium permanganate is a strong oxidizing agent and is used for sulfide control, but it is combustible[33] and expensive[37]. Ozone can oxidize H_2S to sulfur or sulfite, but enough contact time is required which might be difficult[33].

b. Chlorination: Chlorine applied as chlorine gas, sodium hypochlorite or calcium hypochlorite oxidizes sulfide, thus reducing sulfide concentration[33, 34]. However, the process is effective only if chlorine is applied at a particular location from the point of sulfide build-up[33].Also, chlorine is a hazardous chemical[33, 37].

c. Iron salts: Many metals react with sulfide to form metallic sulfides, thus preventing H_2S release[33, 34, 37]. The most commonly used iron salts are ferrous or ferric chloride and sulfate[33, 34, 37]. Ferrous or ferric chloride can suppress the concentration of sulfide for up to 20 miles from the point of application[33, 34].

d. Alkali addition: At high pH H_2S mostly exists as hydrosulfide (HS^-), thus preventing the release of hydrogen sulfide to the sewer headspace[33]. Addition of sodium hydroxide or any other alkali helps to

achieve this[33]. However, organic acids flowing in the wastewater quickly reduce the pH of the wastewater even after the addition of the alkali[33]. So, shock dosing of caustic soda is done to inactivate the bacteria in the slime layer that convert sulfate to sulfide and inhibit sulfide production[33, 37].

e. **Sodium nitrate:** Usually in the absence of oxygen, bacteria tend to give preference to nitrate over sulfate as an electron acceptor[33, 34, 37]. So, sodium nitrate is added to wastewater to limit sulfate reduction because the bacteria will not reduce sulfate until all of the oxygen and nitrate in the wastewater is exhausted[33, 34, 37].

f. **Sewer ventilation:** Natural ventilation reduces hydrogen sulfide in the sewer and also reduces the humidity levels. Natural ventilation is provided through house vents[39] in the USA. The mechanism for natural ventilation to occur includes relative difference in air density between the sewer atmosphere and outside air, frictional drag of the wastewater at the air liquid interface, rise and fall of wastewater level within the sewer and changes in barometric pressure along the sewer[39]. Forced ventilation can be done with the use of fans[39].

However, all these methods are expensive and require special handling. In addition, these methods need to be applied regularly for them to function well, which adds to the cost. Thus, these methods prove ineffective in long run.

2.4.4 DESIGN CHANGES IN SEWER SYSTEMS

Proper design selection can help mitigate the corrosion effect in sewer systems[54]. Though initial cost of designing could add to cost, in long run this approach is economical, and would improve life of the infrastructure[54]. The various design procedures available to reduce corrosion is listed below:

a. **Slope:** Sewers having long runs with minimum slope are prone to sulfide generation due to residence times, poor oxygen transfer, and deposition of solids[54]. Steeper slopes increase turbulence thus aerating the wastewater and maintaining aerobic conditions in sewer systems[54]. In a 1950's study conducted in small collecting sewers of 15 and 20 cm (6 and 8 in) diameter in southern California, it was found that steeper

flows recorded average sulfide concentrations relatively low compared to flatter slopes[54]. Designing of the slope depends on various factors such as flow, EBOD, topography, subsurface conditions, depth of service laterals, pipe size and material[54]. However, slope requirements cannot be always achieved because steeper slopes result in deeper sewer lines thus adding to a pumping cost[54]. So, there is a trade-off required between economical cost and slope requirements[54].

b. **Velocity:** It is recommended that minimum velocity of 0.6 m/s (2 ft/s), should be maintained to achieve self-cleaning action in sewers regardless of the pipe size[54]. However, it is difficult to maintain the minimum velocities during low flow, thus resulting in organic solids depositions and development of anaerobic zones[54]. Another method would be to prevent the settling of suspended particles from settling out on the invert by maintaining minimum boundary shear stress[54].

c. **Pipe size**: A larger pipe size with flow rate and slope to a smaller pipe size will reduce the mean hydraulic depth, thus increasing surface area for reaeration[54]. However, smaller pipe has lesser detention time thus avoiding sulfide buildup, but this benefit of smaller pipe is reduced to greater ratio by the slime-supporting pipe wall to volume of wastewater[54].

d. **Drops and falls**: Drops for wastewater containing little sulfide help in reaeration of wastewater, thus reducing anaerobic conditions, however in presence of sulfide, drops can incorporate turbulence which will release sulfide to the wastewater[54]. Thus, it is preferable to avoid drops in wastewater with high sulfide concentrations[54]. Also, drops are prone to clogging, or stoppages due to debris collection over the drop pipe[54].

e. **Junctions and transitions**: Junctions and transitions allow for solid deposition and sulfide release due to turbulence[54]. Turbulence at junctions is created due to the large difference in velocity, flow of the various sewer lines meeting at the junction[54], and grade changes between upstream and downstream sewer lines[54]. Thus, junctions or transitions are built to change the velocity of merging stream gradually[54].

f. **Pumping stations**: Wastewater is stored at pumping stations for longer periods, thus creating conditions favorable for anaerobic oxidation[54]. Thus, a pumping station should be designed to contain adequate bottom slopes and suctions piping for continuous removal of deposited solids[54]. Wet wells should be as small as possible to minimize detention time[54]. Also, sulfide control methods such as aeration, chemical addition, oxygen injection methods can be used to reduce the sulfide concentrations in the wet well[54].

g. **Siphons**: Siphons or inverted siphons or depressed sewers convey wastewater under sewers or highways, or conduits to regain lost elevation[54]. Siphons are always flowing full and under pressure thus are potential site for sulfide generation[54]. Oxygen injection, chemical additions are commonly used methods to prevent sulfide generation in siphons[54].

From the above discussion, it is clear that design such as drops, siphons, junctions should be avoided when possible. Also, it is important to understand the characteristics of the wastewater and topography of the location to better design sewer slopes, and velocity. Design precautions can be undertaken only for future sewer systems, these precautions do not help mitigate the problem of corrosion in millions of manholes already present in the US. Also, design criteria cannot be achieved for every manhole constructions due to concerns of terrain topography, sewage quality, and cost.

2.5 EXISTING MODELS FOR PREDICTING SULFIDE AND CORROSION IN SEWERS

There are few available methods which try to predict the amount of sulfide generated in the sewers and the corrosion rates. Currently available models are listed below:

a. **1977 Pomeroy-Parkhurst Equation:** The equation can predict sulfide levels in partially- filled sanitary sewers[33]. The model is presented in Equation 2.11 below is used for the partially filled gravity sewers[33]:

$$\frac{d[S]}{dt} = M'[EBOD]r^{-1} - N(sv)^{0.375}[S]d_m^{-1} \quad \text{.....Eq. 2.11}$$

where

$\frac{d[S]}{dt}$ = -rate of change of total sulfide in mg/L - hr,

M' =effective sulfide flux coefficient for sulfide generation by slime layer in m/hr,

EBOD = effective BOD in mg/L,

r = hydraulic radius in ft,

N = empirical factor accounting sulfide losses by oxidation and escaping to atmosphere,

[S] = total sulfide concentration in mg/L,

s = slope of energy gradient in ft/ft,

v = mean wastewater velocity in ft/s, dm = mean hydraulic depth in ft.

For completely filled sewer pipes (force mains), Pomeroy developed another equation which is listed below[33]:

$$\frac{d[S]}{dt} = 3.28(M)[EBOD](1 + 0.48r)r^{-1} \ldots\ldots \text{Eq. 2.12}$$

where

S = total sulfide in mg/L,

t = flow time in hr,

M = specific sulfide flux coefficient in m/hr,

[EBOD] = effective BOD_5 in mg/L,

r = hydraulic radius in ft.

It has been observed that sulfide buildup increases with velocity, but Pomeroy's completely filled sewer pipe equation does not take into account the effect of velocity and sulfate concentration, thus making the model weak[48].

For calculating the rate of corrosion theoretically in cementitious pipes, Pomeroy developed Equation 2.13 below[33]:

$$C_{avg} = \frac{0.45k\phi_{sw}}{A} \ldots\ldots \text{Eq. 2.13}$$

where

C_{avg} = average rate of penetration in mm/yr or in/yr,

k = coefficient of efficiency for acid reaction (range from 0.3 to 1.0),

ϕ_{sw} = flux of H_2S to the pipe wall in g/m^2-hr,

A = alkalinity of the cement bonded material, expressed as $CaCO_3$ equivalent.

However, the model is good only in specific conditions[33]. For conditions when the dissolved oxygen concentration is higher than 0.5 mg/l, this model does not perform well[33]. If the wastewater characteristics vary significantly within a sewer system, each section has to be modeled separately, and then the final sulfide concentrations need to be calculated through mass balance, thus making the process very tedious[33].

b. **1982 Kienow, Pomeroy, and Kienow Model:** Since the 1977 Pomeroy-Parkhurst equation was very long and tedious, the 1982 graphical method was developed[33]. In this method the 1977 Pomeroy-Parkhurst equation is presented on a graph in the form of curves (each pipe diameter has its specific curve) to evaluate the limiting total sulfide concentration (S_{lim}) and factor combining hydraulic variables (β), which can then be used to predict the sulfide level at in sewer pipe using a nomograph[33]. The graphical method is based on Equation 2.14 below[33]:

$$S_{lim} - S_2 = (S_{lim} - S_1)\beta^{L/C_3} \ldots\ldots \text{Eq. 2.14}$$

where

S_{lim} = limiting total sulfide concentration in mg/L,

S_1 = upstream total sulfide concentration in mg/L,

S_2 = downstream total sulfide concentration in mg/L,

L = length of reach in meters/1,000 or feet/1,000,

C_3 = constant (1 for US units, 0.3048 for metric units),

β = factor combining hydraulic variables.

Two types of nomographs are used based on the β values[33]. For β values greater than 0.90, Nomograph A is used and for β less than 0.90, Nomograph B is used[33]. The calculated L and β values are spotted on the graph and then the lines are drawn connecting the two points to the unmarked line[33]. Then the line is extended to the $S_{lim} - S_2$ and $S_{lim} - S_1$ lines[33]. The corresponding $S_{lim} - S_2$ is noted and S_{lim} subtracted to find the downstream sulfide deficit[33]. The nomographs are shown in Fig. 2.8 below[33].

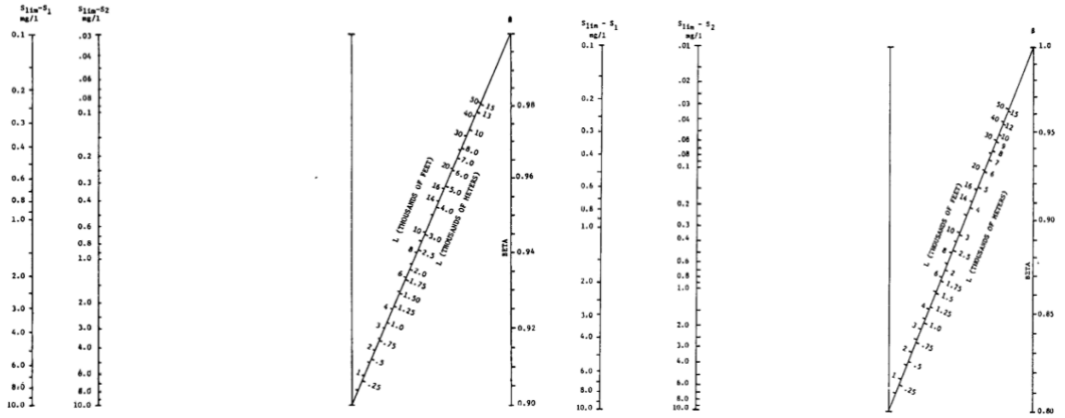


Fig 2.8 Shows nomograph A and B[33]

c. **SULF.BAS computer model:** SULF.BAS is based on the 1977 Pomeroy-Parkhurst equation, and is a Basic language computer model, which can predict sulfide build-up in the sanitary sewer system, and thus assess the need for corrosion protection[33]. The program can also predict corrosion in force-main or full pipe sections[33].

d. **Pomeroy Z formula:** This formula (Eq. 2.15) is good for predicting the sulfide build-up in small gravity sewers[33]. It should not be used for sewers greater than 600 mm or 24 in in diameter[33]. This formula just provides crude estimates and is advised not be used anymore[33].

$$Z = \frac{EBOD}{S^{0.50}Q^{0.33}}\frac{P}{b} \ldots\ldots \text{Eq. 2.15}$$

where

Z = defined function,

S = hydraulic slope,

Q = discharge volume in ft^3/s,

P = wetted perimeter in ft,

B = surface width in ft.

For Z values <5000, sulfide is rarely generated; for Z values >5000 but <10,000, sulfide is marginally generated, and for Z values >10,000, sulfide generation is common[33].

e. **EPA Flow-Slope Relationship:** Fig 2.9 below shows the relation between the slope, flow and the amount of sulfide that might be generated[37]. This flow and slope graph can be utilized for sulfide prediction in sewer pipes[37]. This formula just provides the crude qualitative guidelines and is advised not be used anymore[33].

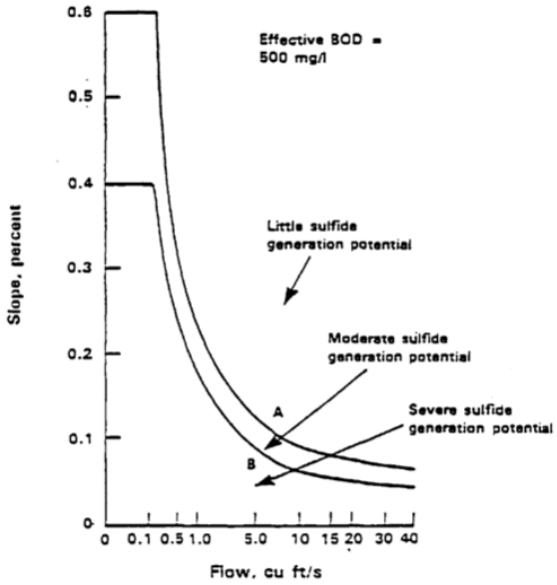


Fig 2.9 Flow and Slope relationship[33]

f. **Thistlewayte's Equation:** Sulfide concentrations in wastewater pipes can be predicted using the Thistlewayte's equation[48]:

$$\frac{d[S]}{dt} = 0.50 * 10^{-3} u[BOD]^{0.8}[SO_4]^{0.4}(1.14)^{T-20} r^{-1} \ldots\ldots \text{Eq. 2.16}$$

where

$\frac{d[S]}{dt}$ = rate of sulfide generation,

BOD = concentration of biological oxygen demand in mg/l,

T = wastewater temperature in Celsius,

r = hydraulic radius in m,

u = velocity of the stream in m/sec.

g. **Boon and Lister Equation:** Sulfide concentrations in wastewater pipes can be predicted using the Boon and Lister equation[48]:

$$\frac{d[S]}{dt} = 0.228 * 10^{-3}[COD](1.07)^{T-20}r^{-1}(1 + 0.37D) \quad \text{……Eq. 2.17}$$

where

$\frac{d[S]}{dt}$ = rate of sulfide generation,

COD = concentration of chemical oxygen demand in mg/l,

T = wastewater temperature in Celsius,

r = hydraulic radius in m,

D = diameter of pipe in m.

Boon and Lister equation does not take into account the effect of sulfate concentration and velocity which would impact the sulfide generation and volatilization repectively, thus making the model weak[48].

h. **Wells Model:** Wells(2014) developed a model to predict corrosion rates and life of standard reinforced sewer pipe[42]. However, the model is applicable in aggressive sewer conditions and higher temperatures[42]. The model utilized two sets of concrete coupons (one new and one old set taken from a 70 year old sewer pipe) to mirror conditions of new concrete, and old concrete which has already undergone extensive corrosion, and placed them in inverted positions in the sewer systems[42, 44, 49, 51]. According to Wells[51], Equation 2.18 below determines the corrosion rate in the sewer[51].

$$C = A * [H_2S]^{0.5} * \left(\frac{0.1602H - 0.1355}{1 - 0.9770H}\right) * e^{-\frac{45000}{RT}} \quad \text{……Eq. 2.18}$$

where

C = rate of corrosion in mm/yr,

$[H_2S]$ = concentration of hydrogen sulfide in sewer atmosphere in ppm,

H = relative humidity of the sewer atmosphere,

R = universal gas constant in 8.314 $Jmol^{-1}K^{-}1$,

T = sewer temperature in K,

A = constant, 207,750.

Also, based on Wells et al. (2014) analysis, they proposed that corrosion loss can be expressed as bilinear with two parameters, t_i (time for corrosion to initiate, or incubation time) and r (longer-term steady state corrosion rate)[42, 49, 51]. According to their study, if the corrosion process has already started, it will not change much with time[49].

Using these parameters, Wells predicted the life of the concrete in aggressive and high temperatures conditions[42]. For example, for a sewer with t_i = 0.9 and r = 12 mm/year, a concrete depth of 100 mm will have a life of 9.2 years.

However, this method is very specific to particular conditions in the sewer pipe. Also, initial conditions in a sewer system need to be known for finding the t_i and r values, thus necessitating inspections[42]. The model is good for predicting the life of sewer systems but cannot be used to identify the sewer structures needing urgent repair.

i. **Corrosion prediction based on H_2S uptake rate model:** This model is developed using Wells et al. (2014) research, and estimates corrosion rates based on laboratory studies, which employed batch reactors containing two sets of concrete coupons (one new and one old set taken from a 70 year-old sewer pipe) to mirror conditions of new concrete, and old concrete which has already undergone extensive corrosion, and injecting 50 ppm of H_2S gas for 33 months to determine H_2S corrosion rates[43]. Based on the H_2S uptake rates of pre-corroded sewer concrete, the amount of corrosion rates is determined[43]. The H_2S uptake rate for pre-corroded concrete material at controlled laboratory conditions, and for 50 ppm of H_2S gas at 25°C and 100% relative humidity for 33 months was[43]:

$$r_{H_2S} = \mu_{max}(1 - e^{-K[H_2S]}) \quad \text{……Eq. 2.19}$$

where

r_{H_2S} = hydrogen sulfide uptake rate,

μ_{max} = maximum uptake rate, mg- $m^{-2}h^{-1}$,

K = empirical coefficient, ppm^{-1}.

Using the r_{H_2S} the corrosion rate at 25°C and 100% relative humidity for pre-corroded concrete can be determined[43]:

$$C = \frac{r_{H_2S} * 2 * \left(\frac{24h}{day}\right) * \frac{\left(\left(\frac{365day}{year}\right) * 10^{-3}\right)g}{mg}}{M_S * 10^{-3}\left(\frac{m}{mm}\right) * 1m^2} * \left(\frac{1}{A}\right) \quad \text{……Eq. 2.20}$$

where

C = annual loss of concrete in mm/yr,

A = buffering capacity of concrete in mol-H^+(m^3 concrete)$^{-1}$,

M_s = molar mass of sulfur.

Since the model was developed under controlled conditions, this method may not fit real-time field conditions. Also, the model assumes that all of the sulfide adsorbed by the concrete is responsible for the corrosion rate, but that may be not possible in real time scenarios since some of the sulfuric acid might be neutralized by alkalinity or washed away. Also, flow conditions also impact the sulfide uptake drastically, as clear from above discussions; this was not factored into the model, so this model is not suitable for real-time situations.

2.6 CONCERNS WITH CURRENTLY AVAILABLE MODELS

All the models listed above prove the importance of various wastewater and sewer characteristics on the MICC process, but they are not able to encompass all the factors. Also, all the available equations apply to sewer pipes and not to manholes,. Currently no equation exists for estimating corrosion rates of manhole shafts. This research will fill this gap.

2.7 DATA ANALYSIS AND PREDICTIVE MODELING

Very limited studies have been carried out to predict microbially induced corrosion using statistical analysis. The studies below outline the predictive models developed to determine corrosion rates due to MICC.

a. **Corrosion rate prediction using lab scale studies:** Jiang et al. (2014) built a model to predict the rates of corrosion occurring in sewer pipe[47]. R software was used to build the statistical models and backward

selection method was utilized to determine the best model for predicting corrosion in concrete sewer pipe[47]. The pre-corroded concrete used for the study was obtained from the sewer and placed in a chamber to stimulate corrosion under 3 gas phase temperatures (17°C, 25°C, and 30°C), two levels of RH (100%, 90%), and 6 levels of H_2S (0 ppm, 5 ppm, 10, ppm, 15 ppm, 25 ppm, and 50 ppm) [47]. A few specimens of the concrete were partially submerged in the sewage (PSC) and a few were exposed to the gas phase of the sewage (GPC)[47].

Statistical results for GPC**:** Tree analysis *p* values of 0.003 and 0.002 were recorded for 12-24- and 24-45-months exposure of the concrete to increasing H_2S concentrations, respectively, thus signifying the importance of H_2S in corrosion rates[47]. The study also reported relative humidity to be a significant factor, with *p* values of 0.03 and 0.01 for 12-24- and 24-45-months exposure to H_2S, but temperature was found to be not significant[47]. Thus, corrosion rates for phase exposed concrete were controlled primarily by H_2S concentration but even relative humidity played a role[47].

Statistical results for PSC**:** Tree analysis *p* values of $1.5*10^{-6}$ and $6.4*10^{-7}$ were recorded for 12-24- and 24-45-months exposure of the concrete, respectively, to increasing H_2S concentrations, thus signifying the importance of H_2S in corrosion rates[47]. The study also reported relative humidity and temperature was found to be not significant[47]. Thus, corrosion rates for partially submerged concrete was controlled only by H_2S concentration[47].

Based on statistical analysis, the two models below were proposed for predicting corrosion rates in partially submerged concrete and gas phase exposed concrete[47].

Partially submerged concrete: $C_r = k.C_{H_2S}^n + C_{ri}$……Eq. 2.21

Gas phase exposed concrete: $C_r = k.C_{H_2S}^n.f_{BET}(RH) + C_{ri}$……Eq. 2.22

where

C_r = corrosion rate in mm/year,

C_{H_2S} = gaseous H_2S concentration,

k = model constants,

n = model constants,

C_{ri} = corrosion caused due to previous exposure,

$f_{BET}(RH)$ = BET sorption isotherm.

The gas phase exposed concrete equation could not be validated, but partially submerged concrete equation was validated by curve fitting with experimental data[47].

Based on the results of the study, we can confidently conclude the importance of hydrogen sulfide concentration, gas phase temperature, and relative humidity on corrosion rates. Also, it is clear from the model that the submerged sections behave differently from the unsubmerged sections, thus indicating that the manhole would tend to have corrosion rates very different from sewer pipes due to the presence of the larger chamber for gas phase exposure of the concrete. However, the study considered corrosion rates in very controlled setting with only 3 parameters, and also did not consider sewer design, physical properties of the wastewater, or the effect of flow velocities and turbulence. In addition, the corrosion rate equation for the GPC was not validated, thus making the model unreliable.

b. **Sulfide build up in filled sewer pipes:** Alani et al. (2013) conducted a study to predict the sulfide build up rates in filled sewer pipes[48]. The method employed the use of evolutionary polynomial regression (EPR), which is a combination of a genetic algorithm and the least square method[48]. EPR works by running the datasets repeatedly through the analysis with various combinations of functions and exponents, which trains the system to select the best model to predict a particular case scenario[48]. The accuracy of the EPR model is confirmed by the coefficient of determination (CD) value[48].

The study considered hydraulic radius to estimate the sulfide flux from the slime layer to the wastewater stream, temperature, concentrations of organic nutrients (COD), and velocity parameters[48]. Also, the final model was compared with already available sulfide prediction equations developed by Pomeroy, Boon and Lister, and Thistlewayte[48]. The data for building the model was taken from 3 different journal articles with

a total of 91 measured values[48]. Random combinations of data were selected for developing and testing the model[48]. The model generated by the EPR is[48]:

$$\frac{d[S]}{dt} = 0.0135[COD]^{0.5}T^{0.5}D^{-1}u^{0.5} \ldots\ldots \text{Eq. 2.23}$$

where

$\frac{d[S]}{dt}$ = rate of sulfide build-up in mg/l/hr,

COD = concentration of biological oxygen demand in mg/l,

T = wastewater temperature in Celsius,

D = internal diameter of the pipe in m,

u = velocity of the stream in m/sec.

The CD value of the model was 84%, which is very high; the CD values for other models were very low in comparison[48]: 18%, 7%, and 19% for the Pomeroy, Thistlewayte, and Boon and Lister models, respectively[48].

The study proves the effectiveness and accuracy of using a statistical tool to develop an equation in comparison to using an empirical study alone. A statistical model can increase the reliability of the equations, and its application. However, the model developed in the study used laboratory data and thus it may not be effective in a field situation.

c. **DDM models for prediction of corrosion rate:** In 2019, Li et al. (2019) tried to predict the time to initiate corrosion t_i and the rate of corrosion r using 3 data-driven models (MLR, ANN, ANFIS); then the t_i and r can be used to find the life of the concrete using the Tony Wells model. The data for the model was collected from Jiang et al. (2014)[47], which used pre-corroded concrete obtained from sewer and placed in a chamber to stimulate corrosion under 3 gas phase temperatures (17°C, 25°C, and 30°C), two levels of RH (100%, 90%), and 6 levels of H_2S (0 ppm, 5 ppm, 10, ppm, 15 ppm, 25 ppm, and 50 ppm) [47]. A few of the concrete specimens were partially submerged in the sewage (PSC) and a few were exposed to the gas phase of the sewage (GPC)[47]. The model was built using environmental factors like temperature, relative humidity, and concentration of gaseous hydrogen sulfide, and also considering if the concrete used

for the data collection was partially submerged in the wastewater or was exposed to the gas phase of the chamber (the location variable for the model)[52]. 60 data points were used to develop a model to predict t_i and 72 data points were used to develop a model for predicting r[52].

Before building the model, interaction between the different terms was studied, and it was found that there existed an interaction between location and relative humidity for building t_i model, and between location and H_2S for building the r model[52].

Multiple Linear Regression (MLR): Four different models for both times to initiate corrosion and corrosion rate were built.

Model for t_i: Four different models were built, one considering the full data set, another using only GPC data set, the third only using the PSC data set, and the last one with the full dataset with the inclusion of the interaction term, with R^2 values of 0.54, 0.76, 0.42, and 0.62 respectively[52]. The 4 models built for prediction of time needed to initiate corrosion are provided below[52]:

Full Dataset: $t_i = 96.34 + 1.68 * \text{Location} - 0.18 * H_2S - 0.54 * RH - 0.84 * T$……Eq. 2.24

GPC Dataset: $t_i = 147.7 - 0.160 * H_2S - 1.01 * RH - 1.08 * T$……Eq. 2.25

PSC Dataset: $t_i = 44.94 - 0.208 * H_2S - 0.0708 * RH - 0.592 * T$……Eq. 2.26

Full Dataset with interaction:

$t_i = 96.34 + 46.07 * \text{Location} - 0.184 * H_2S - 0.538 * RH - 0.835 * T - 0.467 * Location * RH$……Eq. 2.27

where

t_i = time to initiate corrosion,

Location = placement of the concrete in the wastewater chamber,

H_2S = Gas phase H_2S concentration,

RH = relative humidity,

T = Temperature of the gas phase.

According to the MLR models, the moisture content plays an important role in determining the time for initiation of corrosion, as the pore structure of the concrete will provide a pathway for the moisture, carbon dioxide, and hydrogen sulfide gas to penetrate into the concrete and cause corrosion[52]. Also, since the PSC are in contact with the wastewater, the initiation of corrosion will be faster in comparison to the GPC, which will receive moisture only from the sewer gas phase[52].

Model for r**:** Four different models were built, one considering the full data set, another using only the GPC data set, the third only using the PSC data set, and last one with the full dataset with the inclusion of the interaction term, with R^2 values of 0.61, 0.51, 0.64, and 0.71, respectively[52]. The 4 models built for prediction of corrosion rate are provided below[52]:

Full Dataset: $r = -0.173 - 0.453 * \text{Location} + 0.0282 * H_2S + 0.0087 * RH - 0.0157 * T$……Eq. 2.28

GPC Dataset: $r = -1.96 + 0.0119 * H_2S + 0.0293 * RH + 0.0016 * T$……Eq. 2.29

PSC Dataset: $r = 1.61 + \ 0.0446 * H_2S - 0.119 * RH - 0.0299 * T$……Eq. 2.30

Full Dataset with interaction:

$r = -0.173 - 0.167 * \text{Location} + 0.0283 * H_2S + 0.0869 * RH - 0.0157 * T - 0.0164 * Location\ RH$……Eq. 2.31

where

r = corrosion rate,

Location = placement of the concrete in the wastewater chamber,

H_2S = Gas phase H_2S concentration,

RH = relative humidity,

T = Temperature of the gas phase

According to the MLR model, the concentration of the hydrogen sulfide gas had a higher influence on the rate of corrosion than the other factors, and the effect of hydrogen sulfide on PSC was greater than GPC[52]. For both Artificial Neural Network (ANN) and Adaptive Neuro Fuzzy Inference Systems (ANFIS), the models built had a R^2 values of 0.8872 and 0.9384, respectively, for t_i and 0.9102 and 0.9402 for r, but further improvement in the model is needed by incorporating real time sewer data[52].

The DDM models do provide a good estimate of rates of corrosion by considering various environmental factors, and also the concrete conditions, but the models are based on the factors studied in the lab and would need validation via field studies.

2.8 GOAL AND OBJECTIVES OF THE STUDY

In US alone there are 12 million manhole structures, serving 190 million people,[38] Extensive corrosion of these manholes could result in catastrophic failures such as street collapses or sewer blockages[40], resulting safety hazards to the public and economic losses. Though various equations and models are available to reflect conditions in a sewer pipe, limited focus has been given to manholes. The design of the manholes is very different from that of a sewer because the manholes have a larger headspace in comparison to sewer pipes. Also, manholes are the points of intersection for sewer pipes, and thus suffer from high turbulence which is an important factor to determining the release of hydrogen sulfide to the headspace. Since manholes are closer to the ground surface they allow easy diffusion of the oxygen into the manhole airspace which can result in increased corrosion rates, but such a scenario is avoidable in a sewer pipes since they are buried deep into the earth. Also, the nearness of the manholes to the ground surface could mean that environmental changes would greatly impact corrosion rates in manholes in comparison to sewer pipes. Thus, it is not a good assumption to use sewer pipe corrosion rates equation to predict corrosion rates and life of the manholes.

This research will primarily focus on understanding corrosion rates occurring in manholes. As clear from the above discussion, till now no study has been conducted to understand the factors and their effect on corrosion rate in manholes, thus indicating the importance of this study. Various factors determine the amount of sulfide generated and released to the sewer headspace, and corrosion rates. Also, due to the economic and safety concerns involved with regular inspections of manholes, it is important to develop a model that would reduce the task of manual inspections. This research will provide the user with a tool to predict the manholes which are most likely to corrode based on their design conditions. Unlike sewer pipes, manholes have varying designs such as presence of multiple inlets, multiple drops, location at bends; these factors add to the corrosion rates. Hence, this research will consider both the gas phase conditions in the manhole, liquid phase factors of the wastewater along with the design of the manhole. Since manhole designs vary based on the terrain and location of the sewer pipes, each manhole is very

unique. Also, sewer conditions vary diurnally and seasonally, so to account the factor of time 48 hours of continuous field data will be recorded, and for lab analysis samples for each manhole will be collected every two hours for 48 hours. Also, to account for seasonal changes, a few manholes will be repeated in summer, fall, winter, and spring. Thus, to get a better understanding of the processes occurring in manholes, field data will be collected from 350 manholes in this study.

The primary goal of the project is to develop a predictive model to determine hydrogen sulfide generation and corrosion rates in a manhole depending on design and manhole conditions.

The specific objectives of the projects are:

a. To measure liquid phase parameters (temperature, dissolved oxygen, pH, sulfide, sulfate, and biochemical oxygen demand, velocity), and gas phase parameters (temperature, relative humidity, oxygen, hydrogen sulfide), that are responsible for regulating hydrogen sulfide concentrations for 48 hours, as well the amount of corrosion for 350 manholes in the City of Arlington.
b. To build predictive models for hydrogen sulfide generation and rate of corrosion in manholes, using the manhole physical design and measured liquid and gas phase parameters.

Our hypothesis is that manhole design, wastewater characteristics, manhole gas phase conditions, and weather play a major role in determining the rate of corrosion. So, manholes having drops, turbulent flow, or bends should exhibit higher rates of corrosion and should record higher hydrogen sulfide concentrations, in comparison to manholes with no drops, subcritical flow, and no bends. Also, higher wastewater temperatures or low pH should also result in higher gas phase hydrogen sulfide concentrations and increased rates of corrosion in manholes. In addition, manholes would be expected to register higher hydrogen sulfide concentrations in summer in comparison to winter, and thus greater rates of corrosion.

The predictive model will help to determine the rate of corrosion and hydrogen sulfide concentrations as a function of manhole design, wastewater characteristics, manhole gas phase conditions, and weather.

CHAPTER 3

METHODOLOGY

This chapter will first discuss the methods used to collect the data for 350 manholes for completion of Objective 1, including criteria for selecting the 350 manholes, instruments or sources used to gather the information on the various parameters, and methodology and lab protocols to collect and analyze the required data, and QA/QC of instruments. The chapter will then discuss the methodology used for building the prediction model.

3.1 EXPERIMENTAL DESIGN FOR DATA COLLECTION

3.1.1 350 MANHOLES SELECTION

The manhole design can create turbulence thus resulting in hydrogen sulfide gas volatilization. Thus, manhole design factor was considered for the selection of 350 manholes. In addition to the design considerations, other factors such as nearness to high hydrogen sulfide manholes or presence of lift stations will also result excessive corrosion, so they are also considered. To incorporate various design variations between the manholes, we are grouping the manholes into 6 categories. The 6 categories are:

1. Drop – manholes having a drop should be prone to higher rates of corrosion due to the turbulence added by the splashing of the wastewater.
2. Flow – Change in flow characteristics between the upstream pipe and downstream pipe can result in turbulence and thus higher rates of corrosion.
3. Pipe Size Change – Any variation in the upstream pipe diameter and the downstream pipe diameter could result in change of flow velocity thus increasing chances of H_2S gas volatilization.
4. Bends – A manhole situated at the junction of a bend may face change in flow velocity, and thus increased chance of H_2S gas volatilization.
5. Other parameters

a. Multiple inlets – The presence of multiple of upstream sewer lines pouring into the manhole could add to turbulence.
b. Adjacent to high H_2S manhole – A manhole downstream of a high H_2S manhole may have a probability of recording high levels of H_2S, which can then result in high corrosion rates in favorable conditions.
c. Adjacent to lift station – A manhole downstream to a lift station may record higher volatilization of the H_2S, because lift stations release large quantities of wastewater intermittently to the sewer lines.
d. With or without a guided channel –Guided channels may increase the velocity of the incoming wastewater as the flow area decreases according to the continuity equation of the fluids (Q=VA), thus may be increasing the rates of corrosion.

According to literature review, manholes facing the least amount of turbulence should undergo least amount of corrosion. So, based on this assumption, we selected 22 manholes as controls. Manholes selected to be controls must meet the following criteria to meet the assumption of less turbulence resulting in less corrosion:

1. A standard 0.1' drop only,
2. Subcritical flow of the wastewater,
3. Upstream pipe diameter and downstream pipe diameter must be same (uniform pipe size),
4. Single inlet and outlet,
5. Not be present at a junction which is a bend in the sewer system,
6. Not be near a lift station,
7. Not be near a manhole which has recorded high gas phase H_2S concentrations.

Table 3.1 shows the number of manholes belonging to various design categories, and the information source.

Table 3.1 Manhole selection

Category	Design Considerations	Manholes	Formula	Data Source
DROP	>2 feet	22	DS^* invert level of US^+ sewer pipe – US^+ invert level of DS^* sewer pipe	CleverScan, and Invert levels provided by the City
	Between 0.2 feet to 2 feet	22		
FLOW	Supercritical Flow	22	US^+ and DS^* Froude number >1	Average of Froude number calculated over 24 hours provided by the City.
	Subcritical Flow	22	US^+ and DS^* Froude number <1	
	Hydraulic Jump	22	US^+ Froude number >1 DS^* Froude number <1	
	US^+ subcritical and DS^* supercritical flow	22	Upstream Froude number <1 DS^* Froude number >1	
PIPE SIZE	Larger to smaller pipe	17	US^+ sewer pipe diameter – DS^* sewer pipe diameter	CleverScan, and Diameter provided by the City
	Smaller to larger pipe	17		
BENDS	Between 70° to 110° bend	53	NA	CleverScan and ArcGIS
	<110 °			
	Straight 170 to 190			
OTHER PARAMETERS	Multiple Inlets	53	NA	CleverScan and ArcGIS
	Adjacent to High H_2S	22		Field data collected and ArcGIS

	Adjacent to lift station	22		ArcGIS
	With or without guided channel	12		CleverScan
CONTROL	Standard 0.1' drop Subcritical Flow Uniform Pipe Single inlet and no bend Not near lift station Not adjacent to high H_2S manhole	22	DS^* invert level of US^+ sewer pipe – US^+ invert level of DS^* sewer pipe, US^+ and DS^* Froude number <1, US^+ sewer pipe diameter – DS^* sewer pipe diameter	City data, field study, ArcGIS
Total		**350**		

3.1.2 DATA COLLECTION

48 hours of real time data was collected for the 350 manholes. Two manholes on a given day were installed, and after 48 hours the instrument data was retrieved, and wastewater samples were brought to lab for further analysis. The instruments were then moved to other locations to collect data for another 2 manholes, and this process is carried on until data for 350 manholes is collected. Fig 3.1 shows the installation of the ISCO and the gas phase and liquid phase instruments at 2 locations. Fig 3.2 shows the collected wastewater samples after 48 hours of installation.

Fig 3.1 ISCO and gas and liquid phase instruments installation in field for 48 hours of data collection

Fig 3.2 Collected wastewater samples for lab analysis

3.1.3 PARAMETER SELECTION

Based on the MICC process, we selected 13 parameters that we believe will play a major role in affecting the hydrogen sulfide generation and volatilization, thus causing corrosion of the manholes. The 13 parameters were the predictor variables for our model, and we had 2 response variables. The 13 predictor variables are sorted into 3 different categories:

1. Weather conditions – represent the ambient conditions during those 48 hours of field data collection.
2. Gas phase parameters – represent the conditions in the non-liquid section inside the manhole for the 48 hours of data collection.
3. Liquid phase parameters – represent conditions in the wastewater inside the manhole for the 48 hours of data collection.

In addition to this, age of the manholes was used to determine corrosion in a manhole as a rate.

Table 3.2 provides the list of the predictor variables and response variables selected for the study. Predictor variables are represented with P and response variables are represented with R.

Table 3.2 Parameters measured and method of measurement

Category	Specific Parameters	Measurement Method	Frequency of Measurement	Modeling Variable
WEATHER CONDITIONS	Maximum temperature (F)	NOAA Climatological record from Arlington Municipal Airport, TX, US[(1)]	NA	P1
	Minimum temperature (F)			P2
	Precipitation (in)			P3
GAS PHASE PARAMETERS	Temperature (F)	OdaLog SL 1000 (App-Tek International) or OdaLog SL 50 (App-Tek International) Kestrel® DROP™ D2	Continuous for 48 hours	P4
	Relative humidity (%)	Kestrel® DROP™ D2	Continuous for 48 hours	P5
	Oxygen (%)	ToxiRAE Pro	Continuous for 20 minutes	P6
	Hydrogen Sulfide (ppm)	Odalog SL 1000 (App-Tek International) or Odalog SL 50 (App-Tek International)	Continuous for 48 hours	R1
LIQUID PHASE PARAMETERS	Temperature (F)	Aqua TROLL 600 Multiparameter Sonde Hanna Multiparameter (pH/EC/DO) Probe HI98194	Continuous for 48 hours	P7
	pH	Aqua TROLL 600 Multiparameter Sonde Hanna Multiparameter (pH/EC/DO) Probe HI98194	Continuous for 48 hours	P8

	Dissolved oxygen (mg/l)	Aqua TROLL 600 Multiparameter Sonde Hanna Multiparameter (pH/EC/DO) Probe HI98194	Continuous for 48 hours	P9
	Sulfide (mg/l)	ISCO 6712 Full-Size Portable Sampler to collect sample/lab analysis using EPA Method 9034-Titrimetric Procedure	3 samples collected over 48 hours (Peak 1, Peak 2, Composite)[#]	P10
	Sulfate (mg/l)	ISCO 6712 Full-Size Portable Sampler to collect sample/lab analysis using Spectrophotometer DR2800 and Turbidimetric method 10227 using TNT 864 and TNT 865 reagent set kits	3 samples collected over 48 hours (Peak 1, Peak 2, Composite)[#]	P11
	Biological oxygen demand (mg/l)	ISCO 6712 Full-Size Portable Sampler to collect sample/lab analysis using Standard Method 5210 B	3 samples collected over 48 hours (Peak 1, Peak 2, Composite)[#]	P12
	Velocity	Average of 24 hours of velocity data provided by City[@]		P13
MANHOLE CORROSION	Depth of corrosion	CleverScan	Once (entire height of manhole)	R2
RATE OF CORROSION	Age of manhole (yr)	Installed Date	ArcGIS data provide by the City	NA
		Contract Date		
		Accepted Date		
		Planned Date		

[#]3 samples (Peak 1, Peak 2, Composite)[#] - ISCO is programmed to fill 250ml in 2 bottles every 4 hours. Each bottle is 500ml. So, it takes 48 hours to fill all the 24 bottles. The time of starting the ISCO is noted and the sample from

the 24 bottles is then mixed into Peak 1, Peak 2, and Composite. Peak one represents bottles filled between 4pm-8pm, Peak 2 represents bottles filled between 8pm-12am, and the remaining bottles are considered for composite. This selection of peak times was done based on our initial results of hourly sampling.

[@]Velocity data was modelled for a pipe flowing full for 24 hours. Some of the velocity data was measured and remaining modelled.

3.1.4 FIELD INSTRUMENT AND LAB PROTOCOLS

3.1.4.1 WEATHER CONDITIONS

The temperature maximum, temperature minimum, and precipitation data were collected from the (National Oceanic and Atmospheric Administration) website. The station used to collect the weather condition data is Arlington Municipal Airport, Texas, US with station id GHCND:USW00053907[(1)]. The station is located 32.66361° latitude, and -97.09389° longitude with an elevation of 99%[(1)]. The temperature maximum, temperature minimum, and precipitation values were averaged based on start and end time of the ISCO installation in the field. The temperature data is reported in Fahrenheit and precipitation data is reported in inches.

3.1.4.2 GAS PHASE INSTRUMENTS

Instruments fused or gas phase data collection were dropped to hang at the mid-level of the manhole, to prevent atmospheric conditions from effecting the measurements of the gas phase instruments. Also, this ensures that recorded data actually represents the concentrations of hydrogen sulfide volatilized into the manhole from wastewater. Too close to the wastewater surface may result in instruments getting wet during high flows thus incorrect data recording.

3.1.4.2.1 ODALOG SL 1000 OR ODALOG SL 50 (APP-TEK INTERNATIONAL)

The OdaLog is built to monitor gas concentrations in harsh environments such as wastewater facilities[(2)]. It is a portable device which is corrosion resistant[(2)], and due to this property it was selected to measure gas phase data of hydrogen sulfide and temperature in the manhole.

Table 3.3 Sensor summary for Odalog SL 50 and Odalog SL 1000[(2)]

Model	Range	Temperature Range	Resolution	Accuracy
Odalog SL 50	0 ppm to 50 ppm	-20°C (-4°F) to 50°C (122°F)	0.1 ppm	±2ppm
Odalog SL 1000	0 ppm to 1000 ppm	-20°C (-4°F) to 50°C (122°F)	1 ppm	±2ppm

The instrument can function in the temperature range of -20°C (-4°F) to 50°C (122°F) [(2)] and hydrogen sulfide concentrations of 0 ppm to 50 ppm (Odalog SL 50) with an accuracy of ±2ppm[(2)] and resolution of 0.1ppm[(2)]. Electrochemical sensors are used to detect the concentrations of the hydrogen sulfide gas in parts per million (ppm)[(2)]. The electrochemical sensor is a small container with diffusion barrier consisting of electrolyte and electrodes[(2)]. The chemical reactions cause current flow within the sensor to change in relation to the level of the gas passing through the diffusion barrier, and this current output is displayed and recorded by the OdaLog[(2)]. The presence of other gases can affect the response of the sensors[(2)].

To maintain the accuracy of the OdaLog, the instrument was frequently calibrated as per the instructions provided by the manufacturer. The data from the OdaLog was transferred to a computer using Infrared communications in the form of csv or xls file. For data transfer we were using the IrDA device provided by the App-Tek international. The manhole hydrogen sulfide and temperature data were recorded every 1 minute for 48 hours. The instrument was dropped to hang at the mid-depth level of the manhole for 48 hours. OdaLog SL 50 was used for manholes with concentrations less than 50 ppm; otherwise OdaLog SL 1000 was used. The temperature values were reported in Fahrenheit and H_2S concentration in part per million (ppm). Figure 3.3 shows the OdaLog being used in manhole and with the IrDa sensor to transfer data.

Fig 3.3 OdaLog SL 50 (App-Tek International) in the manhole and OdaLog with the IrDA

3.1.4.2.2 KESTREL® DROP™ D2

The Kestrel Drop D2 (Figure 3.4) can be used for both temperature and relative humidity measurement. For accurate measurement of relative humidity, it should be given about 20 to 90 minutes to equilibrate[(3)]. Once the instrument is switched on, it stays on. The Kestrel DROP data can be transferred to a csv or an xls file using Kestrel Drop App on Android or an IOS phone via a Bluetooth. The manhole relative humidity and temperature data were recorded every 1 minute for 48 hours. The Kestrel has a battery life of about 4 months, after which the battery was replaced. The instrument was dropped to hang at the mid-depth level of the manhole for 48 hours. The temperature values were reported in Fahrenheit and relative humidity was reported in percent (%).

Fig 3.4 Kestrel Drop for Temperature and Relative humidity measurement

3.1.4.2.3 TOXIRAE PRO

ToxiRAE Pro (Fig. 3.5) is a wireless instrument which can be used to monitor toxic gases and oxygen deficiency/enrichment[(4)]. The instrument is easy to use since it has an LCD display with measurement units, datalog, battery status[(4)]. The instrument needs regular charging, so after every field trip, the instrument will be put to charge. The instrument can work in various modes; we will be using the basic submode under the normal operational mode. We utilized this instrument for measuring the gas phase oxygen levels in the manholes. The measurement was done every 1 minute for 20 minutes. The stored data was transferred to a csv file using the ProRAE Studio II software. Regular calibration of the instrument was carried out as per the manual. The instrument was dropped to hang at the mid-depth level of the manhole. The oxygen concentration in the manhole was measured in mg/l.

Table 3.4 ToxiRAE Pro specifications[(4)]

Oxygen sensor range	0 to 30% Vol
Resolution	0.1% Vol
Response Time	15 seconds
Temperature Range	-20° C to 50° C
Atmospheric pressure range	±10%(

Fig 3.5 ToxiRAE Pro Oxygen sensor

3.1.4.3 LIQUID PHASE INSTRUMENTS

3.1.4.3.1 AQUA TROLL 600 MULTIPARAMETER SONDE

Aqua Troll (Fig. 3.6) can measure multiple parameters and has an LCD screen, internal memory of 16MB, and an additional SD card storage of 16GB[(5)]. The operating temperature for Aqua Troll is -5°C (23°F) to 50°C (122°F), and a usable depth of 650 ft[(5)]. Aqua Troll was utilized for measuring the liquid dissolved oxygen levels, liquid pH, and liquid temperature. Aqua Troll was lowered into the wastewater for 48 hours to record liquid wastewater temperature, dissolved oxygen concentration, and pH. The temperature was reported in Fahrenheit, and dissolved oxygen in mg/l. Aqua Troll can be lowered into the manhole by using the In-Situ Rugged Cable System. The data was logged every 1 minute for 48 hours. The sensors were replaced and calibrated as per the manufacture's manual. The batteries of the Aqua Troll were replaced when required. The data stored in the Aqua Troll was transferred in the form of a csv or xls file by connecting the Bluetooth to the VuSitu Mobile App or by connecting to the Win-Situ 5 software using a cable or Bluetooth.

Since the Aqua Troll is lowered into the wastewater, it might collect solid materials, so the Aqua Troll was cleaned before being installed in a new location. Table 3.5 describes the sensors and their working in detail.

Table 3.5 Sensor summary for Aqua TROLL[5]

Sensors	Recommended Calibration Frequency	Potential Interferents	Resolution	Accuracy	Methodology
pH/ORP	10 to 12 weeks	Sodium Salts/ ions that are stronger reducing agents than hydrogen or platinum	0.01 pH unit	±0.1 pH units or better	Standard Methods 4500-H+, EPA 150.2
RDO	12 months	Temperature, atmospheric pressure, salinity, chlorinity	0.01 mg/l	±0.1mg/l from 0 to 8mg/l ±0.2mg/l from 8 to 20mg/l ±10% of reading from 20 to 50mg/l	EPA-approved In-Situ Methods (under the Alternate Test Procedure process): 1002-8-2009, 1003-8-2009, 1004-8-2009
Temperature	NA	NA	0.01°C	±0.1°C	EPA 170.1

Fig 3.6 Aqua TROLL 600 Multiparameter Sonde

3.1.4.3.2 HANNA MULTIPARAMETER (PH/EC/DO) PROBE HI7698194 WITH METER HI98194

Hanna (Fig. 3.7) is a portable logging multiparameter system. Hanna's operating temperature is -5 to 55°C, maximum depth of 20m (66ft)[(6)]. The probe has a multistrand-multiconductor shielded cable with internal strength member rated for 68kg (150lb) intermittent use[(6)]. It was utilized for measuring the liquid dissolved oxygen levels, liquid pH, and liquid temperature. Hanna was lowered into the wastewater for 48 hours to record liquid wastewater temperature, dissolved oxygen concentration, and pH. The temperature was reported in Fahrenheit, and dissolved oxygen in mg/l. The data was logged for every 1 minute for 48 hours. Hanna meter HI98194 used to record the data, and the batteries of the meter were replaced as needed. The sensors were replaced and calibrated as per the manufacture's manual. Since the Hanna is being lowered into the wastewater, it might collect solid materials, so it was cleaned before being installed in a new location. Table 3.6 describes the sensors and their working in detail.

Table 3.6 Sensor Summary for Hanna[6]

Sensors	Range	Resolution	Accuracy
pH/mV	0.00 TO 14.00 pH; ±600.0 mV	0.01 pH/ 0.1mV	±0.02 pH/ ±0.5 mV
Dissolved Oxygen	0.0 to 500.0%' 0.00 to 50.00 ppm (mg/l)	0.1%, 0.01 ppm(mg/l)	0.0 to 300.0%: ±1.5% of reading or ±1.0% whichever is greater; 300.0 to 500.0%: ±3% of reading 0.00 to 30.00 ppm (mg/l): ±1.5% of reading or ±0.10 ppm (mg/l) whichever is greater; 30.00 ppm(mg/l) to 50.00 ppm (mg/l): ±3% of reading
Temperature	-5.00 to 55.00°C; 23.00 to 131.00°F; 268.15 to 328.15K	0.01°C, 0.01°F, 0.01K	±0.15°C; ±0.27°F; ±0.15K

Fig 3.7 Hanna Multiparameter probe HI7698194 with the Meter HI98194

3.1.4.3.3 ISCO 6712 FULL-SIZE PORTABLE SAMPLER

ISCO 6712 is a comprehensive sampling system, can be easily moved from site to site, and is durable and corrosion resistant[7]. Some important sampler features we will be using include top cover, controller, center section, adjustable distributor arm, discharge tube and support spring, composite tube guide for composite sampling, tub, Real-Time Displays, peristaltic pump, vinyl and PTFE suction lines, LD90 liquid detector[7]. ISCO was used to collect the samples of the wastewater flowing through the manhole. The wastewater was collected through a tubing attached to the ISCO on one end and a filter unit on the other end. The filter end of the tube was lowered into the manholes, and the tube's length was adjusted based on the depth of the manholes. There are 24 bottles of 500 ml volume in each ISCO. We used 2 ISCOs to collect samples from two manholes for 48 hours. The ISCO was programmed to collect samples intermittently over 48 hours. The ISCO was powered using the Model 946 Lead-Acid Battery[8].

For BOD and sulfate analysis we used raw fresh wastewater, but for sulfide analysis we added a preservative (sodium hydroxide + zinc acetate). So, the odd bottles in the ISCO were added with 5 ml of preservative to maintain the sample quality. The ISCO was programmed to fill 2 bottles with 125 ml of wastewater (1/4 full) at a given time. Two bottles (1 even + 1 odd numbered bottle) were filled every 4 hours after which the ISCO adjustable distributor arm moved to fill the next 2 bottles. This way within 48 hours all the 24 bottles were filled. To preserve the physical and chemical properties of the wastewater, we added dry ice to the center section of the ISCO tub when temperatures are above 25.5°C (78°F). For temperatures below 25.5°C (78°F), normal ice was added, as our initial findings showed that ice maintained necessary 4°C temperature for 48 hours at temperatures below 25.5°C (78°F). After 48 hours, the filter end of the ISCO suction pipe was cleaned and the samples were brought back to the lab and stored in a refrigerator for further analysis. Table 3.7 shows the ISCO 6712's specifications.

Table 3.7 ISCO 6712 specifications[7]

Controller weight with pump tube	13.2 lbs (6.0 kg)
Controller Dimensions	Length: 10.3 in (26.0 cm) Width: 12.5 in (31.7 cm) Height: 10 in (25.4 cm)
Temperature Ranges	Operational: 32 to 120°F (0 to 49°C) Storage: 0 to 140°F (-18 to 60°C)

Fig 3.8 shows the ISCO with the 24 bottles and adjustable distributor arm.

Fig 3.8 ISCO 6712 Full-Size Portable Sampler with 24 bottles of 500 ml volume and Adjustable Distributer Arm

3.1.4.4 LAB ANALYSIS INSTRUMENTS AND PROCEDURE

3.1.4.4.1 LAB SAMPLE PREPARATION

From our preliminary analysis of 7 manholes around the City of Arlington, we found that the peak hours were 4PM to 8PM and 8PM to 12AM. The peak hours represent the time when we noticed greater concentrations of hydrogen sulfide and the wastewater volume was higher. Based on this analysis we split our samples into three categories:

Category 1: Peak 1 represent wastewater samples collected from 4PM to 8PM.

Category 2: Peak 2 represent wastewater samples collected from 8PM to 12AM.

Category 3: Composite represents all the wastewater samples collected except the ones collected during Peak 1 and Peak 2 hours.

These samples were further divided into the odd and even samples. The odd bottles contained preservative, whereas the even bottles did not contain the preservative. The time of starting the ISCO is noted, based on that time we started our calculation to determine the time each of the ISCO sample bottles were filled and divide them into categories. For example, if the ISCO was installed at 8 am then the sample bottle division would be as represented in Table 3.8.

Table 3.8 Sample preparation

Time	**Odd Sample Containing Preservative**	**Even Sample Without Preservative**	**Sample Preparation**
	ISCO Bottle Number	**ISCO Bottle Number**	
8AM – 12PM	1,13	2,14	Composite
12AM – 4PM	3,15	4,16	Composite
4PM – 8PM	5,17	6,18	Peak 1
8PM – 12AM	7,19	8,20	Peak 2
12AM – 4AM	9,21	10,22	Composite
4AM – 8AM	11,23	12,24	Composite

If a particular ISCO sample bottle falls into 2 categories, it will be considered for the category which represents the maximum amount of time or equally distributed between the two categories if such a determination is not possible.

3.1.4.4.2 SAMPLE PRESERVATION

a. **Preparation for BOD and sulfate analysis:** According to the EPA, the samples for BOD and sulfate analysis needs to be stored at 4°C. For sulfate samples stored in 4°C, can be held for 28 days and for BOD they can be held for 48 hours[9,10].

b. **Preparation for sulfide analysis:** The ISCO bottle samples used for the sulfide analysis were collected over a period of 48 hours during which time the sulfide ions may react with the oxygen or other dissolved ions in the wastewater. To prevent this, we added 5ml of preservative to the odd bottles. According to EPA, volume of zinc acetate added depends on the nature of the sample, the amount of precipitate formed, and the assurance of excess of zinc ions[10,11,12]. So, to meet these requirements for all the samples, a conservative value of approximately 25 ml is selected. This preservative maintains the physical and chemical conditions of the samples so they can be stored for weeks to give time for lab analysis[9,10,11]. The zinc reacts with the sulfide to form zinc sulfide, which precipitates and is not easily oxidized when stored for days or for weeks, thus ensuring reliable analysis[10]. Also, this preservative eliminates interferences from other metal ions, sulfite, iodide, and many other soluble substances[10,11].

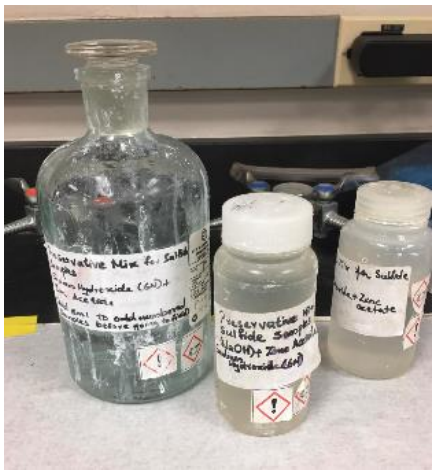


Fig 3.9 Prepared Preservative for sulfide sample preservation

3.1.4.4.3 SULFIDE – EPA METHOD 9034 TITRIMETRIC PROCEDURE

EPA Method 9034 is used to determine the acid soluble and acid insoluble sulfide in a sample. Odd numbered ISCO bottle samples were utilized for this analysis. The sample already contains the preservative to prevent the loss of sulfide ions. The samples were then filtered using the G4 grade of glass fiber filter with a particle retention of 1.2 µm to separate the soluble and insoluble sulfide. Iodine is added to the sample to convert the sulfide to sulfur, which was measured by titration using sodium thiosulfate as a titrant until the blue iodine starch complex disappears[9,10,13]. Duplicates were done for every sample.

3.1.4.4.3.1 IODOMETRIC TITRATION

200 ml of wastewater sample was taken and kept for vacuum filtration. G4 glass fiber filters with particle retention of 1.2 µm was used to collect the insoluble solids. The filtrate collected was used to determine the soluble sulfide and the solids on the filter paper was used for insoluble sulfide measurement. The total sulfide in the sample was reported as the sum of soluble and insoluble sulfide. The vacuum filtration apparatus was utilized for the separation of the soluble and insoluble sulfide. Table 3.9 shows the procedure used to determine the soluble and insoluble sulfide concentrations in mg/l.

Table 3.9 Soluble and insoluble sulfide determination procedure[9,13]

Soluble Sulfide	Insoluble Sulfide
Filtrate is added to 20 ml of 0.025N iodine solution	Filter paper containing the insoluble solids is added to the 100 ml of deionized water to which then 20 ml of 0.025N iodine solution is added
Then 2 ml of 6N HCl and 4 drops of starch indicator was added	Then 2 ml of 6N HCl and 4 drops of starch indicator was added
The filtrate is then titrated against 0.025N sodium thiosulfate until the blue color turns colorless	The filtrate is then titrated against 0.025N sodium thiosulfate until the blue color turns colorless

The equations used to determine the soluble and insoluble sulfides are presented below.

((ml of I_2*I_2 Normality) – (ml of Thiosulfate*Thiosulfate Normality))*(32.06 g/2 eq)*1000 / Sample volume in ml = Sulfide in mg/l……Eq. 3.1[9]

Total sulfide in mg/l = Soluble sulfide in mg/l + Insoluble sulfide in mg/l……Eq. 3.2

3.1.4.4.3.2 STANDARDIZATION OF IODINE

25 ml of 0.025N iodine solution was taken in a flask and 2 ml of 3N HCl was added. The iodine solution was titrated with the 0.025N sodium thiosulfate until the color changes to yellow, after which 4 drops of starch indicator was added, and the solution was titrated again until the color changes to colorless. The normality of the iodine solution was calculated based on the formula below[13].

Sodium thiosulfate Normality * Sodium thiosulfate volume used = Iodine Normality * Iodine Volume taken……Eq. 3.3[13]

Fig 3.10 shows the glass fiber filter used for the analysis and the iodometric titration process.

Fig 3.10 Vacuum filtration of the sample and iodometric method for sulfide determination

3.1.4.4.4 SULFATE – TURBIDIMETRIC METHOD 10227

For sulfate determination we are using the HACH 10227[17,18] method, which is based on the EPA prescribed method 9038[15]. For this we used the HACH TNTplus 864, which measures sulfate concentrations in the range of 40 mg/l to 150 mg/l[18], and HACH TNTplus 865, which measures sulfate concentrations in the range of 150 mg/l to 900 mg/l[17]. Even-numbered ISCO bottle samples stored at 4°C were utilized for this analysis[9,10]. Duplicates are done for every sample. Since the samples contain many floating impurities which may impact the accuracy of the Spectrophotometer, grade 410 filter paper with particle retention of 1µm was used to filter the sample before performing the analysis. DR2800 spectrophotometer was used to measure the sulfate concentration. Table 3.10 shows the procedure followed to determine the sulfate concentration in the wastewater in mg/l.

Table 3.10 TNTplus 864 and TNTplus 865 procedure[17,18]

TNTplus 864	TNTplus 865
Add 5 ml of the sample to the vial, and then add spoonful of Reagent A	Add 2 ml of the sample to the vial, and then add spoonful of Regent A
Invert and shake the vial for 2 minutes	Invert and shake the vial for 1 minute, and then start reaction time of 30 seconds
Clean the vial and read the sulfate concentration using spectrophotometer	Clean the vial and read the sulfate concentration using spectrophotometer

The wastewater samples were first analyzed with the TNTplus 864, and only if the sulfate concentration registers higher than 150 mg/l used the TNTplus 865.

3.1.4.4.4.1 ACCURACY CHECK OF SPECTROPHOTOMETER

Accuracy of the spectrophotometer was checked every month using a sulfate standard solution with known concentration of 100 mg/l. The spectrophotometer if working accurately will read the sulfate concentration value within ±5% error[16]. Fig 3.11 shows the TNTplus 864 and TNTplus 865 kits and the process of sulfate determination.

Fig 3.11 HACH 10227 method for sulfate determination

3.1.4.4.5 BIOLOGICAL OXYGEN DEMAND – EPA STANDARD METHOD 5210B

Biological oxygen demand (BOD) method is used to determine the relative oxygen requirements. According to Standard Methods, this process is utilized to measure oxygen used to biochemically degrade organic compounds, oxidize inorganic compounds, and oxidize reduced forms of nitrogen[10]. There are different methods for BOD determination such as 5-day period BOD method (5210B), ultimate BOD method (5210C), and Respirometric method (5210D)[10]. For our study we used the 5-day period BOD method (5210B) to determine the biological oxygen demand (BOD) of the wastewater collected in the even numbered ISCO bottle samples. Duplicates were done for each sample.

In the 5210B method we measure the change in the dissolved oxygen concentration due to microbial decay of organic matter in the wastewater stored in a 300 ml bottle for 5 days in a dark room at 20°C. The change in the dissolved oxygen between the zeroth day and fifth day was used to evaluate the BOD of the wastewater.

3.1.4.4.5.1 BOD 5210B PROCEDURE

We used the 300 ml recommended BOD bottles for this analysis. The bottles were cleaned and dried for quality control after every analysis. For BOD analysis the pH must be between 6 to 8[10,19]. During our field data collection, we recorded the pH, and the wastewater does record pH in these ranges, so we will not be checking the pH in the lab again. Since we used wastewater, we will not be seeding the BOD bottle as there are enough microbes in the wastewater to degrade the organic matter. Also, we did not add a nitrification inhibitor to the samples since we are concerned with total oxygen demand exerted on the wastewater rather than only the carbonaceous biochemical oxygen demand. The total oxygen demand provided us with a better perspective of the anaerobic conditions in the wastewater.

From our initial analysis of 10 manholes, the average BOD recorded was 351.1 mg/l thus it is clear that the wastewater contains enough BOD, so a sample volume of 2 ml of the wastewater was added to the 300 ml BOD bottles[20]. Table 3.11 below provides the minimum sample volume to be used for different waters provided by HACH[20].

Table 3.11 Minimum sample volume requirements for BOD[20]

Sample Type	BOD, mg/l	Volume, ml
Strong trade waste	600	1
Raw and settled sewage	300	2
	200	3
	150	4
	120	5
	100	6
	75	8
	60	10
Oxidized effluents	50	12
	40	15
	30	20
	20	30
	10	60
Polluted river waters	6	100
	4	200
	2	300

For BOD the bottles were filled with the reagent water a little over the neck. Care was taken to prevent the formation of bubbles in the bottles. Initial DO was measured as soon as the bottles are filled with the reagent water, after which the bottles were sealed. DO was measured using BOD LDO Probe connected to the HQd Portable Meter. The probe had a stirrer attached at its end, which was used to stir the samples before recording the readings. The probe cap was washed with deionized water after each measurement. The bottles were sealed using a water seal, and then the cap of the bottle was sealed with parafilm to avoid evaporation of the water seal. Before sealing, the bottles were checked for the presence of bubbles, as they result in incorrect recording of dissolved oxygen values.

The BOD bottles were stored at laboratory temperature which is usually at 20°C to 25°C temperature for 5 days. After 5 days the parafilm was removed and the 5^{th} day DO was recorded; care was taken not to stir the sample while taking 5^{th} day readings. Blanks were used to eliminate the oxidation occurring due to the dilution water. The BOD in mg/l of the samples was calculated by the formulas below.

Final BOD in mg/l = Sample BOD in mg/l – Blank BOD in mg/l……Eq. 3.4

Sample BOD in mg/l = (Sample 0^{th} day BOD – Sample 5^{th} day BOD)/Dilution Factor……Eq. 3.5

Blank BOD in mg/l = (Blank 0^{th} day BOD – Blank 5^{th} day BOD)/Dilution Factor……Eq. 3.6

Dilution Factor = Volume of the wastewater/ Volume of the Reagent water = 2/300 = 0.00667……Eq. 3.7

3.1.4.4.5.2 BOD LDO PROBE – MODEL LBOD10101

We will be using the BOD LDO Probe Model LBOD10101 for recording the DO values. The probe is a luminescent dissolved oxygen (LDO) sensor with integrated stirring system. It has a dissolved oxygen range of 0.05 to 20.0 mg/l (ppm), and an accuracy of ±0.05 mg/l for 0.0 to 10 mg/l and ±0.1 mg/l for greater than 10 mg/l[(21)]. It needs a stabilization time of 10 seconds when stirred[(21)]. The instrument was calibrated as per the manual's requirements. The probe is connected to the HQd meter to record and report the DO values. To maintain the probe in good condition, it was stored in BOD bottle filled partially with the deionized water, when not in use[(21)].

3.1.4.4.5.3 HQD PORTABLE METER

The BOD LDO Probe was connected to HQd Portable Meter for data collection. Meter batteries were changed when required. The meter had an option to record and save data to be later transferred to an USB drive[(22)]. Fig 3.12 shows the dilution water preparation and 0^{th} and 5^{th} day analysis for 2 wastewater samples.

Fig 3.12 BOD reagent preparation and 0th Day and 5th Day BOD determination

3.1.4.5 DEPTH OF CORROSION MEASUREMENT

The CleverScan device is regularly used in inspections of manholes. It provides flat screen images, and a dense point cloud as the output[(23)]. The dense point cloud can be used along with 3D modelling to understand the structure and geometry of the manholes. Also, it provides a shape file which consists of the X,Y,Z coordinates of the manholes, which can be used to determine the diameter of the manholes. It has four laser scanners and 5 HD cameras to capture high resolution images.

In our research, CleverScan was used for 2 major purposes. Firstly, it was used to measure the current diameter of the manholes, which helped us understand the current conditions inside the manhole, as well as how much corrosion has taken place inside the manhole. Shape files were used to collect the X,Y, and Z coordinates of the manholes. We will be rejecting diameter which is 10% less than the dx_{max} and dy_{max}, from the top manhole in the recorded data of the shape files, because those data could be inaccurate, as upper sections of the manhole are usually built narrower (in the form of the cone).

Assumption: Lower sections of the manhole are least corroded:

Comparing average of the lower sections of 20% data points (diameters in x and y directions, d_x and d_y) with the remaining average of manhole diameters in x and y directions (d_x and d_y). The lower sections of the manhole represent the section of the manhole near the water surface.

(d_x averaged over the upper remaining height of manhole)– (d_x averaged over the lower 20% datapoints of the height of the manhole) = change in diameter……Eq. 3.10

(d_y averaged over the upper remaining height of manhole)– (d_y averaged over the lower 20% datapoints of the height of the manhole) = change in diameter……Eq. 3.11

Fig 3.13 shows the width (dx) and height (dy) of the manholes. Since the manholes are not perfect circles 2 different diameter values were obtained. The differences in the diameter of the 20% with the rest of the manhole diameter provided us with the corrosion that had already taken place in the manhole. The reason this process was selected was because age of manholes varied anywhere from 120 years to 7 years, and city did not have records for all the manholes, thus the corrosion determination was made assuming either the upper sections or lower sections of the manhole were most corroded.

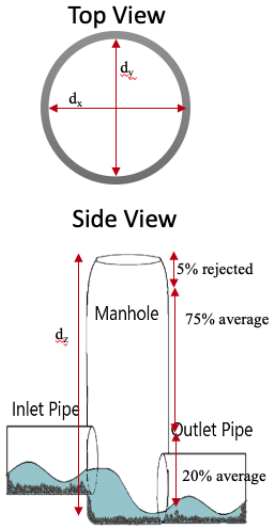


Fig 3.13 Corrosion rate determination: Lower 20% most corroded

Secondly, the CleverScan was used for determining the design of the manholes. CleverScan provides the side view, top view, of the manholes. Also, CleverScan has tools (ruler for diameter determination) which allow for determination of the height of the inlets and outlets, diameter of the inlets and flow direction, and number of inlets. Thus, CleverScan was used to determine the design of the manholes, which is used in building in predictive model

for hydrogen sulfide generation and corrosion rate. Fig 3.14 shows CleverScan tool been used for design determination for one of the manholes.

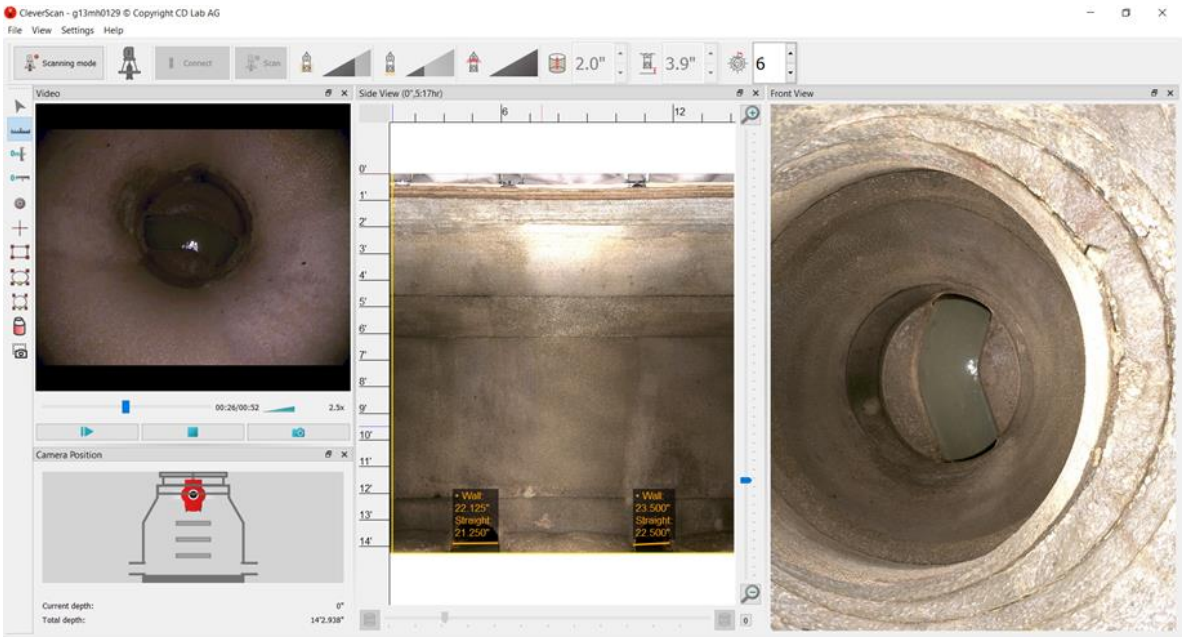


Fig 3.14 CleverScan being used design determination

To determine the design category into which a given manhole fit, we used various sources of information. The primary method was using the CleverScan videos and shape files; however, in some cases CleverScan determination was not possible due to video snow, the presence of a drop with high gushing wastewater flow, or the dark environment in deep manholes. In these cases, a combination of ArcGIS data, CleverScan recordings, and design files provided by the City were used for manhole design determination. Fig 3.15 shows CleverScan been used in the field and the point cloud generated.

Fig 3.15 CleverScan being used in the field for manhole diameter measurement and the point cloud

3.1.5 QA/QC FOR THE INSTRUMENTS

3.1.5.1 GAS PHASE INSTRUMENTS QA/QC

The below table shows the QA/QC followed for the gas phase instruments.

Table 3.12 Gas phase instruments QA/QC

Instrument	**Determination**	**Calibration**	**Maintenance**
OdaLog SL 1000 or OdaLog SL 50 (App-Tek International)	Continuous logging	Do 4-5 fresh air days after every 11 days of sampling. Then do pre-use span with 50 ppm H_2S span gas as in manual. If accuracy not within 10% then calibration must be done. Factory calibration once/year. Battery lasts 8 months.[(2)]	
Kestrel® DROP™ D2	Continuous logging	NA	Periodically change battery[(3)]
ToxiRAE Pro	Continuous logging	Every 6 months[(4)]	Recharge battery[(4)]

3.1.5.2 LIQUID PHASE INSTRUMENTS QA/QC

Table 3.10 below shows the QA/QC followed for the liquid-phase instruments.

Table 3.13 Liquid phase instruments QA/QC

INSTRUMENT	SAMPLING	DUPLICATES	BLANKS	CALIBRATION	MAINTENANCE
Aqua TROLL 600 Sonde	Continuous logging	NA	NA	Check once in 6 months with calibration solution[5]	Cleaning & storing overall sonde: Rinse sonde thoroughly, with warm water and soap, then rinse. Allow to air dry. Put the restrictor in storage mode with 15 mL of water[5]. Cleaning & storing pH sensor: Clean pH sensor with cold water. The pH sensor must be kept moist. Stored with storage solution[5]. Cleaning & storing DO sensor: Leaving sensor cap on, rinse sensor water and wipe with a soft cloth, then perform a 2-point calibration. DO sensor can be stored dry[5].

Hanna Probe HI7698194	Continuous logging	NA	NA	Quick calibration method will be used for pH, DO, temperature will be done every 6 months using quick calibration solution[6]	Cleaning & storing pH sensor: Clean pH sensor with cold water. The pH sensor must be kept moist. Stored with quick calibration solution [6]. Cleaning & storing DO sensor: The DO membrane will be replaced every 2 months, and electrolyte solution every 1 month[6].
ISCO 6712 Sampler	24 bottles in 48 hours	NA	NA	NA	End of each logging session: charge battery, clean the unit, inspect tubing for wear[7].
Sulfide	Samples collected for 48 hours	Each sample done twice	NA	Standardization of iodine done with standard sodium thiosulfate every time iodine is prepared[10,13]	Regularly acid cleaning for vacuum filter membrane.
Sulfate	Samples collected for 48 hours	Each sample done twice	NA	Every month spectrophotometer accuracy checked with 100 mg/l sulfate solution[17,18]	NA

Biological Oxygen Demand	Samples collected for 48 hours	Each sample done twice	Duplicates of blank run with every sample	LBOD probe is calibrated every month[21]	NA

3.2 DATA ANALYSIS

The liquid phase parameters, gas phase parameters, weather data, and manhole design data were used to develop the equation to predict the MICC rate in the manholes.

Table 3.2 in the above section provided the list of the predictor variables and response variables selected for the study. Predictor variables were represented with P and response variables are represented with R.

The data collected through field and lab studies were used to build a multiple regression model using the process outlined by Kutner[24]. Python software was used to carry out the analysis. Two models, one predicating the rate of corrosion, and other for determining the hydrogen sulfide concentration, were developed. The level of significance selected for the model building is $\alpha = 0.10$. Of the collected data, 80% of the manholes were used for building the predictive model and remaining 20% manholes were used for validating the model. The following steps were considered to develop a reliable model.

a. Development of a preliminary model, along with checking model assumptions (normality, constant variance, and correlation) and diagnostics (outliers, leverage, variance of inflation).
b. Transformations to satisfy all the model assumptions, if needed.
c. Determination of any interaction terms with the use of partial regression.
d. Standardizing the variables and rechecking the model for correlation.
e. Searching for best models using the backwards deletion, best subsets, or stepwise regression methods.
f. Selecting the model that best describes the corrosion process and verifying it for model assumptions and diagnostics.

g. Validating with the remaining 20% of manholes.

3.2.1 MULTIPLE LINEAR REGRESSION

Multiple linear regression using the Python software was used for developing the rate of hydrogen sulfide generated and the rate of corrosion equation. The analysis results in 3 equations similar to the form represented by Eq. 3.12, Eq. 3.13 and Eq. 3.14. Eq.3.12 represents the rate of hydrogen sulfide produced in the manhole, and Eq. 3.13 and Eq. 3.14 represents rate of corrosion taking place in the manhole along the dx and dy axes, respectively.

$$Y_1 = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots \ldots\ldots \beta_{35} x_{35} + \beta_{36} x_{36} + \varepsilon \ldots\ldots \text{Eq. 3.12}$$

$$Y_2 = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots \ldots\ldots + \beta_{36} x_{36} + \beta_{37} x_{37} + \varepsilon \ldots\ldots \text{Eq. 3.13 – along the dx axis}$$

$$Y_3 = \beta_o + \beta_1 x_1 + \beta_2 x_2 + \cdots \ldots\ldots + \beta_{36} x_{36} + \beta_{37} x_{37} + \varepsilon \ldots\ldots \text{Eq. 3.14 – along the dy axis}$$

Where variables are defined as listed in Table 3.14.

Table 3.14 Predictor and response variables for multiple linear regression

Parameters			**Variable**
Flow Rate, MGD			x_1
Velocity, ft/s			x_2
Depth of Flow, ft			x_3
Manhole Depth, ft			x_4
Average Ambient Temperature, °F			x_5
Precipitation inches			x_6
Design Parameters	Drop	>=2’	x_7
		>=0.2’, <2’	x_8
		<=Std 0.1’	x_9
		Maximum drop height, ft	x_{10}
	Pipe Size Change	Smaller to larger	x_{11}

		Larger to smaller	x_{12}
		Uniform	x_{13}
		Maximum Pipe size change, inches	x_{14}
	Flow Type	Supercritical flow	x_{15}
		Subcritical flow	x_{16}
		Hydraulic jump	x_{17}
		Subcritical Upstream, supercritical Downstream	x_{18}
	Bends	<80° Bend	x_{19}
		80-110° Bend	x_{20}
		>110° Bend	x_{21}
		Multiple Inlets	x_{22}
		No Bend 170° - 190°	x_{23}
	Adjacent to High H_2S manhole	Upstream to high H_2S manhole	x_{24}
		Downstream to high H_2S manhole	x_{25}
	Adjacent to lift station	Upstream to a lift station	x_{26}
		Downstream to a lift station	x_{27}
	Other parameters	Guided channel	x_{28}
		Non-guided channel	x_{29}
Liquid-Phase Parameters	Average Temp., °F		x_{30}
	Average DO, mg/L		x_{31}
	Average pH		x_{32}
	Average Sulfide, mg/L		x_{33}
	Average Sulfate, mg/L		x_{34}
	Average BOD, mg/L		x_{35}
Gas-Phase Parameters	Average Temp., °F (H_2S File)		x_{36}
	Average H_2S, ppm		Y_1, x_{37}

Corrosion change (lower 20% most corroded)	Diameter Change Width, ft/ Age in years	Y_2
	Diameter Change Height, ft/ Age in years	Y_3
Age Years		

CHAPTER 4

RESULTS AND DISCUSSION

This chapter will first discuss results and report on the manholes data that was procured during the study. Next the chapter will demonstrate trends observed among the various manhole designs, physical and chemical properties of the liquid and gas phases, hydrogen sulfide generation, and rate of corrosion. Last, the chapter will describe how multiple linear regression will help in developing better relationships among the various parameters.

4.1 DATA COLLECTION RESULTS

Our first objective was to measure the gas and liquid phase parameters as well as corrosion for 350 manholes, a large enough sample set with various manhole designs. Table 4.1 shows the number of manholes sampled during the 2 ½ years of data collection, which totaled 366. In the 366 total, several manholes were counted multiple times because they were sampled in different seasons, with accordingly different measurements of temperature and other parameters.

Table 4.1 Manholes sampled by season

Parameter	**Fall 2017**	**Spring 2018**	**Summer 2018**	**Fall 2018**	**Spring 2019**	**Summer 2019**	**Fall 2019**	**Target**	**Achieved**
Manholes	73	77	44	27	66	54	25	350	366

Table 4.2 demonstrates the number of manholes done in each design category. Initially, our assumption was that each manhole would have a unique design, but upon field study it was noticed that each manhole could fit into various design categories; thus, for most design categories, the target number of manholes was greatly exceeded, except for manholes adjacent to a lift station, and manholes with subcritical flow upstream and super critical flow downstream. Only 7 manholes were sampled in the lift station category was because there are only 4 lift stations in the City of Arlington, with only 7 manholes around the lift stations which were accessible and usable for data

collection. The City of Arlington does not have any manholes in the subcritical flow upstream and supercritical flow downstream for which hydraulic modeling was available, so this category could not be included in the study and will not be further discussed.

Table 4.2 Target and Actual Manholes Sampled in Each Design Category

Category	Design Type	Target	Achieved
Drop	>=2’	20	49
	>=0.2’, <2’	20	76
	<=Std 0.1’	N/A	340
Pipe Size Change	Smaller to larger	15	162
	Larger to smaller	15	26
	Uniform	N/A	268
Flow	Super-critical flow	20	41
	Sub-critical flow	20	260
	Hydraulic jump (upstream super, downstream sub)	20	35
	Sub-critical upstream, super-critical down-stream	20	0
Bends	<80° Bend	50	4
	80-110° Bend		52
	>110° Bend		66
	Straight 170 – 190°	N/A	68
Other parameters	Multiple Inlets	50	171
	Adjacent to H_2S manhole	20	52
	Adjacent to lift station	20	7
	Non-guided	10	62

	Guided	N/A	241
	Control	20	33

Note: Cells highlighted in peach are the designs which did not meet the minimum target number.

4.2 VARIABLES AND DATA SELECTION FOR HYDROGEN SULFIDE AND CORROSION RATE MODEL-BUILDING

For further analysis and development of the hydrogen sulfide equation, data sets from 146 manholes will be used. Of the original 366 manholes, only 146 had complete data sets, for reasons listed below, and summarized in Table 4.3:

a. Manhole age unknown: Manholes included in our study belonged to various time periods, with the oldest manholes built in 1899, and the youngest manholes built in 2014. Since some manholes are old, city does not possess sufficient hydraulic data (flow data, velocity, depth of flow) or installation date for the manholes, thus rendering those manholes not useful for the study.

b. Manhole design uncertain: In some cases, design data available from the city did not match field conditions. We used CleverScan scan and shape files to identify manholes design, but as discussed above, this was not always possible as CleverScan's efficiency is impacted in conditions of heavy wastewater flow or inlets submerged in wastewater.

c. Manhole data collected on different days: In some cases, when an instrument malfunctioned, we were able to complete data collection on a different day. These manholes were not use in the analysis, to better capture the effect of temperature and environmental conditions on hydrogen sulfide generation and corrosion.

d. Data logged partially or intermittently for one or more liquid-phase parameters: The instruments were installed in the field for 48 hours; however, in some cases waste (tissue, debris on the sensors) entangled the instruments, resulting in incomplete or intermittent data recorded (typically for liquid-phase

temperature and dissolved oxygen). Since these manholes did not have a complete 48-hours of data, they were excluded from the analysis.

e. Data not logged for one or more liquid-phase parameters (excluding O_2, RH): In some cases, no data was collected for one or more liquid-phase parameters, due to debris entangling the instruments, flow conditions impacting the instruments, weather conditions (such as summer weather resulting in ISCO batteries not able to pump enough samples for analysis), or instruments out for repair. We attempted to go back and collect this missing data on different days; however, in most cases, the manhole was no longer accessible due to construction issues or the instrument for logging the liquid phase data was still not working.

Table 4.3 Dataset selection for hydrogen sulfide model-building

Description of Manholes	**Number of manholes**
Total Sampled	366
a) Manhole age unknown in addition to hydraulic data not provided	25
b) Manhole design uncertain in addition to missing hydraulic data and age	2
c) Manhole data collected on different days	43
d) Data logged partially or intermittently for one or more liquid-phase parameters	22
e) Data not logged for one or more liquid-phase parameters (excluding O_2, RH)	94
Complete sets for hydrogen sulfide model building	146

Table 4.4 presents the manholes used in each design category used for developing the hydrogen sulfide prediction equation. The number of complete data sets exceeded the target for all categories except the larger to smaller, hydraulic jump, and adjacent to high H_2S manhole, which were 5, 2, and 3 manholes short, respectively (in addition to sub-critical upstream/super-critical down-stream and adjacent to lift station, for which there were not enough manholes available in the sewer system with hydraulic modeling data available). Being somewhat short in these 3 categories was not anticipated to substantially impact model accuracy. 20% of the manholes were used for model validation; 28 manholes for hydrogen sulfide model validation, and 25 for corrosion model validation.

Table 4.4 Manhole data sets for developing hydrogen sulfide and corrosion rate models, by design category

Category	Design Type	Target	Complete data set for:	
			H_2S model	Corrosion rate model
Drop	>=2’	20	20	17
	>=0.2’, <2’	20	30	28
	<=Std 0.1’	N/A	136	122
Pipe Size Change	Smaller to larger	15	68	62
	Larger to smaller	15	10	10
	Uniform	N/A	111	96
Flow	Super-critical flow	20	20	17
	Sub-critical flow	20	108	98
	Hydraulic jump (upstream super, downstream sub)	20	18	15
	Sub-critical upstream, super-critical down-stream	20	0	0
Bends	<80° Bend	50	2	2
	80-110° Bend		21	17
	>110° Bend		27	23
	Straight 170 – 190°	N/A	30	28
Other parameters	Multiple Inlets	50	66	60
	Adjacent to H_2S manhole	20	17	16
	Adjacent to lift station total	20	2	2
	Non-guided	10	30	29
	Guided	N/A	91	91
	Control	20	30	18
	Validation	20%	28	25

For the further analysis and developing of corrosion rate equation, we will be using data from 130 manholes, which had complete data sets including CleverScan data. We began the project in September 2017; however, the CleverScan arrived in October 2018. Thus, we had to revisit the manholes sampled from Sept. 2017 – Oct. 2018 for CleverScan data collection. During the first year, many manholes became inaccessible due to new construction. Some were closed and replaced, and others occurred on private property, which we were not able to revisit. Of the 146 manholes with complete data sets for hydrogen sulfide modeling, for 16 manholes we were not able to collect CleverScan data, leaving 130 complete data sets for corrosion rate modeling. As shown in Table 4.4, for corrosion rate model building, the number of datasets falls short in most categories, which could reduce the accuracy of the model.

4.2.1 VARIABLES NOT INCLUDED IN MODEL-BUILDING

Gas phase oxygen data was not included because the sensors could not survive harsh manhole conditions; thus, oxygen data was only able to be collected at 31 manholes. This would result in severe data loss. In addition, the oxygen values for the 31 manholes do not vary much, as clear from the graph below. The values mostly fluctuate between 18.2% to 21% oxygen, with average oxygen percentage of 20.2% and a standard deviation of 0.65%. Also, there does not appear to be any correlation between hydrogen sulfide and oxygen concentration, as shown in Fig 4.1. Thus, this variable will not be considered in further analysis.

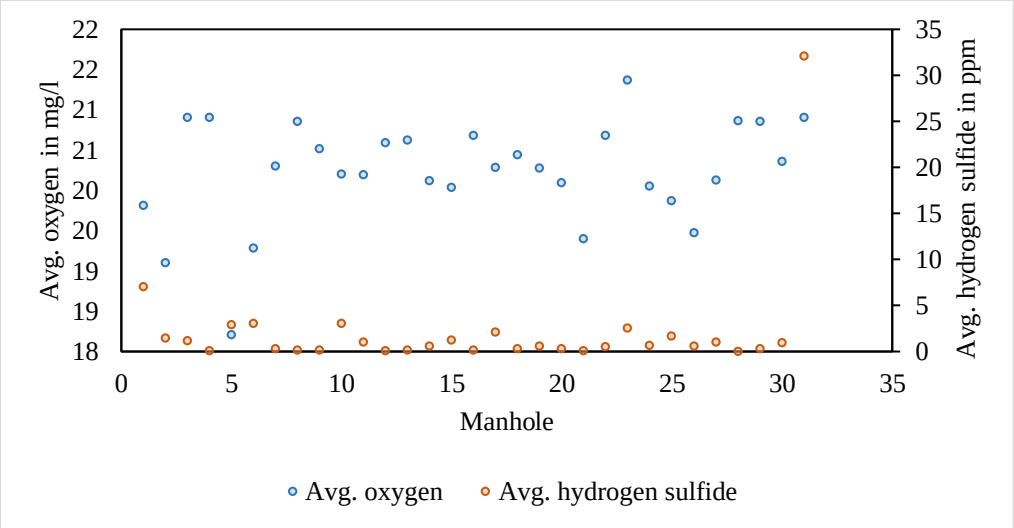


Fig 4.1 Average oxygen concentration for 31 manholes, along with hydrogen sulfide concentrations

Of the 146 manholes for which datasets for hydrogen sulfide model-building were otherwise complete (except for O_2), 118 had relative humidity measurements. Most of the average relative humidity values for the 118 manholes are close to 100% with an average of 95% and standard deviation of 11.9%, as shown in Fig 4.2 below. Also, there does not appear to be any correlation between hydrogen sulfide and relative humidity, as shown in Fig 4.2. Thus, the parameter will no longer be considered for analysis. Table 4.5 shows the averages and standard deviation for the oxygen and relative humidity for the manholes for which they were measured.

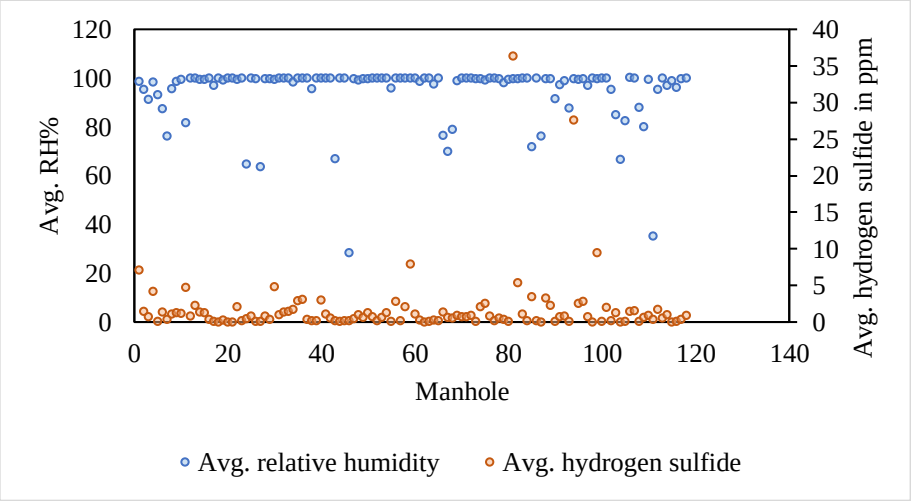


Fig 4.2 Average relative humidity for 118 manholes, along with hydrogen sulfide concentrations

Table 4.5 Average oxygen and relative humidity values for manholes for which they were measured

Statistic Parameter	Oxygen	Relative Humidity
Average	20.24	94.63
Standard Deviation	0.65	11.96
Maximum	21.37	100.08
Minimum	18.21	28.27

4.3 HYDROGEN SULFIDE GENERATION AND FACTORS

Of the 146 manholes, 90.4% recorded an average gas phase hydrogen sulfide between 0 to 3 ppm. The highest recorded hydrogen sulfide concentration was 36.28 ppm and the lowest was 0 ppm, with an average gas phase hydrogen sulfide concentration of 1.73 ppm, as seen in Fig 4.3. In addition, these values may not represent the maximum hydrogen sulfide in the manhole shaft, as the Odalog sensors were installed halfway along the depth of the manhole to avoid the sensors getting damaged due to the wastewater flow, and also to record average concentrations along the depth of the manhole rather than maximum values. The hydrogen sulfide data collected is

right skewed, thus signifying the data is not normally distributed; thus, there is need for transformation to carry out multiple linear regression.

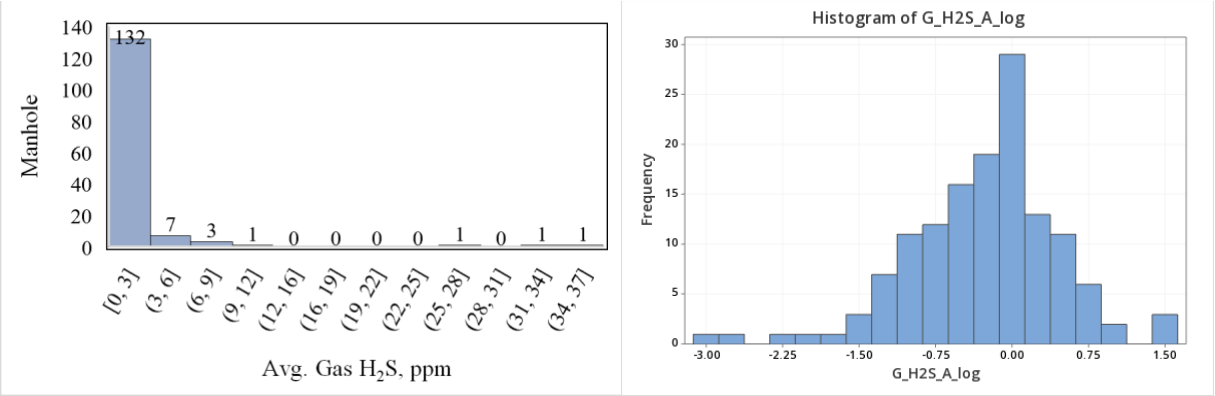
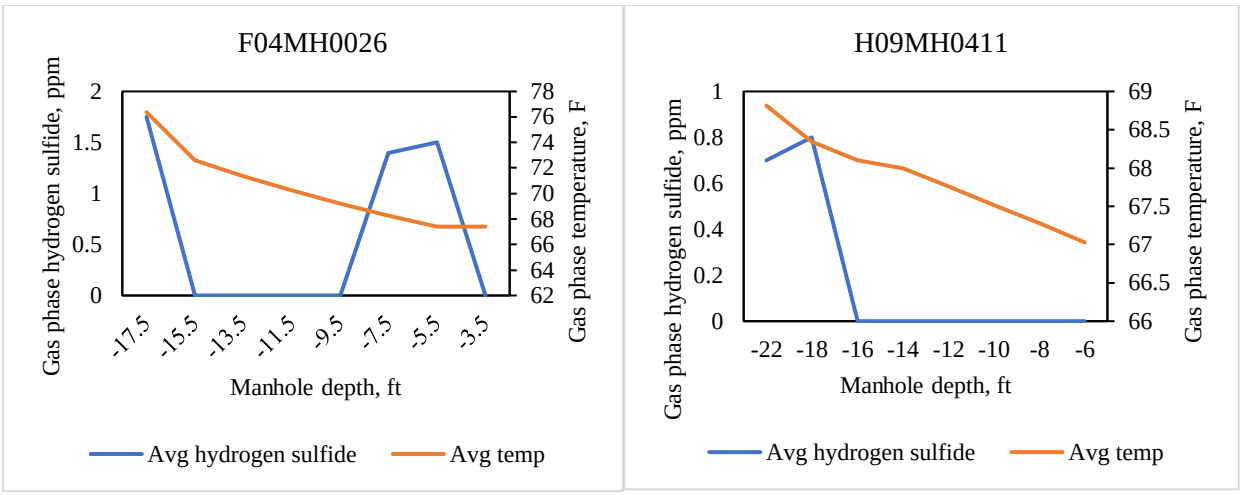


Fig 4.3 Histogram for average gas phase hydrogen sulfide measured in 146 manholes

4.3.1 VARIATION OF HYDROGEN SULFIDE WITH DEPTH

For several manholes, hydrogen sulfide concentration and temperature measurements were made along the depth of the manhole, as shown in Fig 4.4. The temperature increased with the increase in the manhole depth; this is reasonable, as the wastewater is usually at a higher temperature than ambient temperature. Romanova[45] found ambient temperatures to be 3.5°C lower than wastewater temperatures. In addition, as seen in Fig 4.4, the highest hydrogen sulfide concentrations were recorded 1 to 5 feet from the inlet or wastewater flow. In Fig 4.4, the manhole F04MH0026 with depth of 23.88 feet has 3 inlets at height of 17.2 feet, 10.15 feet, and 3.87 feet, and in the graph, we notice 3 clear peaks 1 to 3 feet above the inlets. Similarly, manhole H09MH0411 with depth 22.45 feet does not have a drop, so we notice the peak at 4 feet above the inlet. Third, manhole G12MH0226 with the depth of 18.21 feet has a high drop of 7 feet, with a hydrogen sulfide peak at 7 feet declining until it meets another peak due to the other inlet at the base of the manhole. It is expected that the highest hydrogen sulfide concentrations would occur near the inlets. The reason for peaks typically occurring a few feet above the inlets could be “plume rise” of hydrogen sulfide due mechanical turbulence caused by the high velocity flow of wastewater, or thermal buoyancy due to the elevated wastewater temperature.



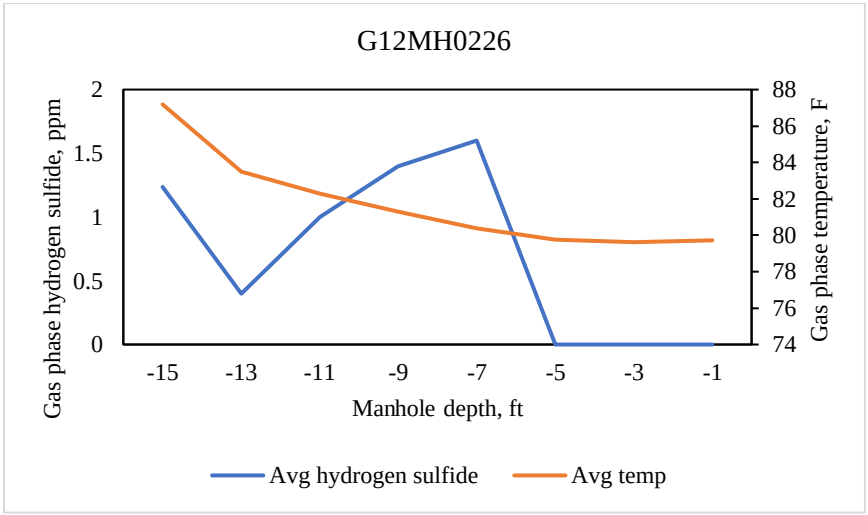


Fig 4.4 Depth effect on temperature and average hydrogen sulfide concentration

From the above graphs, it is clear that the height of inlets impacts hydrogen sulfide concentrations.

4.3.2 EXAMPLES ILLUSTRATING EXPECTED EFFECT OF MANHOLE DESIGN ON HYDROGEN SULFIDE

The above discussion on effect of manhole depth on temperature and hydrogen sulfide concentration indicates the effect of design on manholes. Based on those results, we selected varied design manholes to illustrate the

relationship between the various gas and liquid parameters over 48 hours and the effect of designs on hydrogen sulfide generation.

4.3.2.1 EXAMPLE EFFECT OF DROP ON GAS-PHASE HYDROGEN SULFIDE

We distributed our manholes in 3 different sub-categories of inlet drop height, with high drop being greater than 2 feet, low or medium drop being between 0.2 to 1.9 feet, and no drop being 0 to 0.19 feet. Fig. 4.5 shows example hydrogen sulfide concentrations for manholes with 3 different drop heights (9.08 feet drop, 0.260 feet drop, and no drop (0 feet)). In all 3 cases, manholes were chosen with subcritical flow and uniform pipe size diameter, and not adjacent to a hydrogen sulfide manhole or a lift station. The high drop manhole had a second inlet at zero feet height, and guided flow; the medium and no-drop manholes had non-guided flow.

In this example, the high drop manhole did record higher hydrogen sulfide surges in comparison to other drop designs; this result is in accordance to the results discussed by ASCE in their report[33] stating that designs which cause excessive turbulence in wastewater could result in higher sulfide release in the manhole's headspace. For the manhole with no drop, the hydrogen sulfide values are negligible with no severe surges in the hydrogen sulfide concentration over the course of 48 hours. In addition, a diurnal pattern in all the 3 drop designs is clearly visible with higher hydrogen sulfide concentration in the afternoons and late evenings, and lowest temperatures during the early morning, similar to patterns observed by Wells in a prior sewer system study[42].

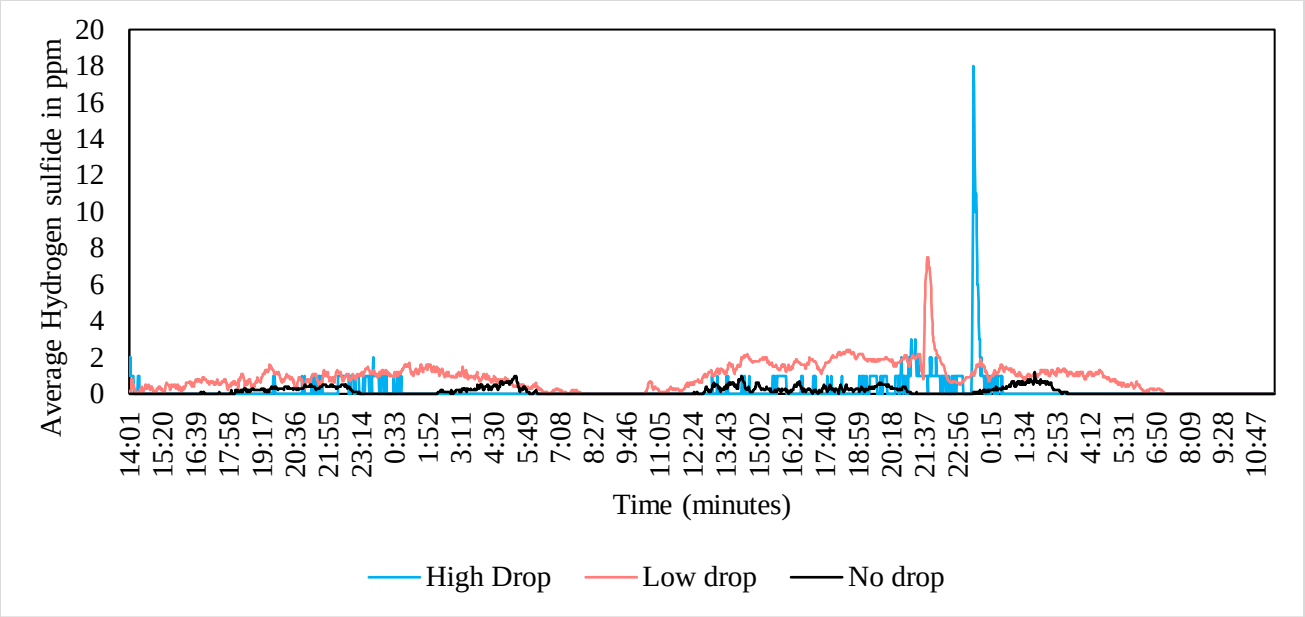


Fig 4.5 Hydrogen sulfide concentration for example manholes with drops of different heights

4.3.2.2 EXAMPLE EFFECT OF PIPE SIZE CHANGE ON GAS-PHASE HYDROGEN SULFIDE

We distributed our manholes in 3 different pipe size changes of smaller to larger pipe size (implies the inlet pipe diameter is smaller than outlet pipe), larger to smaller pipe size (implies the inlet pipe diameter is larger than outlet pipe), and uniform (both inlet and outlet pipe have same diameters). Pipe size change usually results in a change in wastewater velocity, which could result in turbulence and thus more dispersion of hydrogen sulfide into the manhole's headspace. Fig. 4.6 shows the hydrogen sulfide concentration for manholes with 3 different pipe size changes (outlet pipe is 3 inches larger than inlet pipe -smaller to larger, outlet pipe is 4 inches smaller than inlet pipe -larger to smaller, and uniform). For all 3 manholes, the inlet was at zero feet height and flow was subcritical, and the manholes were not adjacent to a high hydrogen sulfide manhole or lift station. Smaller to larger pipe size change manhole had guided flow with a bend angle greater than 110 degrees; the larger to smaller pipe size change manhole had straight flow in a guided channel; and the uniform pipe size manhole had non-guided flow.

From the figure, it is clear that larger to smaller pipe size changes resulted in higher hydrogen sulfide being recorded in comparison to uniform pipe size designs. This indicates that excessive turbulence in wastewater could result in higher sulfide release into the manhole's headspace; this result is in accordance to the results discussed by ASCE in

their report[33]. However, the difference in both the smaller to larger pipe size change manholes and uniform pipe size change manholes is not very different, except that we can see intermittent higher hydrogen sulfide concentration in smaller to larger pipe size change manholes. These minimal higher values of smaller to larger pipe size change manhole shows that there is an effect of diameter change even though if small. Also, the diameter difference for the inlet and outlet pipe for smaller to larger pipe is only 2 inches; maybe a higher difference in diameter result in higher turbulence. Also, in both the smaller to larger pipe size change and larger to smaller pipe size change, we see intermittent surges in the hydrogen sulfide concentrations. This implies that any type of pipe size change causing velocity change and thus turbulence in wastewater will result in higher hydrogen sulfide release, irrespective of its type. For the manhole with uniform pipe size change, the hydrogen sulfide values are negligible with no severe surges in the hydrogen sulfide concentration over the course of 48 hours. Also, a clear diurnal pattern is seen here as well, with higher hydrogen sulfide concentrations between 11 AM to 7 PM in larger to smaller pipe and 10 PM to 8 AM in smaller to larger pipe.

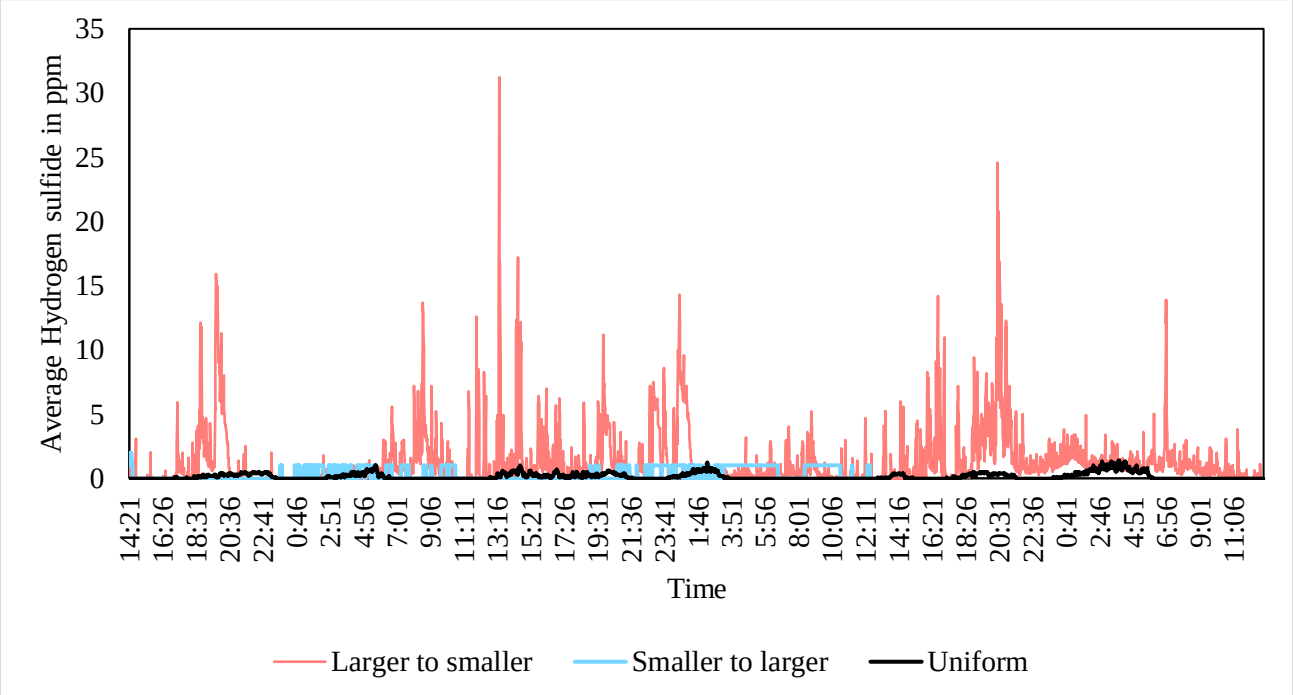


Fig 4.6 Hydrogen sulfide concentration for varied pipe size changes in manholes

4.3.2.3 EXAMPLE EFFECT OF ANGLE OF FLOW/MULTIPLE INLETS ON GAS-PHASE HYDROGEN SULFIDE:

According to study a conducted by EPA, sharp bends or manholes with colliding flows will have greater turbulence and thus higher hydrogen sulfide generation[33]. Thus, for this study, 3 types of angles of flow/bends were considered, along with straight flow and multiple inlets, as listed below.:

a. Acute angle bend (bend less than 80 degrees)
b. Right angle bend (bend angle between 80-110 degrees)
c. Obtuse angle bend (bend angle between 110 -170 degrees)
d. Straight flow (flow direction is between 170-180 degrees)
e. Multiple inlets flow (more than one inlet).

Fig. 4.7 shows hydrogen sulfide concentrations from example manholes in the sub-categories above. All manholes had inlets at zero feet height and uniform diameter for the inlet and outlet pipe. All had subcritical flow, with the exception of the acute angle bend, which had supercritical flow. All manholes had guided flow, with the exception of the straight flow manhole, which had non-guided flow. The manholes were not adjacent to high hydrogen sulfide manholes or lift stations.

In Fig. 4.7 (a), manholes with multiple inlets have the highest hydrogen sulfide generation, in comparison to others bends, which is likely due to the turbulence caused by the varies flow streams converging into the manhole. In Fig. 4.7 (b), among bends, we observed that acute bends had the highest overall hydrogen sulfide manhole concentration in the headspace, but the right-angle bend had 2 instantaneous high peaks, which could have been caused due to instantaneous heavy flow into the manhole. Except for the instantaneous peaks, the right-angle bend and straight flow show similar trends. Although higher turbulence would be expected for the manhole with the right-angle bend, the velocity for the straight flow was higher (3.0 ft/sec compared to 0.2 ft/sec for the manhole with the right-angle bend), which also resulted in more turbulence. The obtuse bend has low hydrogen sulfide concentration trend in comparison to straight flow. The reason for this is because the straight flow manhole has a flow velocity of 3.0 ft/s whereas the obtuse bend manhole has a flow velocity of 1.8 ft/s.

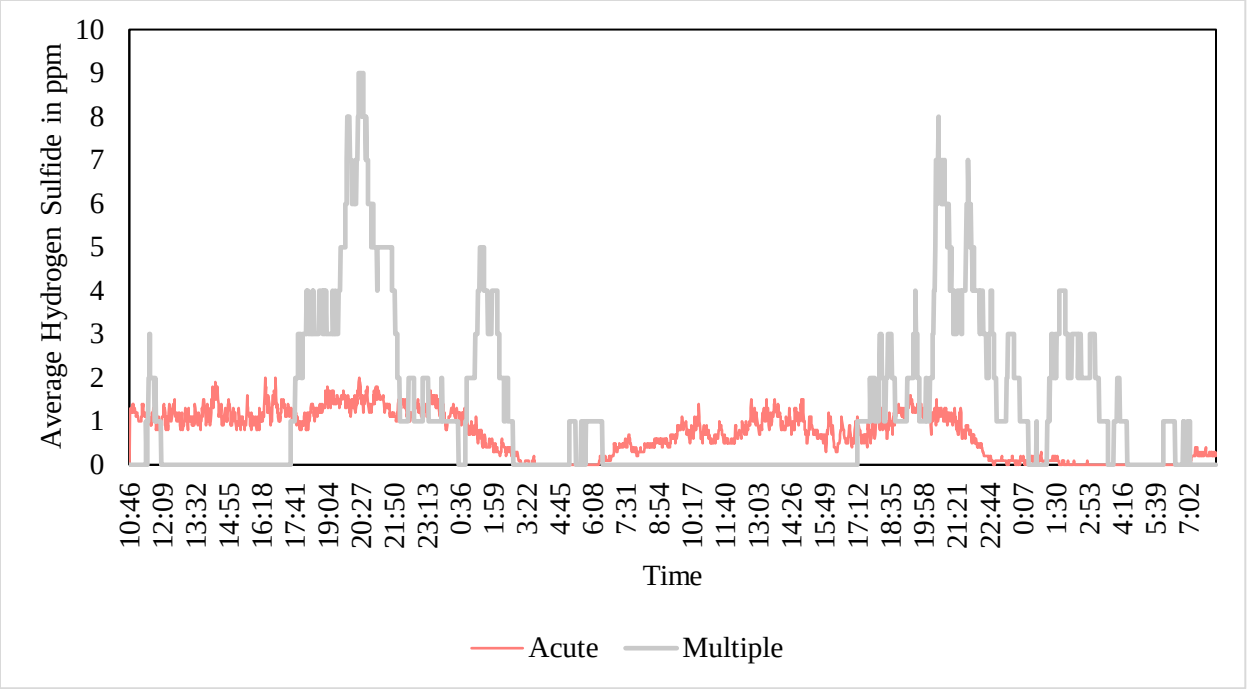


Fig 4.7a Hydrogen sulfide concentration for varied bends and flow direction

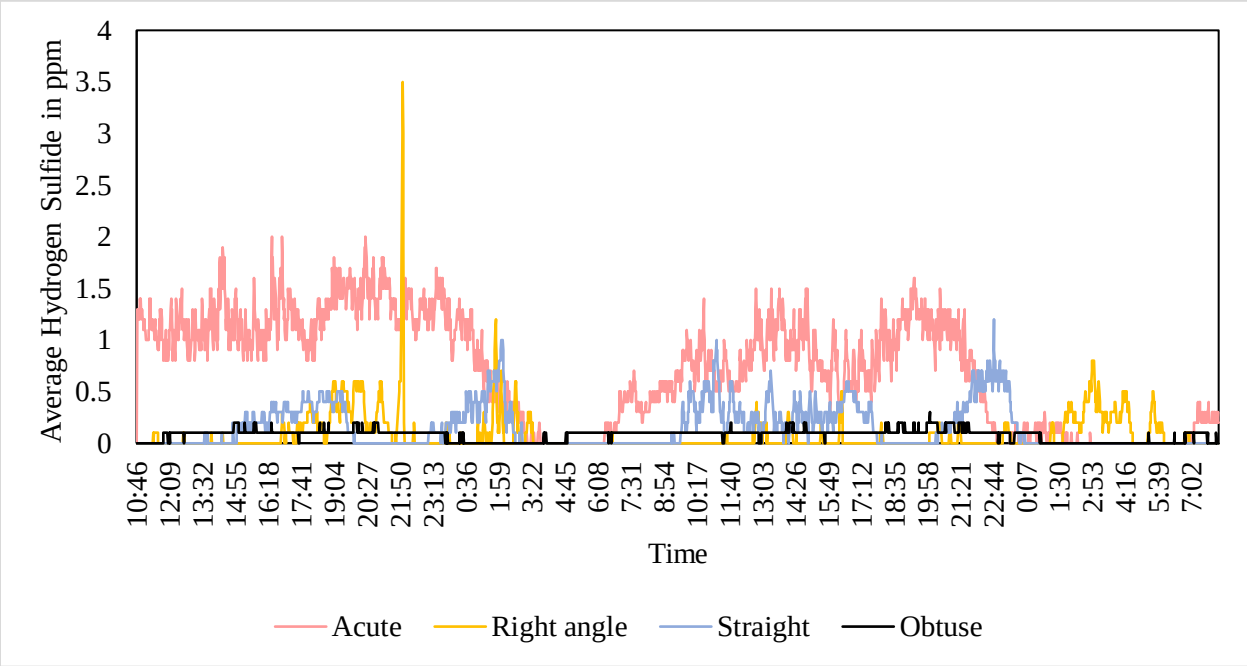


Fig 4.7b Hydrogen sulfide concentration for varied bends and flow direction

4.3.2.4 EXAMPLE EFFECT OF PRESENCE OF GUIDE ON GAS-PHASE HYDROGEN SULFIDE

Presence (walls on the side of the manhole to provide a clear flow direction to the wastewater), may result in friction and turbulence, resulting in increased hydrogen sulfide release into the manhole's headspace. On the other hand, absence of a guide may create turbulence due to splashing when wastewater enters the manhole, and due to a "sudden contraction" when flow must re-enter a pipe on the downstream side of the manhole. Also, presence of the guide may reduce the flow area thus increasing the flow velocity by the continuity equation for the fluids (Q= VA). Fig 4.8 shows example hydrogen sulfide concentrations from guided and non-guided manholes, which are similar, thus showing there is not much change with the presence of the guide. Each manhole had a single inlet at zero feet height, subcritical flow, and uniform pipe size diameter for inlet and outlet pipe. The manholes were not adjacent to a high hydrogen sulfide manhole or a lift station. The guided flow manhole had an obtuse angle bend, and the non-guided flow manhole had straight flow. Also, here a diurnal pattern is clearly visible.

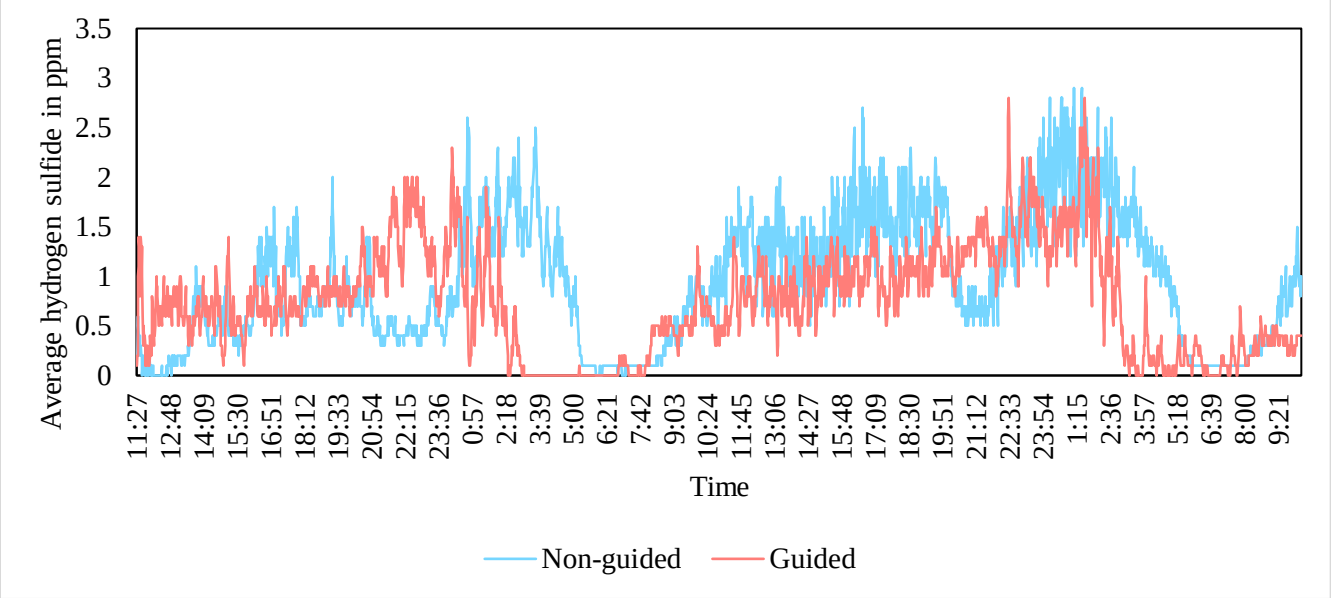


Fig 4.8 Hydrogen sulfide concentration for guided/non-guided manholes

4.3.2.5 EXAMPLE EFFECT OF FLOW TYPE/TURBULENCE ON GAS-PHASE HYDROGEN SULFIDE

Previous studies have shown any design or factor which results in change in turbulence inside the manhole causes higher hydrogen sulfide release to the manhole headspace. Thus, turbulent flow conditions are one of the primary

factors resulting in higher hydrogen sulfide release into the manhole and thus higher corrosion rates. Turbulent flows are caused due to change in slope of the manhole, change in flow velocities, change in depth, presence of drop, change in pipe size, or presence of bends. To assess the impact of flow conditions, we distributed our manholes into 3 flow categories: supercritical flow, where Froude number > 1; subcritical flow, where Froude number < 1, and hydraulic jumps (where a supercritical flow moves into subcritical flow, resulting in a loss in energy). Since the hydraulic jumps experience a change in velocity, we anticipated a higher hydrogen sulfide concentration. Fig 4.9 below shows an example of hydrogen sulfide concentrations from manholes with the 3 flow conditions. All manholes had a single inlet at zero feet height, uniform pipe size diameter for inlet and outlet pipe, and were not adjacent to high hydrogen sulfide manhole or a lift station. The manhole with the hydraulic jump had a right-angle bend; the manhole with supercritical flow manhole had guided straight flow, and the manhole with subcritical flow had straight non-guided flow.

The highest value of hydrogen sulfide recorded in the manhole with the hydraulic jumps was 10.4 ppm, in comparison to 3.1 ppm in the manhole with supercritical flow, and only 1.2 ppm in the manhole with subcritical flow. The maximum concentration of hydrogen sulfide in the hydraulic jump manhole was about 3 times larger than supercritical flow, and 10 times larger than subcritical flow. Also, it is seen that during the day the concentration of hydrogen sulfide is very high is compared to night. This is in accordance with EPA's report [(33)], as people consume more water during the day, producing higher flows which result in higher flow velocities, which in turn result in greater turbulence.

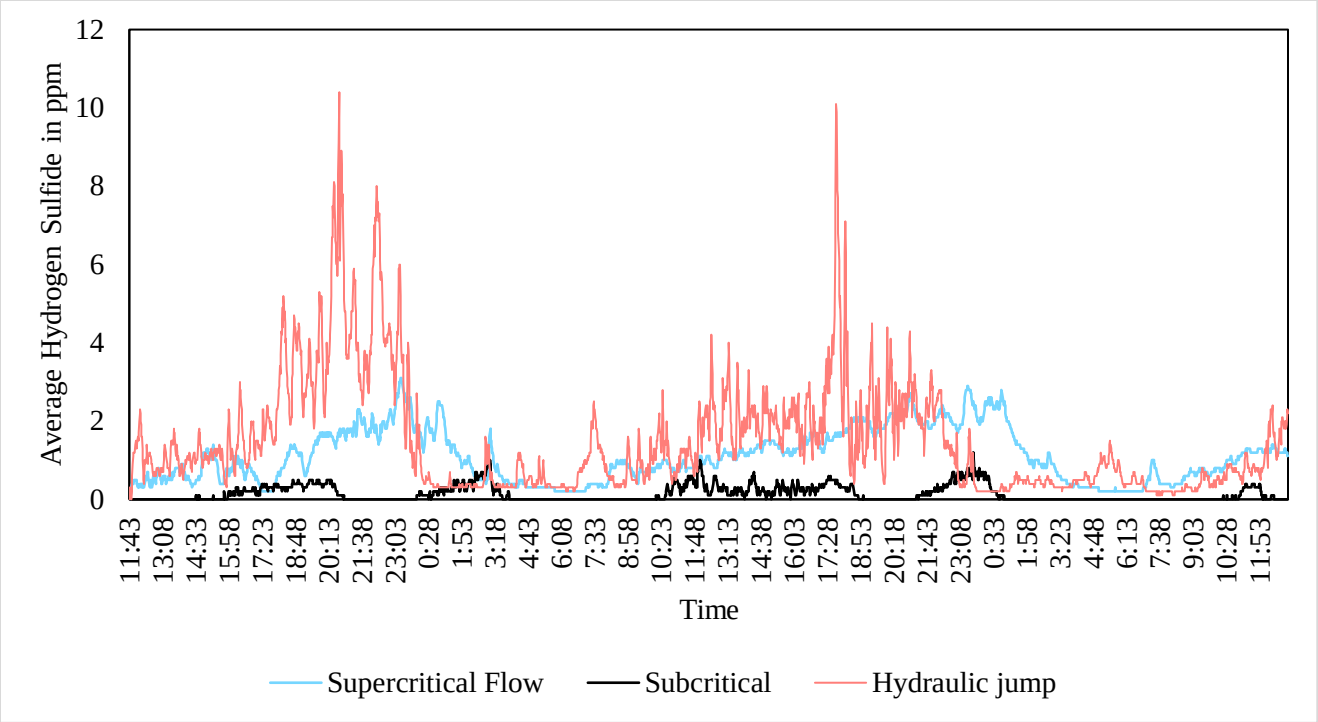


Fig 4.9 Hydrogen sulfide concentration for varied flow types in manholes

4.3.2.6 EXAMPLE EFFECT OF LOCATION OF MANHOLE ON GAS-PHASE HYDROGEN SULFIDE

Location of manhole also plays an important role in the amount of hydrogen sulfide release into the manhole's headspace. When a manhole is located downstream of a lift station, or downstream of a manhole with high hydrogen sulfide concentrations, this manhole can suffer from high hydrogen sulfide concentrations due to the discharge of accumulated hydrogen sulfide downstream, which can ultimately cause corrosion[33, 37]. From our data collection, we found that 90% of the manholes recorded average hydrogen sulfide values less than 5 ppm and maximum hydrogen sulfide values less than 10 ppm, as clearly seen via the orange pareto line in Figs. 4.10a and 4.10b. Thus, manholes which recorded maximum hydrogen sulfide value of 10 ppm or greater, or an average hydrogen sulfide value of 5 ppm or greater, were regarded as high hydrogen sulfide manholes. The manholes downstream from manholes with these high hydrogen sulfide concentrations were considered likely to record higher hydrogen sulfide concentrations.

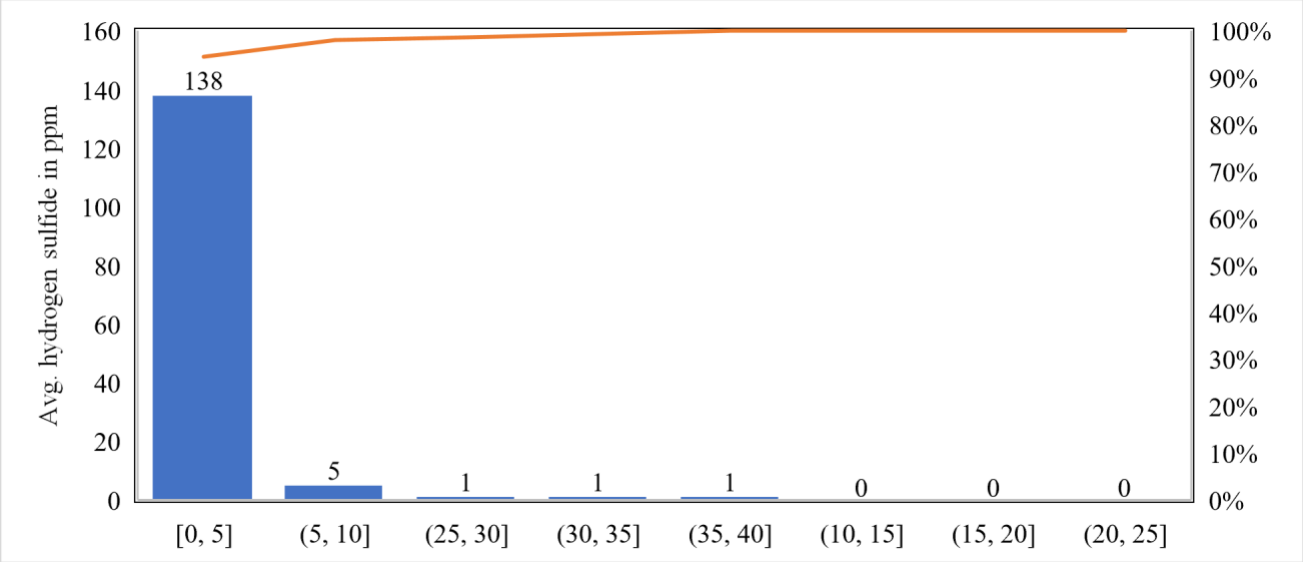


Fig 4.10a Average hydrogen sulfide distribution for 146 manholes

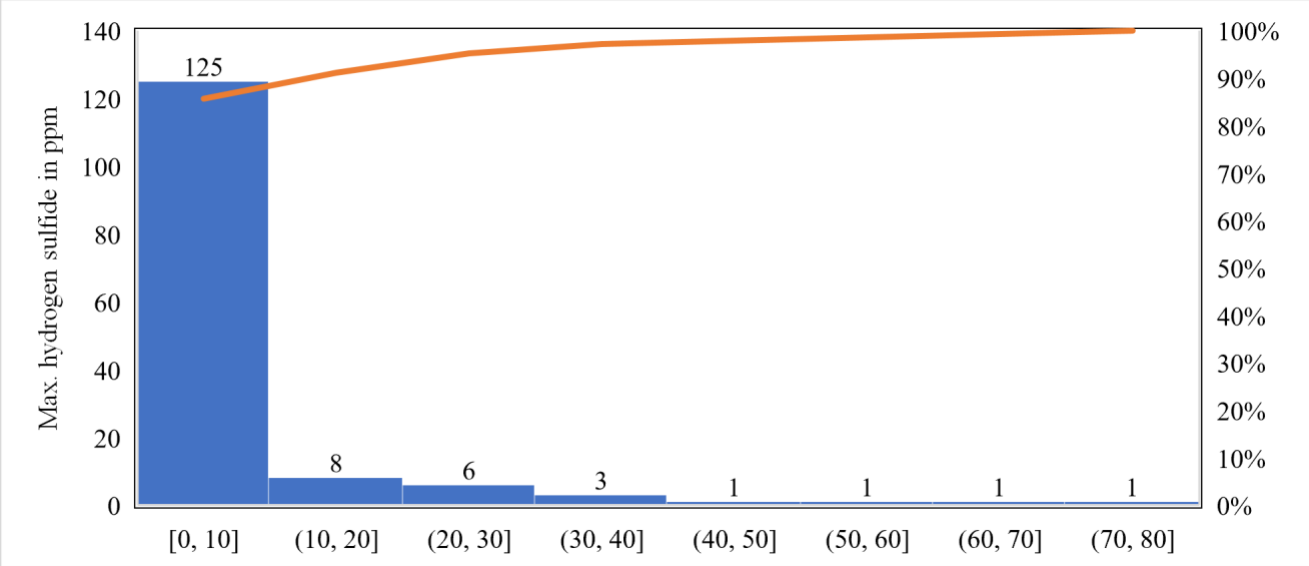


Fig 4.10b Maximum hydrogen sulfide distribution for 146 manholes

Fig 4.11 provides an example of H_2S concentrations recorded at a manhole downstream of the lift station, as well as a manhole downstream of a manhole with high H_2S concentrations. The manhole features in the example for each drop category are presented below:

a. Manhole downstream to a high hydrogen sulfide manhole – One inlet at zero feet height, from the bottom of the manhole. It had subcritical flow, uniform pipe size diameter for inlet and outlet pipe. The manhole was adjacent to a high hydrogen sulfide manhole. The flow had an obtuse angle bend. It was not adjacent to a lift station. The flow was non-guided.

b. Manhole downstream to a lift station – Two inlets at zero feet height, from the bottom of the manhole. It had subcritical flow. One inlet had a uniform pipe size diameter for inlet and outlet pipe and other inlet diameter was smaller than the outlet pipe by 3 inches. The manhole was not adjacent to high hydrogen sulfide manhole. It was adjacent to a lift station. The flow was non-guided.

It is seen that the manhole downstream of the lift station recorded the highest hydrogen sulfide concentration of 78.1 ppm, highest among the 366 manholes. The sudden dip in hydrogen sulfide values in the manhole downstream of the lift station is because lift stations release wastewater intermittently. The manholes downstream from a high hydrogen sulfide manhole recorded the highest value of 5 ppm, which is reasonably high.

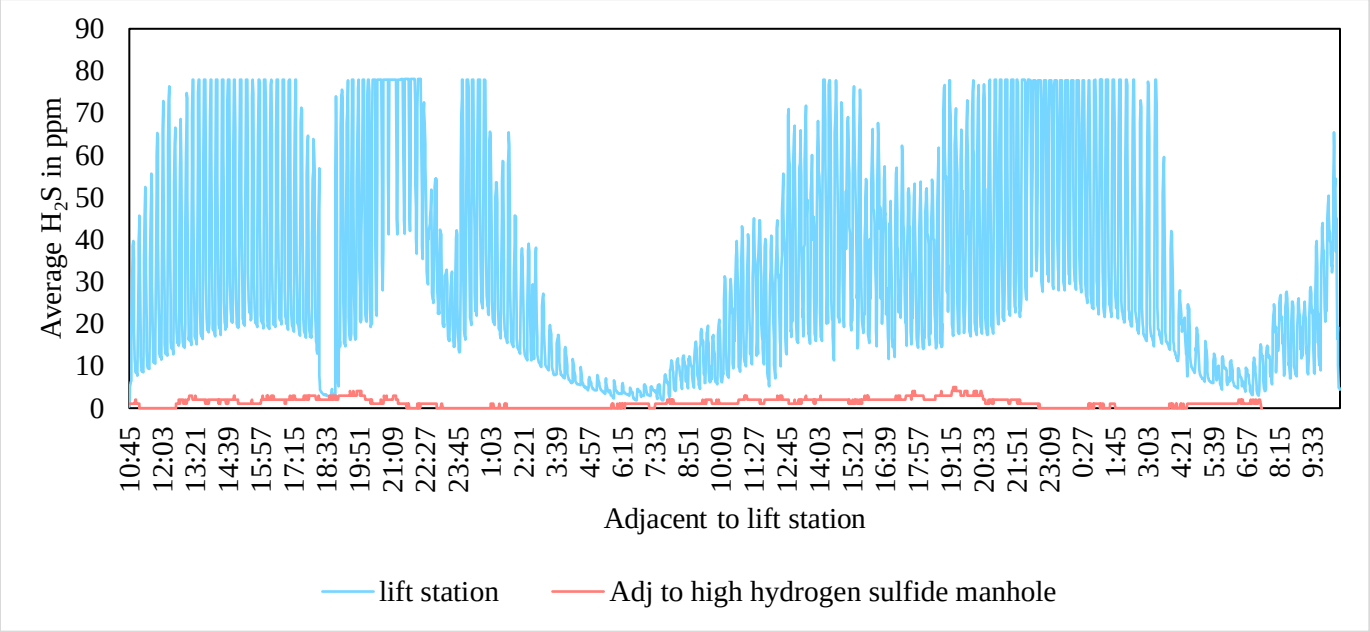


Fig 4.11 Hydrogen sulfide concentration depending on location of manholes

4.3.2.7 SUMMARY OF AVERAGE GAS-PHASE HYDROGEN SULFIDE BY MANHOLE CATEGORY

Fig 4.12 summarizes the average hydrogen sulfide concentrations for 146 manholes based on their designs. From the graph it is very clear that manholes which were downstream of lift stations recorded the highest hydrogen sulfide concentrations (32.1 ppm average), followed by manholes downstream of high hydrogen sulfide manholes (8.6 ppm average), followed by manholes upstream of high hydrogen sulfide manholes (5.1 ppm average) and hydraulic jump (4.1 ppm average). The lowest average hydrogen sulfide concentrations were recorded for acute bends, with only 0.4 ppm, followed by control and supercritical flow with only 0.7 ppm. Controls were manholes with no drop, uniform pipe size change, subcritical flow, and straight flow; thus, they were expected to have the least turbulent flows, which resulted in low hydrogen sulfide values. Only two manholes fell in the acute category, so they may not have been representative. Each manhole is a combination of drop, pipe size change, flow conditions, and presence of bends; thus, other manholes had more turbulent designs overall, resulting in higher average hydrogen sulfide in comparison to supercritical flow and acute bends.

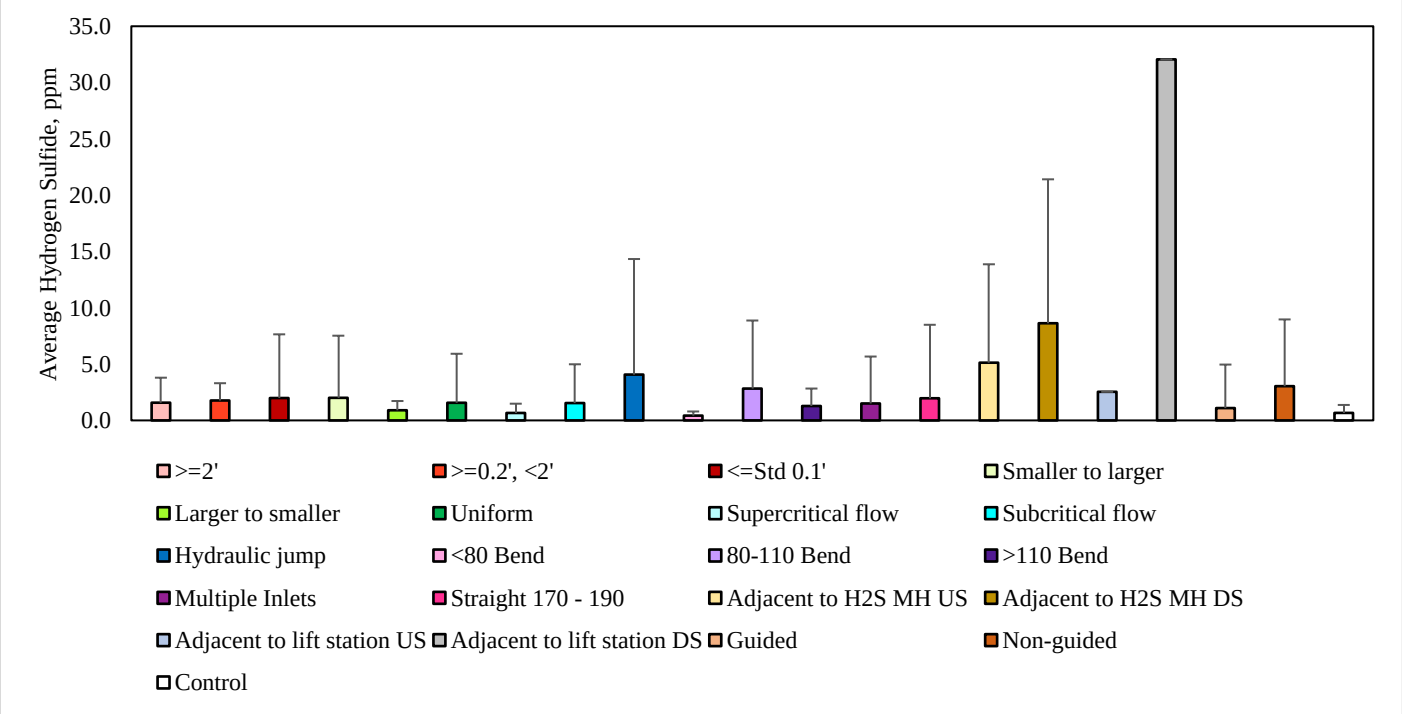


Fig 4.12 Average hydrogen sulfide concentration depending on manhole design

In conclusion, we can infer that:

a. Each manhole consists of a combination of designs; thus, a combination of designs impact hydrogen sulfide concentrations rather than just one factor, as noticed in Fig 4.11. However, radical designs such as hydraulic jump or presence of lift station could still suffer from higher concentrations of hydrogen sulfide even when other design parameters are built to reduce turbulence.
b. Manhole designs (e.g. hydraulic jumps) which result in turbulence generate more hydrogen sulfide, and this would imply that these manholes might undergo faster and severe corrosion rates.
c. Location of manhole plays an important role in determining the hydrogen sulfide concentrations in the manhole's headspace.
d. A diurnal pattern was noticed in all the designs of the manholes (Figs. 4.5-4.9, 4.11), suggesting that temperature is one of the important factors determining hydrogen sulfide concentrations in the manhole, thus higher hydrogen sulfide concentrations in the afternoon in comparison to early mornings.

4.3.3 EXAMPLES ILLUSTRATING EFFECTS OF WASTEWATER CHARACTERISTICS ON HYDROGEN SULFIDE

The relationships between the hydrogen sulfide concentration and the gas phase temperature and liquid phase parameters (temperature, pH, and dissolved oxygen) were also studied. Figs. 4.13 to 4.15 show examples of these relationships. The effects of specific parameters are discussed in the following sections.

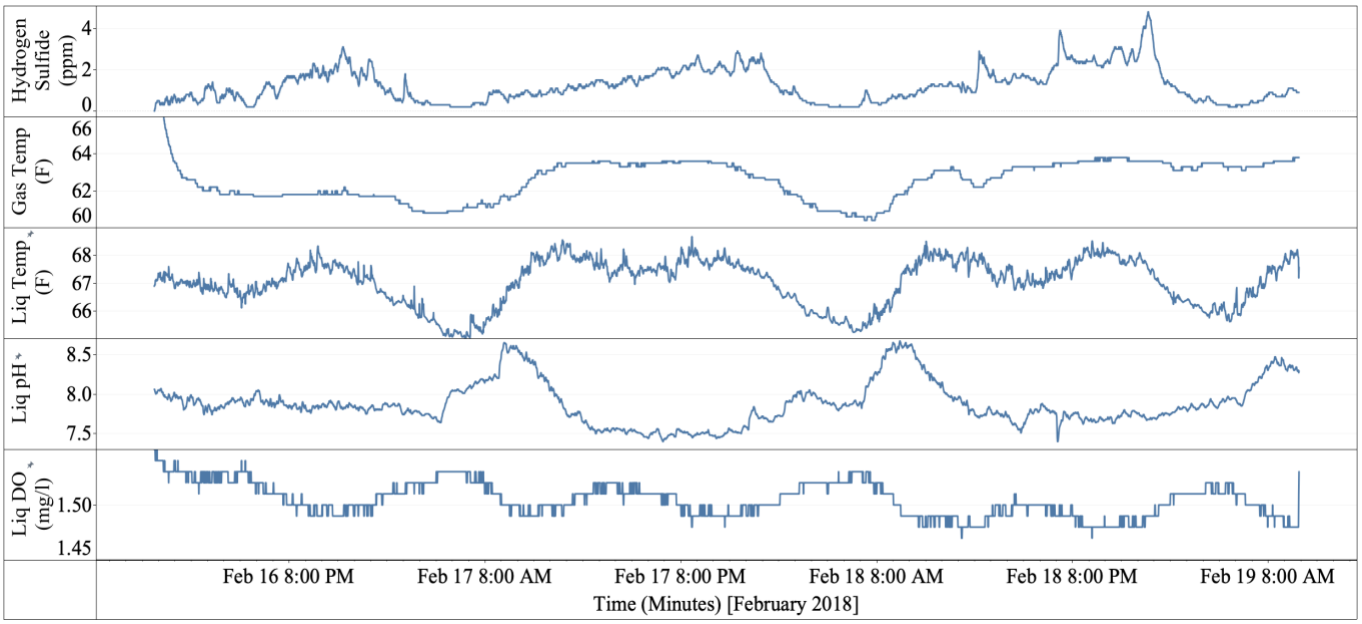


Fig 4.13 Effects of gas and liquid-phase parameters on hydrogen sulfide for manhole with supercritical flow (D14MH0192)

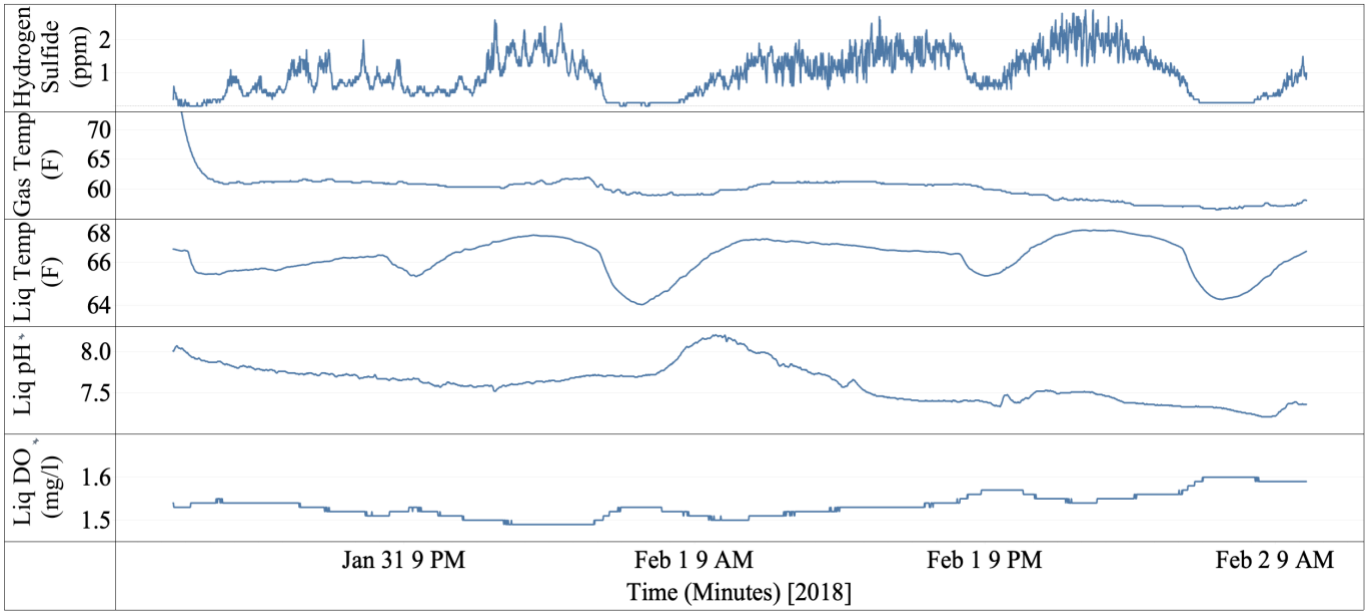


Fig 4.14 Effects of gas and liquid-phase parameters on hydrogen sulfide for manhole with trends for manhole with non-guided channel (E07MH0111)

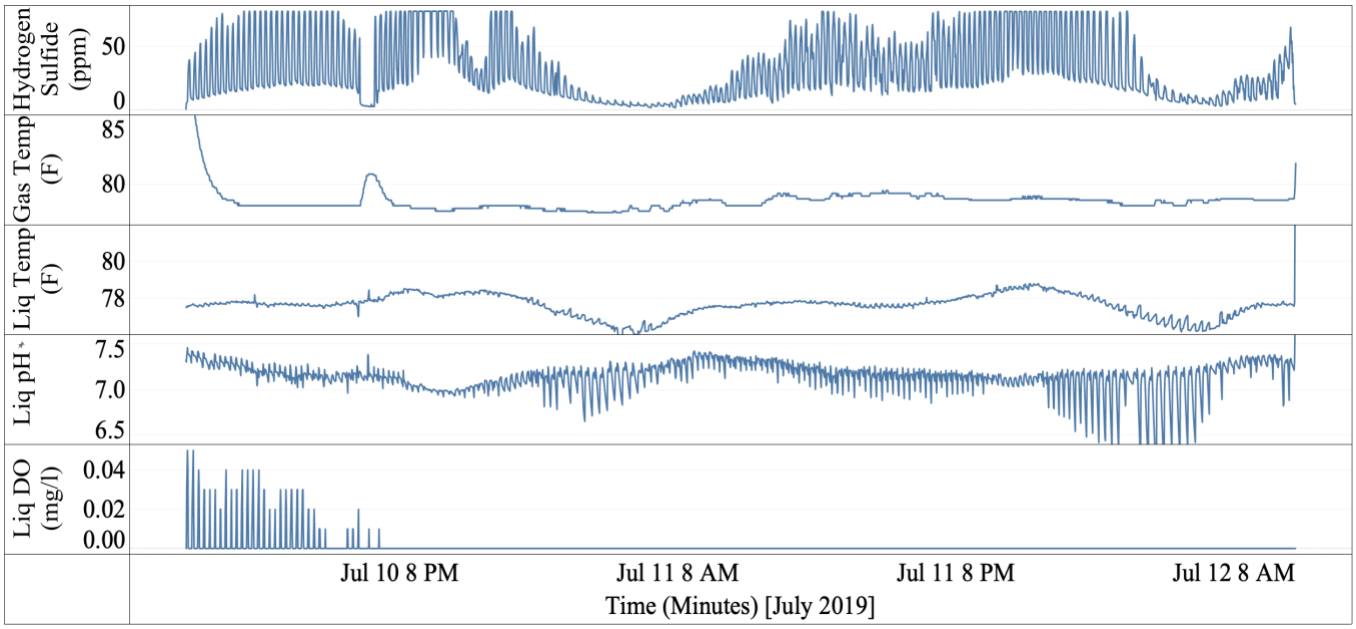


Fig 4.15 Effect of gas and liquid-phase parameters on hydrogen sulfide for manhole near lift station (B12MH0026)

4.3.3.1 EXAMPLE EFFECTS OF TEMPERATURE ON HYDROGEN SULFIDE

From Figs. 4.13 to 4.15, we can infer that there is a direct relationship between the average hydrogen sulfide concentration and average manhole gas phase temperature and liquid phase temperature. This is in accordance with the understanding that increase in temperatures result in decreased hydrogen sulfide solubility due to a decrease in Henry's law constant[37]. At higher temperatures, the hydrogen sulfide gas can easily volatilize into the manhole headspace; thus, warmer climates or higher temperatures or summer season will record higher hydrogen sulfide concentrations. A diurnal pattern is seen in the temperature due to the change in ambient temperatures, similar to patterns observed by Wells in their study of sewer systems[42]. Hydrogen sulfide and temperature values increase during the day, with peaks seen around 6 PM – 8 PM. Lowest values of hydrogen sulfide and temperature are observed in the early mornings, similar to patterns observed in Australia's sewer system[42].

In Fig 4.16a-b and c-d, we see the relation between the average gas phase temperature and average gas phase hydrogen sulfide, and average liquid phase temperature and average gas phase hydrogen sulfide, respectively. Gas phase temperature shows a 3% correlation with the gas phase hydrogen sulfide for 146 manholes (Fig 4.16a); however, when the average gas phase temperature for each design category of the manhole is compared with the

average gas phase hydrogen sulfide concentration (Fig 4.16b), we see a 45% correlation between the parameters. (When a manhole falls into multiple categories, such as high drop and supercritical flow, it is included in both of those averages.) The 45% correlation indicates that manhole design is impacting gas-phase temperature. Heat transfer occurs via turbulence in a matter similar to mass transfer; hence, manholes with greater turbulence (e.g. manholes with drops) would be expected to experience more heat transfer between the gas and liquid phases. Similarly, liquid phase temperature shows a 2% correlation with the gas phase hydrogen sulfide for 146 manholes (Fig 4.16c); however, when the average liquid phase temperature for each design category of the manhole is compared with the average gas phase hydrogen sulfide concentration (Fig 4.16d), we see a 18% correlation between the parameters. Again, manholes with greater turbulence would be expected to experience more heat transfer.

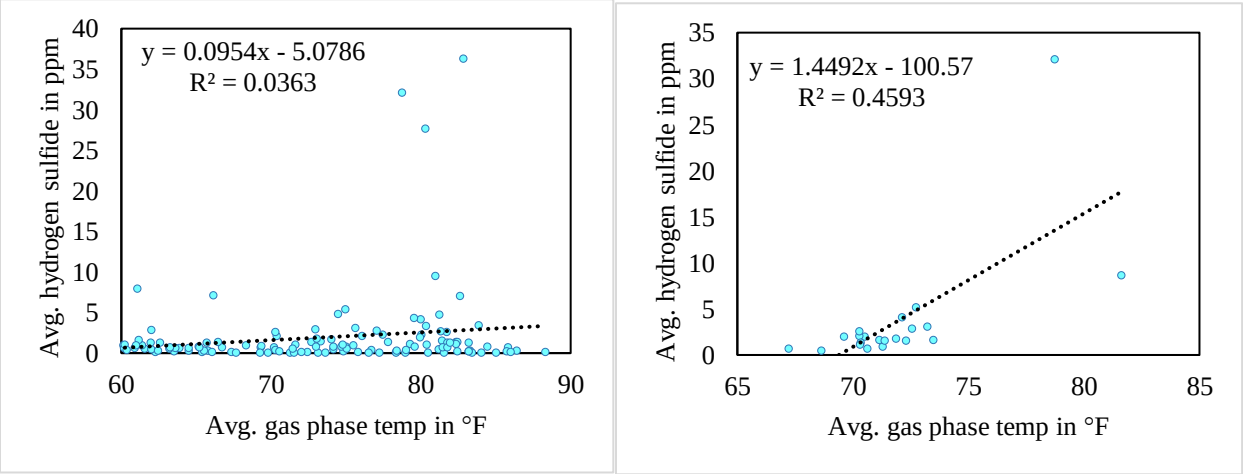


Fig 4.16a and b Average hydrogen sulfide in ppm vs average gas phase temperature in °F

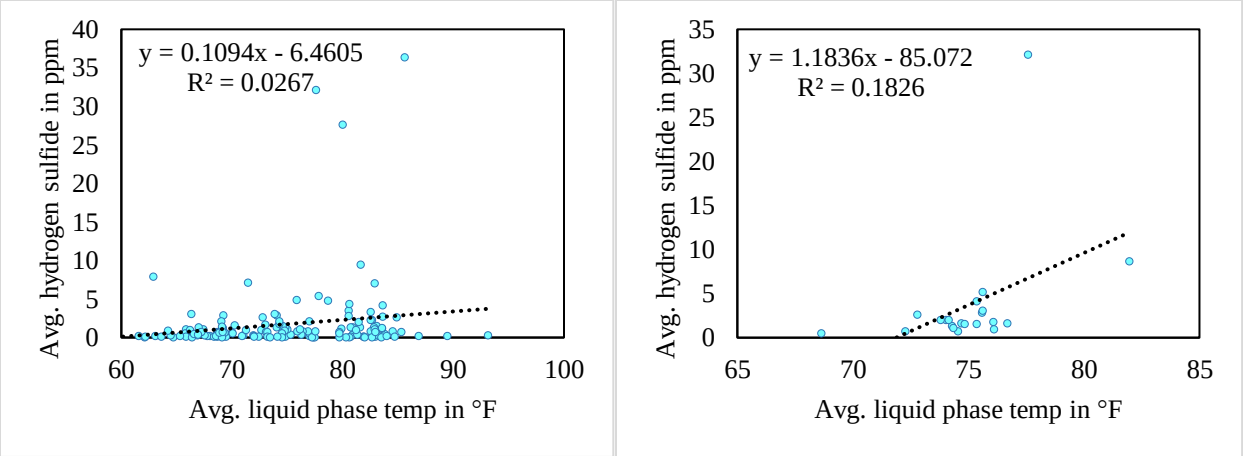


Fig 4.16c and d Average hydrogen sulfide in ppm vs average liquid phase temperature in °F

4.3.3.2 EXAMPLE EFFECTS OF pH ON HYDROGEN SULFIDE

In Fig 4.13 and 4.14, an inverse relationship is noticed between the wastewater pH and hydrogen sulfide. This is because as the pH decreases, the solubility of hydrogen sulfide in wastewater decreases[37], thus resulting in easier volatilization of the hydrogen sulfide from the liquid to gas phase. Also, in Figure 4.13 and 4.14, a diurnal pattern for liquid pH is noticed, with high pH noticed during the early morning of 5 AM to 9 AM. The reason could be the excessive use of water by the residents for the morning routines, or use of sulfate soaps which could increase pH of the wastewater.

A 3% correlation between average pH and average gas phase hydrogen sulfide for 146 manholes is seen (Fig 4.17a). Similarly, the average pH for each design category of the manhole is compared with the average gas phase hydrogen sulfide concentration (Fig 4.17b); we see only 1% correlation between the parameters, indicating no relationship between manhole design and pH.

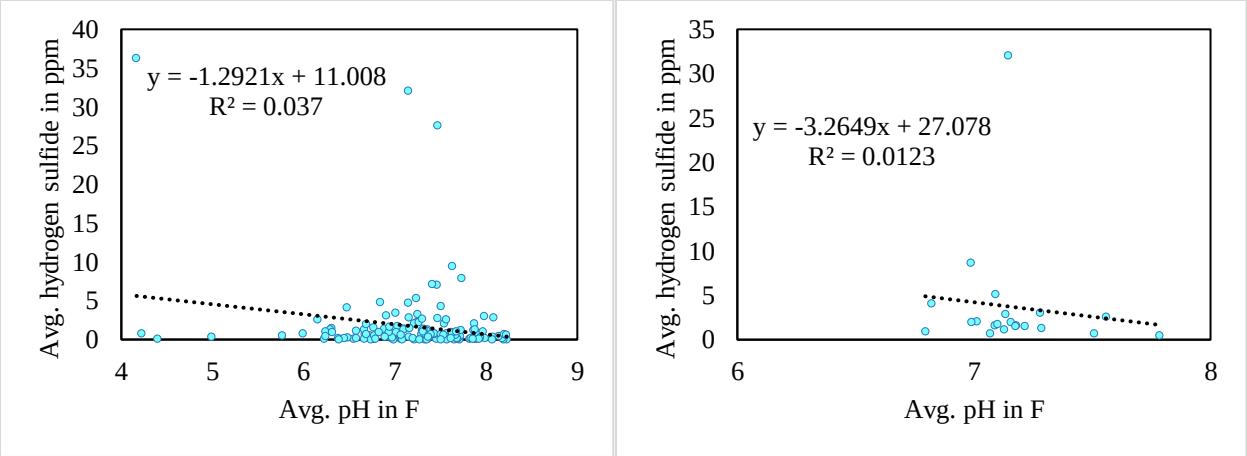


Fig 4.17a and b Average hydrogen sulfide in ppm vs average liquid phase pH

4.3.3.2 EXAMPLE EFFECTS OF DISSOLVED OXYGEN ON HYDROGEN SULFIDE

In the example figures, a direct relationship was seen between the pH and dissolved oxygen for some designs (Figs. 4.13) but not for others (Figs. 4.14, and 4.15), suggesting that design might be incorporating oxygen into the system through agitation or mixing, which has been seen in sewer system with high flow velocities or turbulent flows by EPA and ASCE[34, 35, 37]. Also, the peak of pH seems to follow the dissolved oxygen peak, with dissolved oxygen peak increasing between 1 AM – 6 AM, and peak of pH increasing between 5 AM to 9 AM.

In Fig. 4.18a we a correlation of 0.6% between the average dissolved oxygen and average gas phase hydrogen sulfide; however, in Fig 4.18b a correlation of 24% is seen between the average dissolved oxygen between the various design categories of the manhole and average hydrogen sulfide is seen. It confirms our previous assumption that manhole's design might be incorporating oxygen into the system. Fig 4.18c shows the variations in dissolved oxygen concentrations based on manhole design. We can clearly see that an acute bend is showing the highest dissolved oxygen concentrations of 4.4 mg/l; also, supercritical flow, smaller to larger pipe size change and larger to smaller pipe size change show values higher than the average recorded DO concentration of 1.5 mg/l (Table 4.8). The lowest recorded dissolved oxygen concentrations was for the manhole downstream of lift station; this seems reasonable due to the detention of wastewater and solids settling out in lift station, resulting in anaerobic conditions with relatively low oxygen[33, 37].

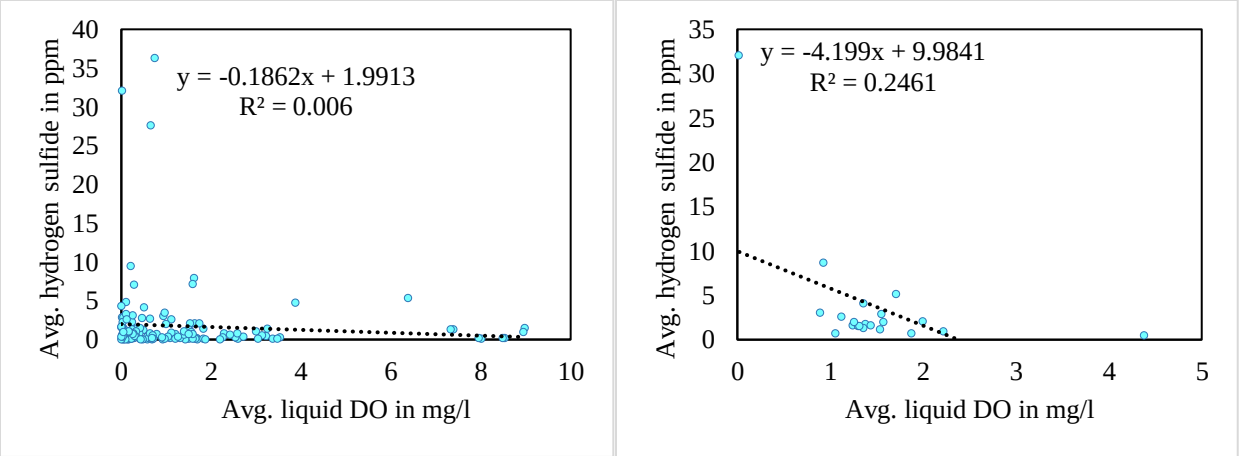


Fig 4.18a and b Average hydrogen sulfide in ppm vs average liquid DO in mg/l

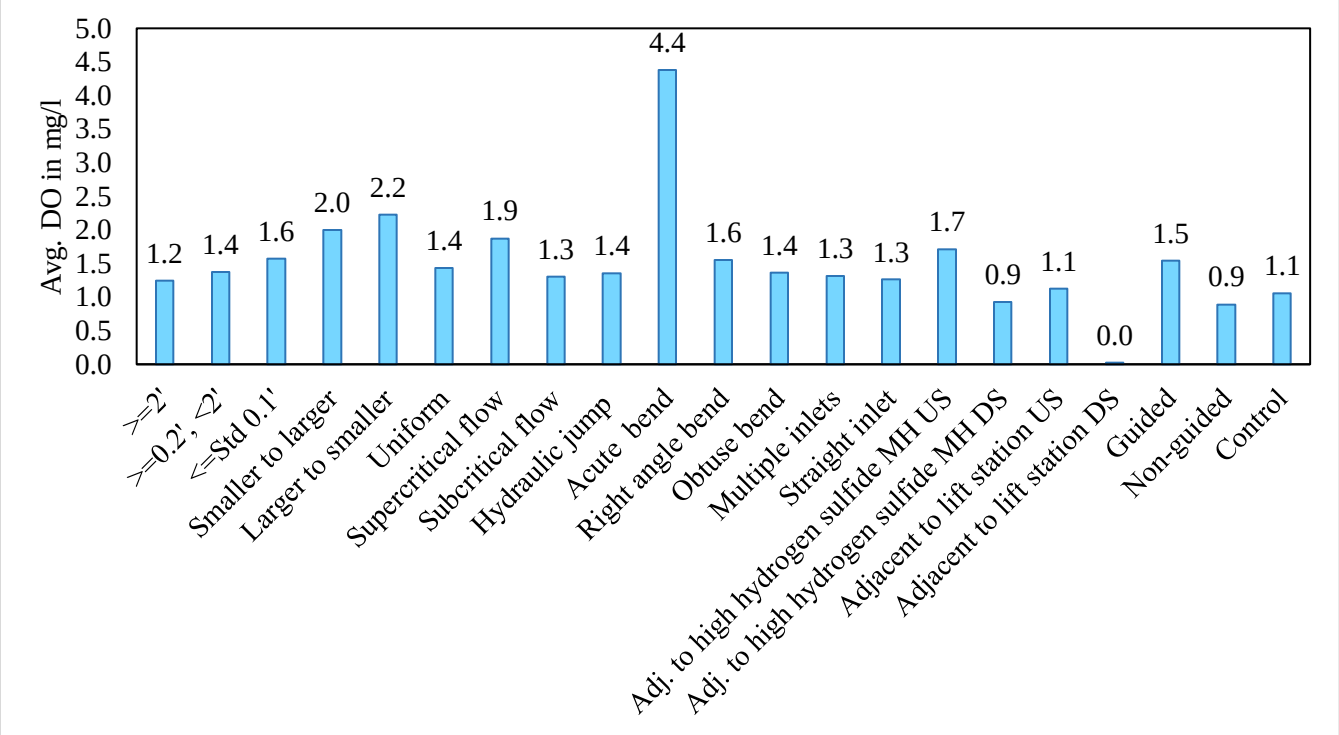


Fig 4.18c Average dissolved oxygen for various manhole categories

4.3.4 EFFECTS OF SULFIDE, SULFATE, AND BOD ON HYDROGEN SULFIDE

A correlation of 0.1%, 0.03%, and 0.1% between average sulfide, average sulfate and average BOD with average hydrogen sulfide for 146 manholes is seen (Fig 4.19 a, c, and e). A reason for such a low correlation could be due to

other factors such as temperature, manhole design, and dissolved oxygen playing a greater role in determining the hydrogen sulfide generation. However, an increase in correlation is seen when the manholes are grouped into various design categories, with sulfide, sulfate, and BOD showing 19, 18%, and 17% correlations with average hydrogen sulfide (Fig. 4.19 b, d, and f), respectively. Manhole design impacts turbulence, which impacts stripping of hydrogen sulfide and oxygen into the gas phase. When oxygen is transferred to the gas phase, this leads microorganisms to use oxygen chemically-bound in sulfate, converting it to sulfide; this would decrease liquid-phase sulfate and increase sulfide. The amount of BOD remaining in the wastewater in turn depends on the amount of DO and sulfate available for oxidation. Other factors such as temperature are also playing a role, as will be discussed in Ch. 5.

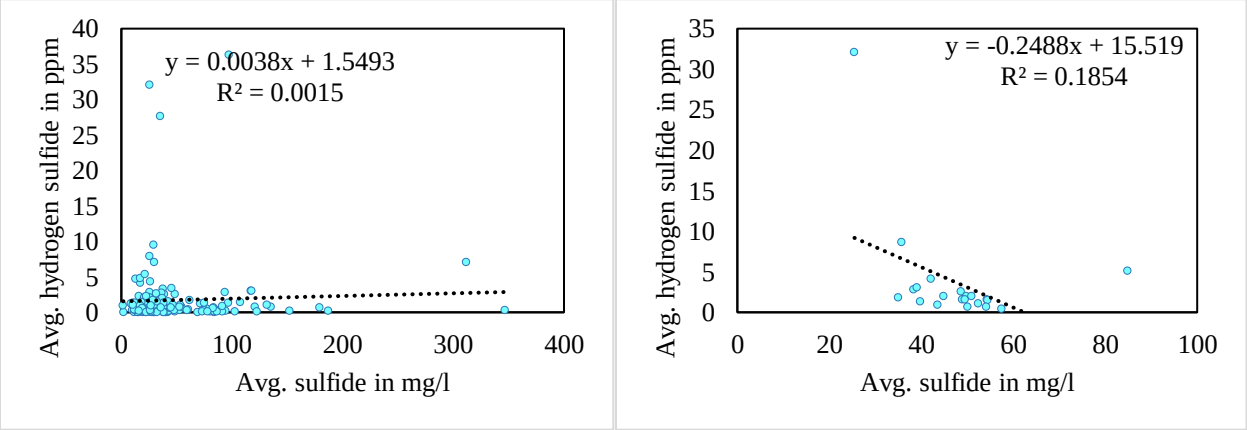


Fig 4.19a and b Average hydrogen sulfide vs. Average sulfide

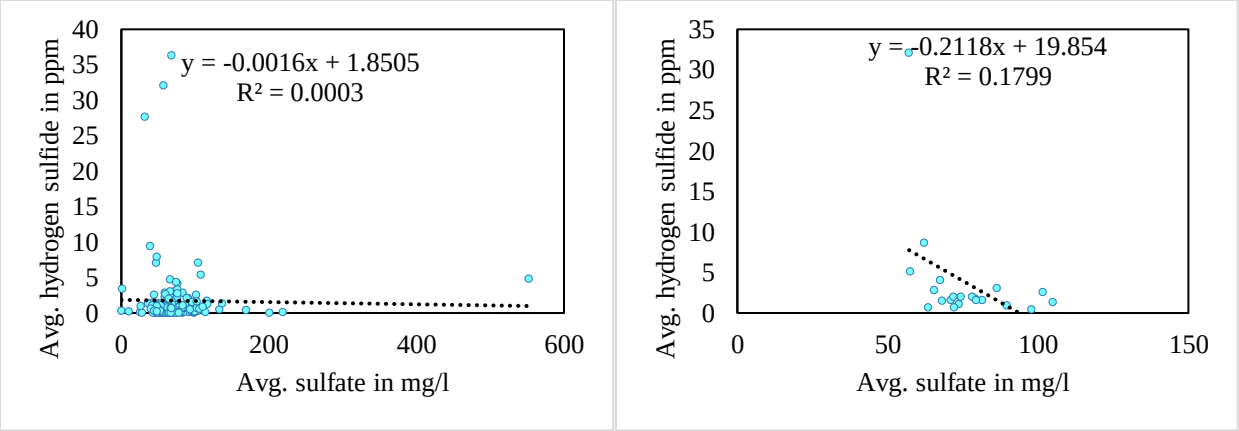


Fig 4.19c and d Average hydrogen sulfide vs. Average sulfate

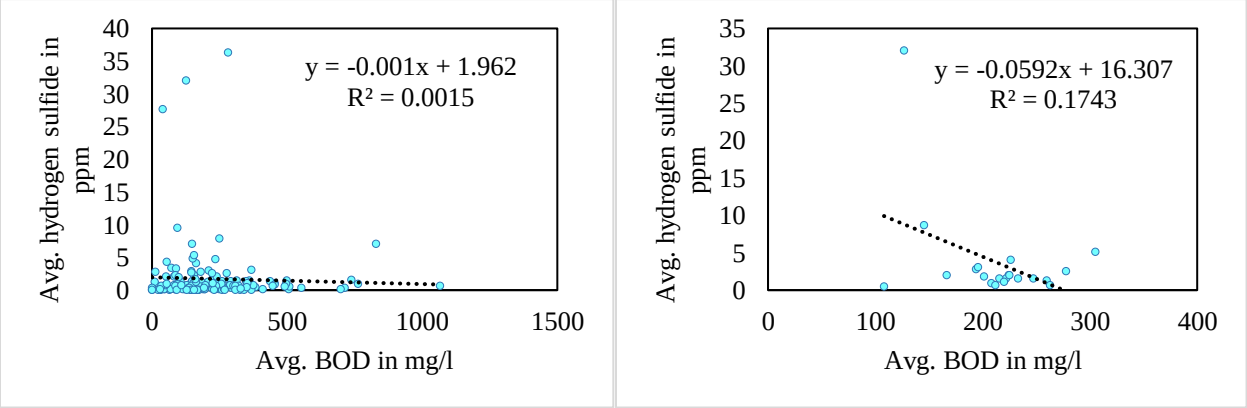


Fig 4.19e and f Average hydrogen sulfide vs. Average BOD

4.3.5 EFFECT OF SEASONS ON HYDROGEN SULFIDE

Since Texas is a warmer state with hot summers, it becomes important to understand the effect of ambient temperatures on hydrogen sulfide volatilization and increased corrosion rates. The section below will focus on the impact of seasons on gas-phase hydrogen sulfide levels in the manhole. Table 4.6 shows the distribution of the manholes in different seasons and Table 4.7 shows the average reported values of the gas and liquid parameters in different seasons.

Table 4.6 Distribution of manholes based on seasons

Seasons	Start date	End Date	Manhole in each season
Winter	21-Dec	19-Mar	80
Summer	20-Jun	21-Sep	62
Fall	22-Sep	20-Dec	86
Spring	20-Mar	19-Jun	64

Table 4.7 Average gas and liquid phase parameters based on seasons

	Ambient	Gas-Phase		Liquid-Phase					
Season	Temp. (°F)	Temp. (°F)	H_2S (ppm)	Temp. (°F)	DO (mg/l)	pH	Sulfide (mg/l)	Sulfate (mg/l)	BOD (mg/l)
Summer	86.82	81.34	3.53	82.25	0.66	7.16	39.01	58.12	168.16
Fall	62.56	76.37	0.99	78.64	1.32	8.73	55.60	72.08	217.28
Winter	47.95	59.60	0.87	66.49	1.80	7.54	50.66	81.02	249.96
Spring	69.23	69.79	2.07	73.14	1.67	6.84	46.37	91.24	242.84

In Fig. 4.20 we can clearly see the effect of seasons on the average gas phase hydrogen sulfide. Manholes sampled in summer or spring recorded the highest gas phase hydrogen sulfide concentrations of 32.05 ppm and 36.28 ppm respectively, with summer showing the highest average gas phase hydrogen sulfide concentration of 3.53 ppm. The lowest average gas phase hydrogen sulfide concentration of 0.87 ppm was recorded for the manholes done during spring (Table 4.9).

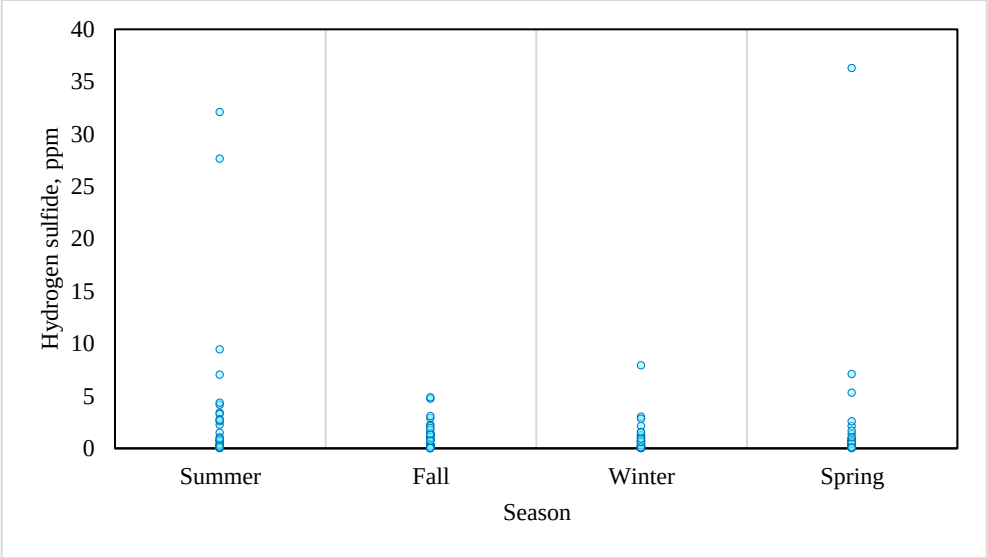


Fig 4.20 Average hydrogen sulfide generated every season

From Fig. 4.21 we can infer that there is a seasonal relationship between the average hydrogen sulfide concentration and average ambient temperatures in different seasons, with summer recording higher gas phase hydrogen sulfide values in comparison to fall and winter. This is in accordance with the understanding that an increase in temperatures results in decreased hydrogen sulfide solubility due to a decrease in Henry's law constant[37]. In addition, a correlation of 88.9% between the average ambient temperatures and average hydrogen sulfide in confirms our results about the effect of seasons on hydrogen sulfide generation.

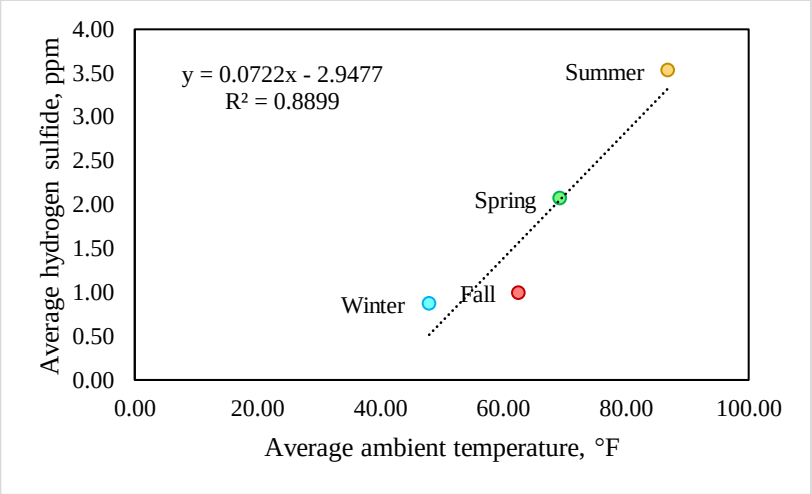


Fig 4.21 Average hydrogen sulfide generated every season

It is seen in Table 4.8 that the gas phase temperature and ambient temperature, and gas phase temperature and liquid phase temperature varies the highest during winter and least during the summer, with average variation between liquid phase temperature and gas phase temperature of 3.48 degrees. This result is like Romanov, who reported a 3.5°C variation between the manhole ambient temperature and liquid temperature[(45)].

Table 4.8 Gas and liquid temperatures for 146 manholes based on seasons, °F

Season	Ambient Temp., (°F)	Gas Phase Temp., (°F)	Liquid Phase Temp., (°F)	Gas Phase - Ambient Temp. (°F)	Liq. Phase - Ambient Temp. (°F)	Liq. Phase - Gas Phase Temp. (°F)
Summer	86.82	81.34	82.25	-5.47	-4.57	0.91
Fall	62.56	76.37	78.64	13.81	16.08	2.26
Winter	47.95	59.60	66.49	11.65	18.54	6.89
Spring	69.23	69.79	73.14	0.57	3.91	3.35
Average	65.17	71.39	74.87	6.22	9.70	3.48

4.4 EFFECT OF MANHOLE DESIGN ON SULFIDE, SULFATE, BOD, AND DISSOLVED OXYGEN, AND CORRELATION AMONG THE PARAMETERS

For the various manhole design categories, the average sulfide, sulfate, and BOD values are shown in Fig 4.22. Average sulfide values for different manhole designs ranged from 25.5 mg/l to 84.9 mg/l, with an overall average sulfide value of 47.6 mg/l. The lowest sulfide value was recorded for manholes downstream from lift stations, which suggests that turbulent flow due to the lift station resulted in excessive hydrogen sulfide volatilization to the gas phase (high value of 32.1 ppm), lowering liquid-phase sulfide levels.

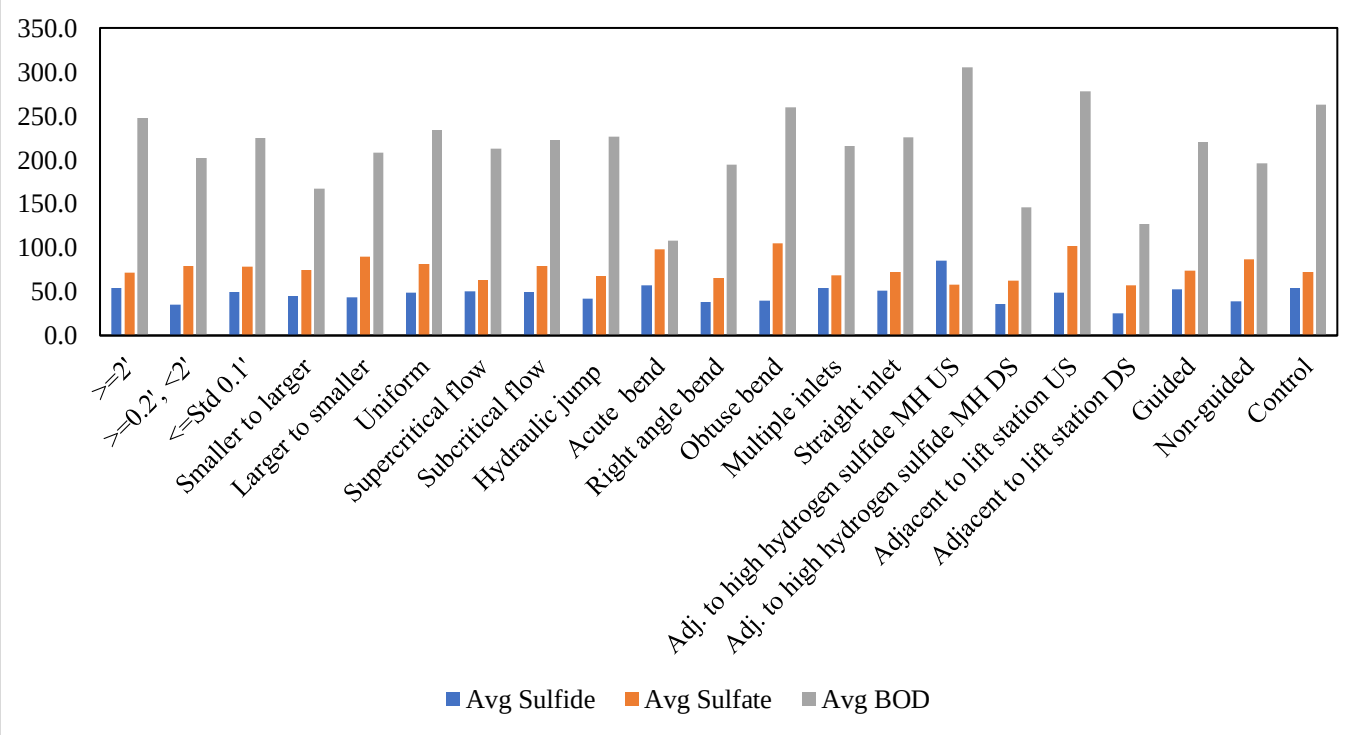


Fig 4.22 Average sulfide, sulfate, and BOD by manhole category

Average sulfate values for different manhole designs ranged from 57 mg/l to 104 mg/l, with an overall average sulfate value of 76.4 mg/l. These values are similar to values measured in municipal wastewater by ASCE, which lies between 30 mg/L to 250 mg/L[34]. The highest average sulfate value of 104.9 mg/l was recorded for right angle bends and lowest value of 57.1 mg/l for manholes downstream of the lift stations. The manhole upstream of the lift station recorded the second highest sulfate value of 101.6 mg/l. Due to absence of DO at the lift station and the high

oxygen demand (second highest BOD at 278.2 mg/l), microorganisms used chemically-bound sulfate for respiration, which reduced the sulfate present upstream of the lift station to sulfide. This generated sulfide was volatilized into gas phase hydrogen sulfide; thus, low concentrations of sulfide were seen in the manhole downstream of the lift station, along with high hydrogen sulfide concentrations in gas phase, as described previously.

Average BOD values for different manhole designs ranged from 237 mg/l to 305 mg/l, with an overall average BOD value of 213.5 mg/l. The highest BOD value of 305.4 mg/l was recorded for manholes upstream of the high hydrogen sulfide manholes, and lowest value of 108.1 mg/l for acute bends.

The manholes upstream of a high hydrogen sulfide manhole recorded the highest average BOD and sulfide concentrations of 305.4 mg/l, and 84.9 mg/l respectively, but low average sulfate concentrations of 57.6 mg/l, which would imply that the sulfate in the upstream manhole is converted to sulfide in the high BOD environment. In addition, the manhole downstream of high hydrogen sulfide manholes recorded low sulfide and BOD values, a trend similar to seen in manhole downstream to lift station. Here also the generated sulfide is likely volatilized into gas phase hydrogen sulfide, thus lowering concentrations of sulfide seen in the manhole downstream of the lift station and increasing hydrogen sulfide concentration seen in the gas phase. The reason for low BOD in manholes downstream of high hydrogen sulfide manholes could be the addition of chemically-bound oxygen from sulfate, which forms sulfide.

Figures 4.23-4.26 show correlations among sulfide, sulfate, BOD, and dissolved oxygen for different design categories. A correlation of only 1% is seen between the average sulfate and average BOD (Fig 4.20c). Both depend on the composition of the wastewater itself. In Fig. 4.20d, a correlation of 29.5% is seen between sulfide and BOD, which makes sense: a high oxygen demand makes it more likely that microorganisms will need to rely on chemically-bound oxygen for respiration, and thus generate sulfide from sulfate. In Fig. 4.20a, a correlation of only 1% is seen between sulfide and sulfate, which is surprising, but decomposition of sulfate to sulfide also depends on various other factors such as temperature and microbial activity. Surprisingly, a weak correlation of 7% was seen between average BOD and average DO as seen in Fig. 4.20b, thus suggesting involvement of other factors such as temperature and pH which affect dissolved oxygen concentration, as discussed above.

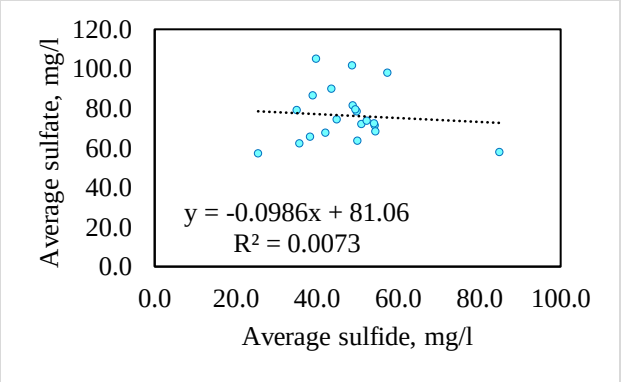


Fig 4.23 Average sulfate vs. average sulfide.

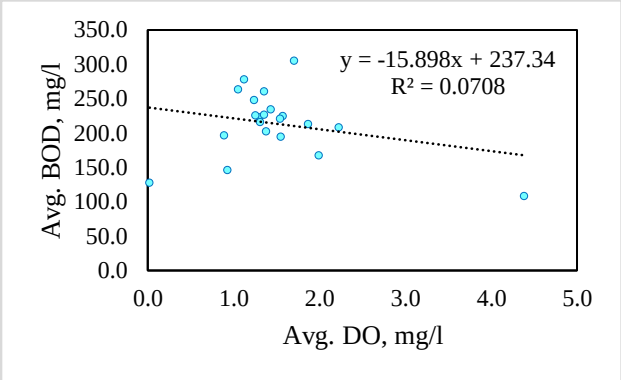


Fig 4.24 Average BOD vs. average DO

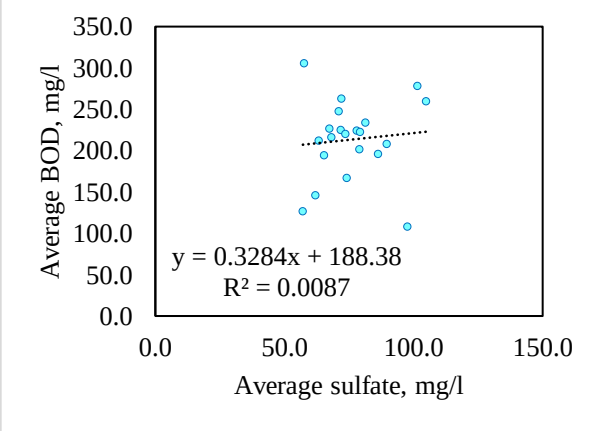


Fig 4.25 Average BOD vs. average sulfate

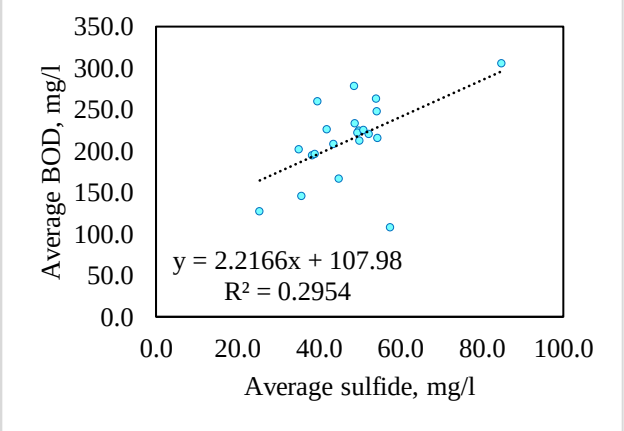


Fig 4.26 Average BOD vs. average sulfide

4.5 GAS- AND_LIQUID-PHASE DATA SUMMARY FOR COMPLETE MANHOLE DATA SETS

Table 4.9 summarizes gas and liquid-phase data for the 146 manholes for which complete data sets were available. Liquid temperatures are the same or slightly higher than gas phase temperatures for all the manhole design categories as seen in Table 4.8, because wastewater temperatures are usually warmer due to increased use of hot water by residential areas[41]. The average dissolved oxygen concentrations in the 146 manholes range from 0 – 4 mg/l, with an average of 1.5 mg/l, which is above the recommended 1 mg/l of oxygen concentration to reduce anaerobic conditions in the manhole[33, 34] The low 0.0 mg/l of dissolved oxygen value was recorded for manholes downstream from lift stations, thus showing the impact lift stations can have on the anaerobic conditions of the wastewater. The measured pH values varied between 6.8 to 7.6, with an average pH of 7.2, which is usual for

wastewater, as fecal pH and fresh urine has a neutral pH[33,45, 57]. Trends in sulfide, sulfate, and BOD were already discussed in Section 4.4.

Table 4.9 Gas and Liquid Parameter Summary for 146 manholes

Manhole Category	Ambient Temp. (F)		Gas Phase				Liquid Phase											
			Temp. (°F)		H_2S (ppm)		Temp. (°F)		DO (mg/l)		pH		Sulfide (mg/l)		Sulfate (mg/l)		BOD (mg/l)	
	Avg	STDEV	Avg	STDEV	Avg	STDEV	Avg	STDEV	Avg	STDEV	Avg	STDEV	Avg	STDEV	Avg	STDEV	Avg	STDEV
>=2'	66.5	15.4	73.5	8.6	1.6	2.2	76.7	6.4	1.2	1.2	7.1	1.1	54.3	66.5	71.1	22.3	247.4	179.4
>=0.2', <2'	68.8	14.9	71.9	6.4	1.8	1.5	76.1	5.1	1.4	1.9	7.1	0.5	35.0	18.3	79.2	19.3	201.8	75.5
<=Std 0.1'	64.2	15.9	70.5	9.6	2.0	5.7	74.0	7.1	1.6	2.2	7.2	0.7	49.8	48.1	78.2	59.7	224.5	184.8
Smaller to larger	65.1	19.2	70.3	10.0	2.0	5.5	73.8	7.0	2.0	2.4	7.0	1.0	44.9	32.2	74.3	25.9	166.9	122.7
Larger to smaller	66.8	15.7	71.3	8.6	0.9	0.8	76.1	4.6	2.2	2.9	6.8	1.2	43.6	29.6	89.8	27.2	208.1	151.9
Uniform	65.5	15.8	71.2	9.5	1.6	4.4	74.7	7.4	1.4	2.1	7.2	0.6	48.9	48.9	81.6	62.7	233.6	195.8
Supercritical flow	63.5	12.1	70.7	8.3	0.7	0.8	74.6	6.1	1.9	2.3	7.5	0.4	50.0	44.0	63.5	21.2	212.4	178.1
Subcritical flow	65.0	17.6	71.4	9.6	1.5	3.4	74.8	7.1	1.3	1.9	7.2	0.6	49.6	51.7	79.4	56.2	222.3	179.3
Hydraulic jump	67.9	14.3	72.1	9.3	4.1	10.3	75.4	7.6	1.4	2.1	6.8	1.2	42.1	27.6	67.5	22.2	226.3	139.0
Acute bend	62.0	36.8	68.6	2.7	0.4	0.4	68.6	0.1	4.4	5.8	7.8	0.6	57.5	37.0	97.8	3.2	108.1	88.2
Right angle bend	67.4	18.6	72.6	10.1	2.8	6.0	75.6	8.2	1.6	2.6	7.1	0.5	38.4	23.1	65.5	21.6	194.4	186.8
Obtuse bend	64.5	14.4	70.3	9.6	1.3	1.6	74.3	6.8	1.4	2.4	7.3	0.6	39.7	30.0	104.9	95.7	259.9	189.0
Multiple inlets	65.6	17.1	72.3	9.7	1.5	4.2	75.4	7.2	1.3	1.6	7.2	0.7	54.4	61.1	68.2	31.0	216.0	154.0
Straight inlet	63.4	15.1	69.6	8.2	2.0	6.5	74.2	6.1	1.3	1.5	7.0	0.8	50.9	42.1	71.9	26.0	225.2	196.1
Adj. to high hydrogen sulfide MH US	65.0	18.3	72.8	12.0	5.1	8.7	75.6	8.2	1.7	2.8	7.1	0.7	84.9	93.5	57.6	19.4	305.4	237.4
Adj. to high hydrogen sulfide MH DS	79.7	10.5	81.6	1.9	8.6	12.8	82.0	2.2	0.9	1.2	7.0	1.1	35.7	23.0	62.1	18.3	145.7	105.5

Adjacent to lift station US	79.2	0.0	70.3	0.0	2.5	0.0	72.8	0.0	1.1	0.0	7.6	0.0	48.7	0.0	101.6	0.0	278.2	0.0
Adjacent to lift station DS	84.2	0.0	78.8	0.0	32.1	0.0	77.6	0.0	0.0	0.0	7.1	0.0	25.5	0.0	57.1	0.0	127.0	0.0
Guided	62.4	16.5	70.3	9.8	1.1	3.9	74.4	7.0	1.5	2.0	7.1	0.8	52.4	48.7	73.8	32.5	220.3	158.2
Non-guided	71.5	15.8	73.2	8.2	3.0	5.9	75.6	6.8	0.9	1.2	7.3	0.5	39.0	29.6	86.4	90.0	196.2	193.1
Control	58.4	15.7	67.2	7.9	0.7	0.7	72.3	5.6	1.1	1.1	7.1	0.5	54.2	36.7	72.2	29.2	263.0	236.4
Average	67.5	15.2	71.9	7.6	3.7	4.1	75.0	5.6	1.5	2.0	7.2	0.7	47.6	37.7	76.4	32.6	213.5	150.1

4.6 CORROSION RATE AND FACTORS

The subsequent sections will discuss corrosion rate determination and the impact of gas phase and liquid phase parameters on the corrosion rate.

4.6.1 CORROSION RATE DETERMINATION

Manholes usually have a conical funnel shape extending for the top 1 to 2 feet, followed by cylindrical shape, as shown in Fig. 4.27. To remove the conical funnel shape, datapoints were removed for which dx < 0.90 dx_{max} and dy < 0.90 dy_{max}, as shown in the Fig 4.28, where dx and dy are diameters measured in the x and y directions, respectively. For corrosion rate determination, the bottom 20% of the datapoints of manhole shaft was used as most corroded, and remaining datapoints as the original diameter of the manhole.

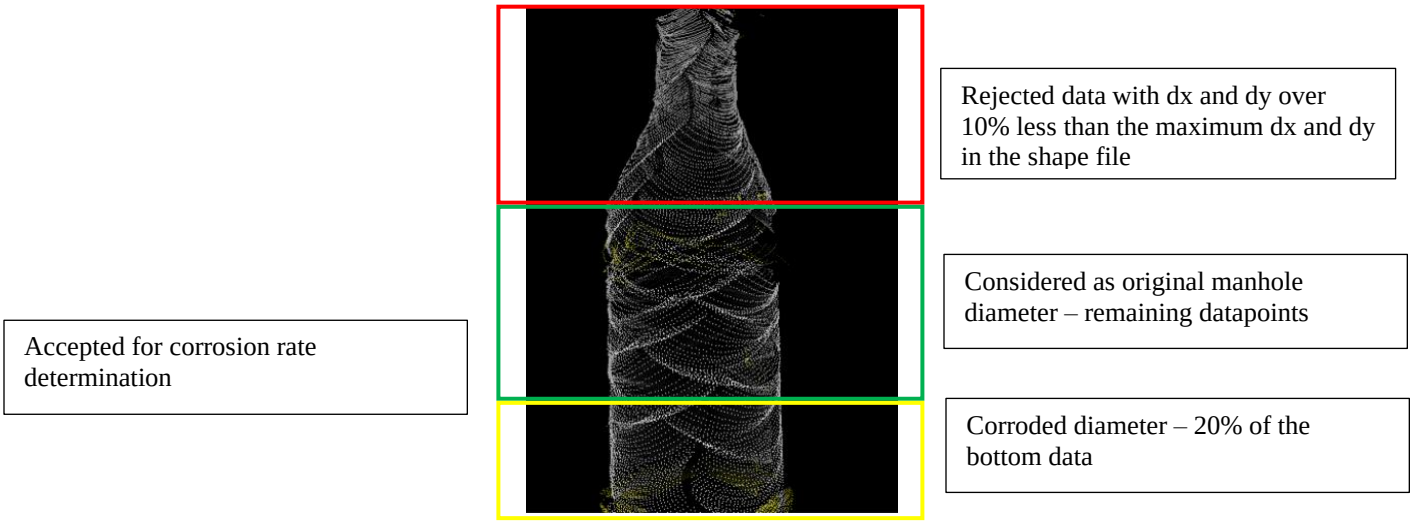


Fig 4.27 Manhole point cloud showing portion used for corrosion rate determination*

Note: The 20% are number of datapoints, which do not correspond with height.

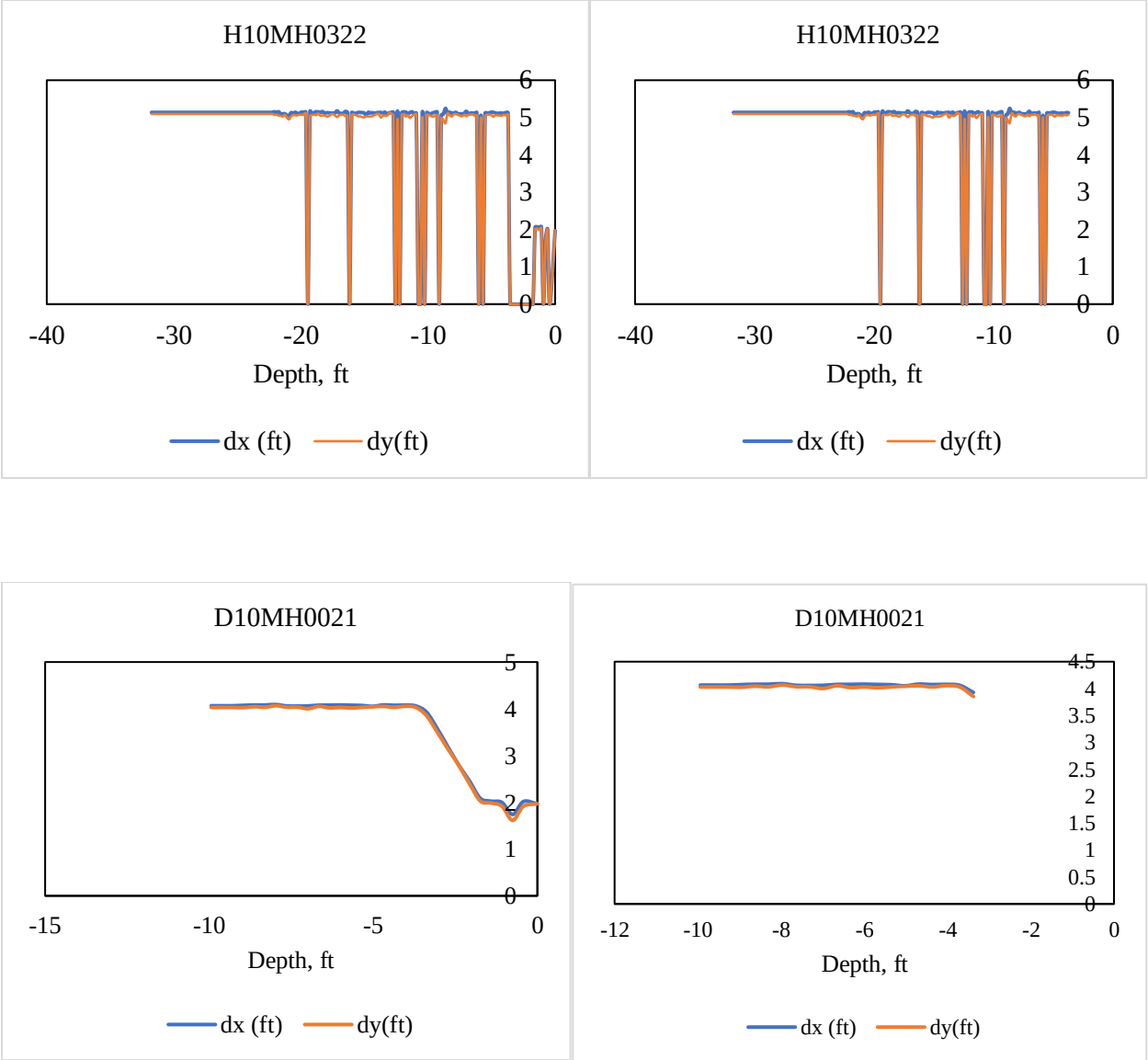


Fig 4.28 Rejecting data 10% less than the maximum dx and dy from top of manhole shape file

In preliminary analysis, we observed increase in temperature with the depth of the manhole, and also higher hydrogen sulfide concentrations were observed 1 to 5 feet above the inlet pipe as seen in Fig. 4.29a, Fig. 4.29c, Fig. 4.29e. Thus, higher hydrogen sulfide concentrations along with increased temperature will allow for increased diffusion near the inlet thus higher corrosion rate might be anticipated at lower sections of the manhole. In Fig 4.29, we see the change in gas phase temperature with the depth of the manhole for 3 different manhole designs. In Fig 4.29a-b represents a manhole with 3 inlets (3 drops), Fig 4.29c-d represents a manhole with 2 inlets (1 high drop at 7 feet, and no drop), and Fig 4.29d-f represents a manhole with no drop. In the Fig. 4.29b, Fig. 4.29d, Fig. 4.29f, we

see the rise in gas phase temperature with manhole depth expressed in percentages, and a sharp increase in gas phase temperature is observed in the lowest 20% of datapoints closest to the water surface. From literature review, we know that corrosion rates increase with temperature due to increased diffusion of H_2S in air and increased chemical and biological sulfide oxidation rates[43]. Thus, our assumption was that we would see higher corrosion in the lower 20% of the manhole datapoints, where high hydrogen sulfide concentrations in high along with a gas phase temperature to allow for increased diffusion and oxidation reaction to occur.

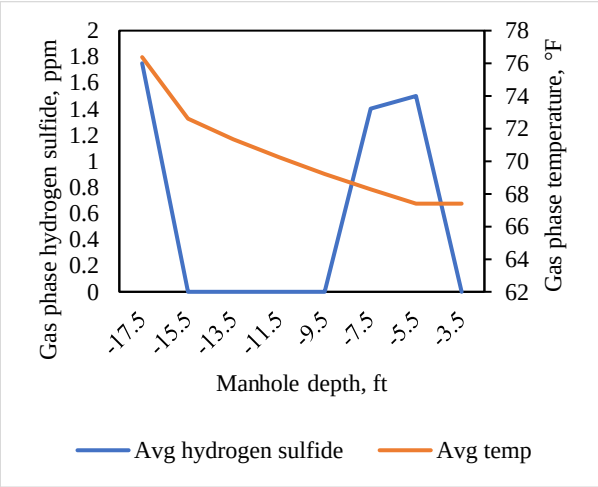


Fig 4.29a F04MH0026: Gas phase temp, H_2S vs depth

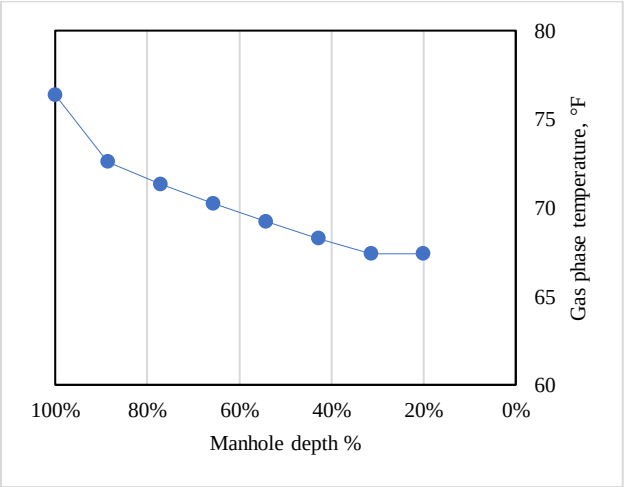


Fig 4.29b F04MH0026: Gas phase temp, H_2S vs depth%

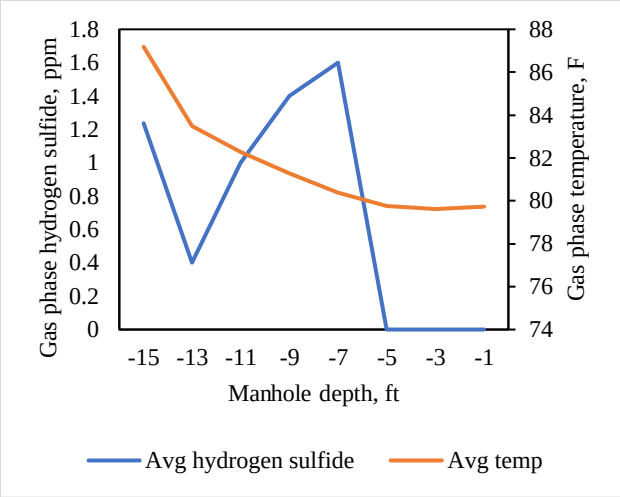


Fig 4.29c G12MH0226: Gas phase temp, H_2S vs depth

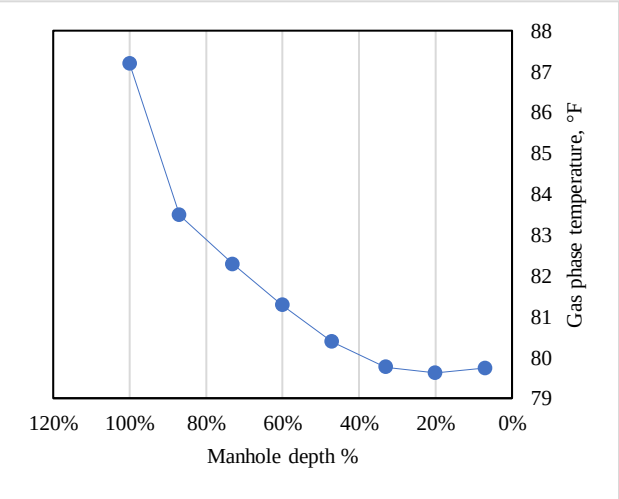


Fig 4.29d G12MH0226: Gas phase temp, H_2S vs depth%

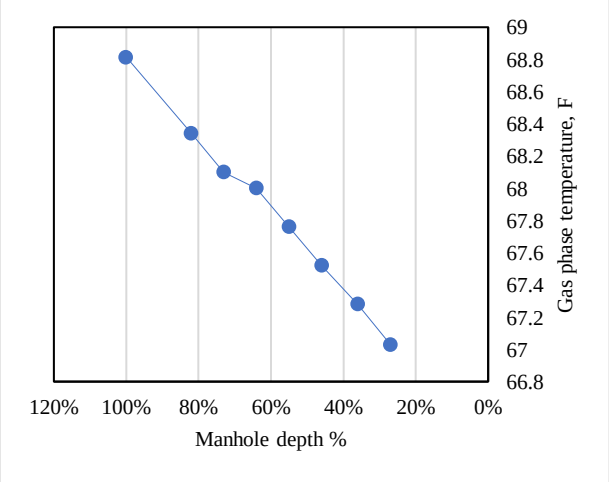


Fig 4.29e H09MH0411: Gas phase temp, H_2S vs depth

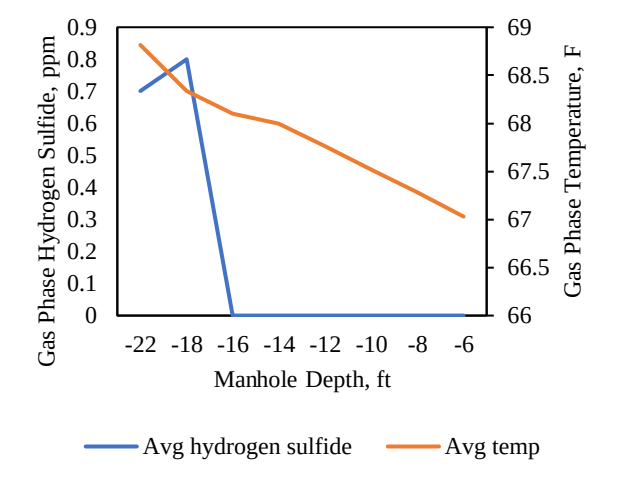


Fig 4.29f H09MH0411: Gas phase temp, H_2S vs depth%

4.6.2 CORROSION RATE FACTORS

Based on the manholes studied, 86.9% of the 130 manholes recorded corrosion rates between 0 to 2.6 mm/yr along the dx and 82.3% of the 130 manholes recorded corrosion rates between 0 to 3 mm/yr along the dy. The highest recorded corrosion rate was 24.4 mm/yr, and 24.0 mm/yr along the dx and dy axis respectively, with an average corrosion rate of 1.7 mm/yr and 2.2 mm/yr along the dx and dy axes respectively, as seen in Fig 4.30. The data

seems is right skewed, thus signifying that the data is not normally distributed thus there is need for transformation to carry out multiple linear regression.

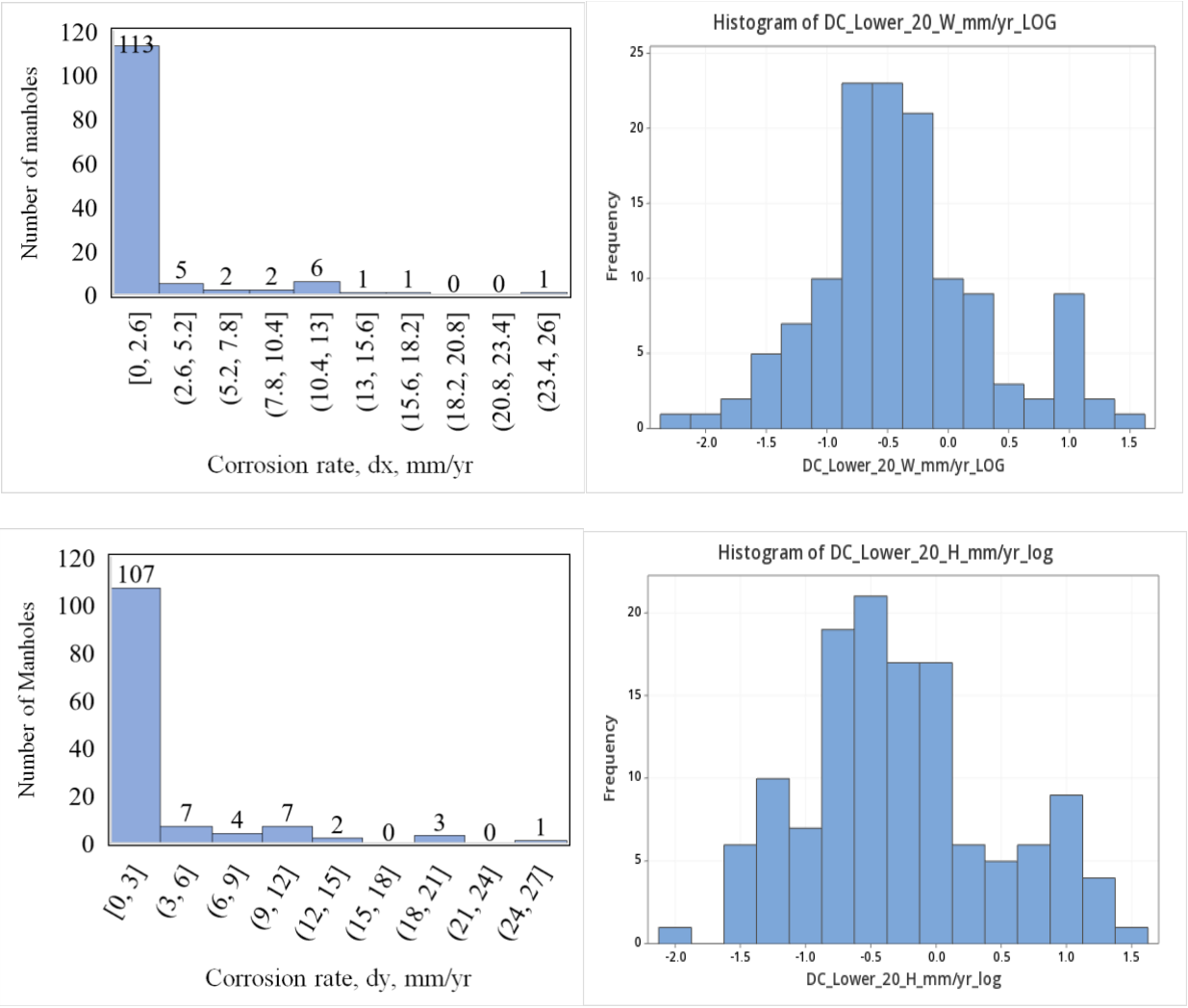


Fig 4.30 Histograms for average corrosion rate along dx and dy in mm/y

4.6.3 EFFECT OF MANHOLE DESIGN ON CORROSION RATE

From the above discussion on hydrogen sulfide generation, we see that there is an effect of design on hydrogen sulfide concentration in the manhole; since corrosion occurs due to the hydrogen sulfide present in the manhole, it becomes important to understand if corrosion rate could also be a function of manhole design. Fig. 4.31a and Fig. 4.31b show the effect of design on average corrosion rates along the dx and dy axis, and Fig 4.31c shows the

average hydrogen sulfide and corrosion rate along dx and dy for individual designs. It seems the effect of design on corrosion rate was different from what was seen with the hydrogen sulfide generation. In hydrogen sulfide generation, we saw that location of the manhole such as downstream of a lift station played an important role, whereas for corrosion rate manhole downstream of the lift station recorded the third lowest average corrosion rates along dx and second lowest along the dy. For manholes upstream from the lift station, we recorded low average hydrogen sulfide concentrations, but the corrosion rates were second highest along the dx and third highest along the dy. Also, the manholes downstream from the high hydrogen sulfide manholes recorded lower corrosion values than manholes upstream from the high hydrogen sulfide values. The highest average corrosion rate was seen for manholes with medium drop (>=2', <2'). Designs which are considered to contribute to lower average gas phase hydrogen sulfide concentrations seems to have showed higher corrosion rates such as uniform pipe size change, subcritical flow, obtuse bend designs, whereas manholes which result in higher hydrogen sulfide generation such as high drop, supercritical flow, right angle bend seem to be showing lower corrosion rates ranging anywhere between 1 mm/yr to 3 mm/yr. This could be because of other factors which might be playing a more prominent role in controlling the corrosion rates.

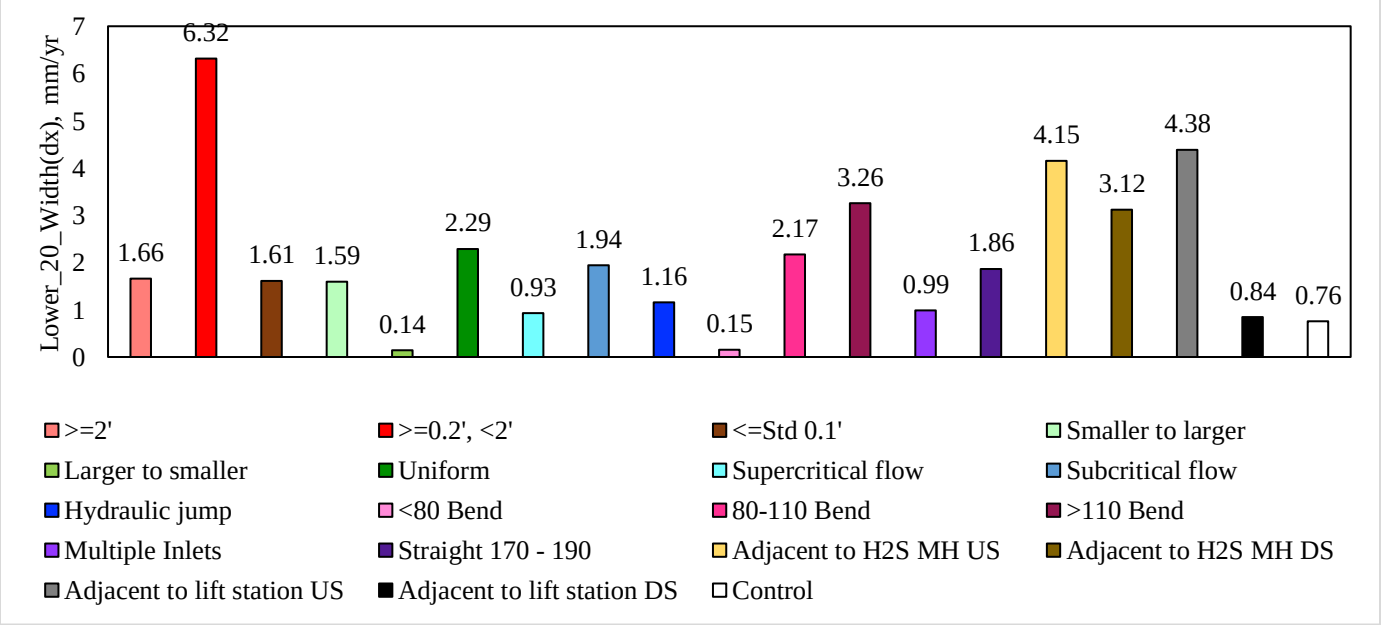


Fig 4.31a Effect of manhole design on average corrosion rate along dx

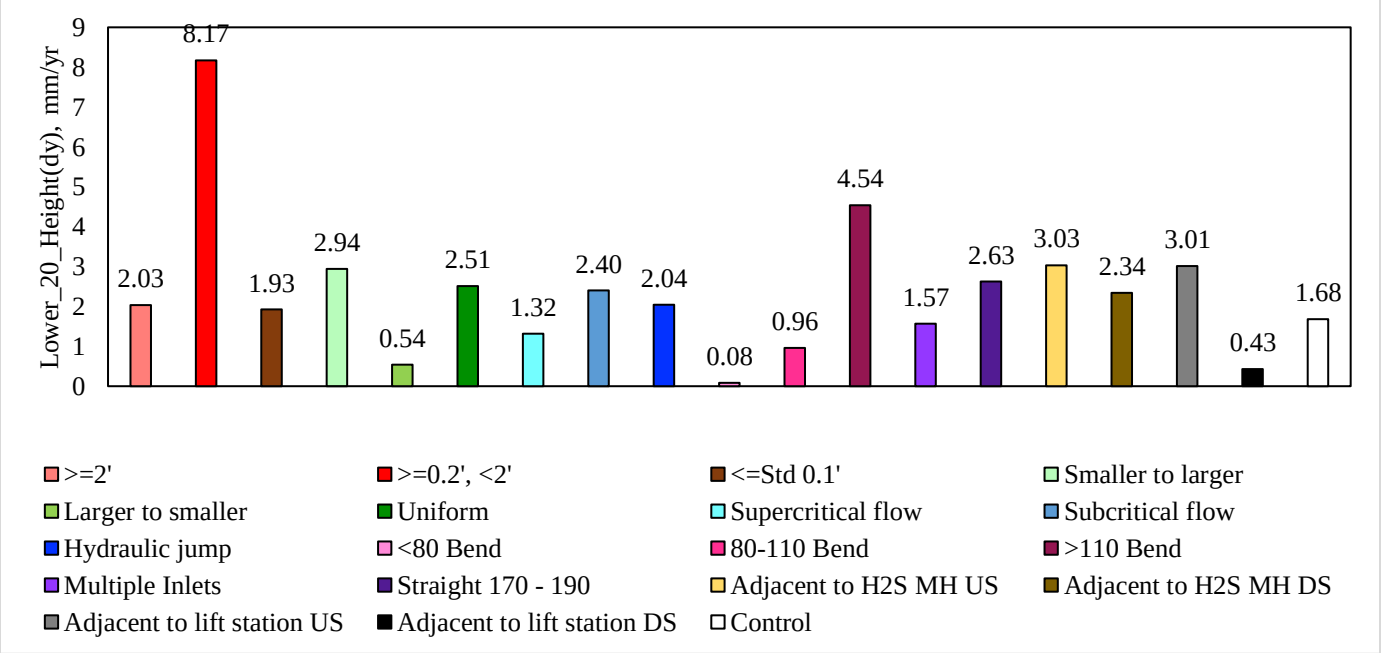


Fig 4.31b Effect of manhole design on average corrosion rate along dy

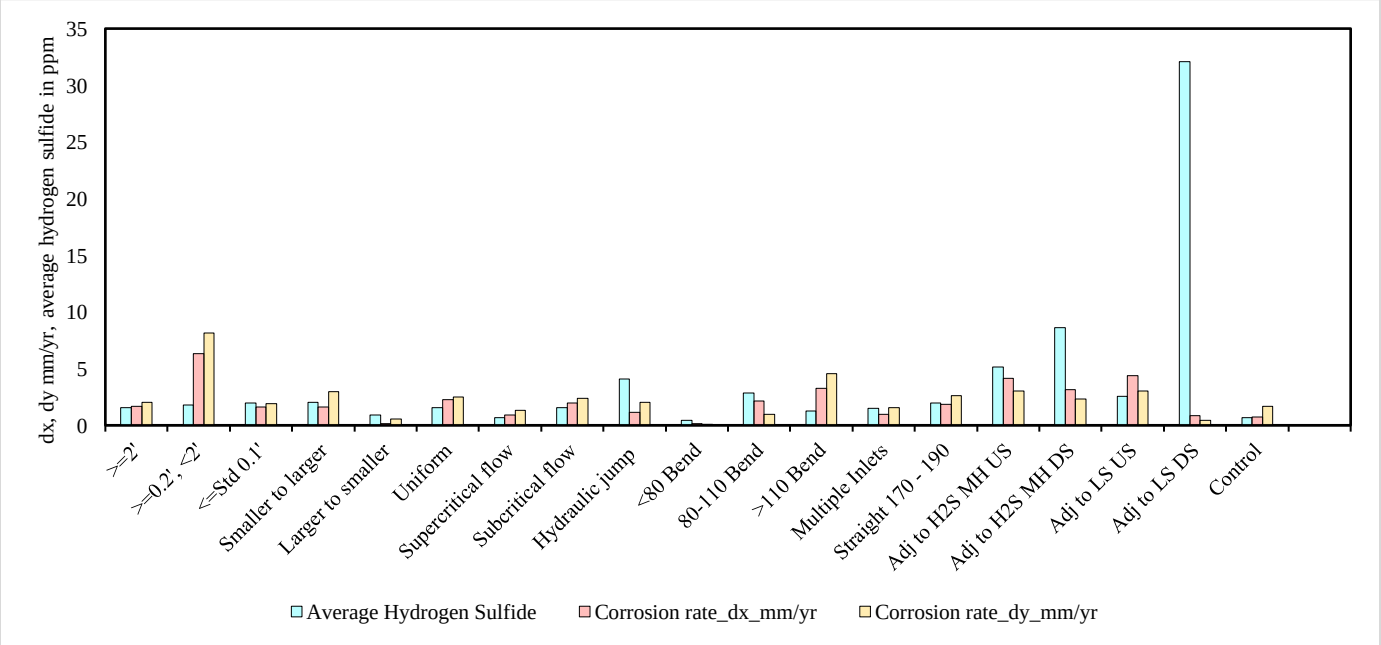


Fig 4.31c Average hydrogen sulfide and corrosion rates for various designs

The corrosion process is controlled by various factors such as microorganisms, type of microbial community, pH of concrete, age of manholes, concrete composition, temperature, and relative humidity, in addition to hydrogen sulfide

concentrations on the manhole's surface, as discussed in the section 2.2.2 of Chapter 2. The start of corrosion process requires neutrophilic sulfur oxidizing bacteria (NSOM) like *Thiobacillus* for sulfuric acid generation, thus lowering the concrete pH. As the pH reduces, another microbial community - acidophilic sulfur oxidizing microorganisms (ASOM) - will start to develop thus accelerating the process. In really corroded manholes, sulfuric acid can even react with concrete compounds silicate and carbonate to form gypsum[37, 39, 41,50]. So, it is clear that as the corrosion increases, the reactions occurring on the concrete surface change, and various factors can dominate during different periods of time. Moreover, corrosion is a slow process in which good microbial community growth is required for constant and continuous corrosion. If for any reason there is a washout or removal of this microbial community, this may impact corrosion rate. We can hypothesize that turbulent designs, such as presence or high drops or hydraulic jump, may result in washing away of the microbial community present on the manhole surface, and this might result in lower corrosion rates; however, in calmer flows such as subcritical flow, or obtuse bends, the microbial community has a chance to thrive, thus increasing corrosion.

In addition, in contrast to the sewer pipelines, the manhole is a vertical structure; thus, the temperature varies along the depth of the manhole. This variation will impact corrosion rates along the depth of the manhole. Furthermore, in very deep manholes, gas-phase oxygen concentrations could be low near the base of the manhole compared to ambient oxygen concentration near the manhole lid (since it is exposed to environment); this might also impact corrosion rates. Moreover, the uptake rate of the hydrogen sulfide by the concrete has been found to change with temperature depending on the age of concrete due to increased diffusion of H_2S in air and increased chemical and biological sulfide oxidation rates[43]: new concrete showed a 17% increase in hydrogen sulfide uptake rate compared to 26% for old corroded concrete with increase in temperature[43]. Thus, old manholes in our dataset might have had different corrosion rates in comparison to new manholes. Also, studies have suggested that the lower pH of older concrete allows for easier hydrogen sulfide uptake because older concrete has lower pH to support growth of ASOM, which will faster oxidize the hydrogen sulfide gas to produce sulfuric acid [43] and this may also impact corrosion rates[42]. Finally, one study has shown that rate of corrosion will be different for sections of the concrete submerged in wastewater and section of concrete exposed to air[47]. We measured corrosion in sections exposed to air only, not submerged. Thus, we can suggest that corrosion rates might be varying along the depth of the manhole with few areas more impacted than the others. Another reasons for corrosion rate been higher in manholes which

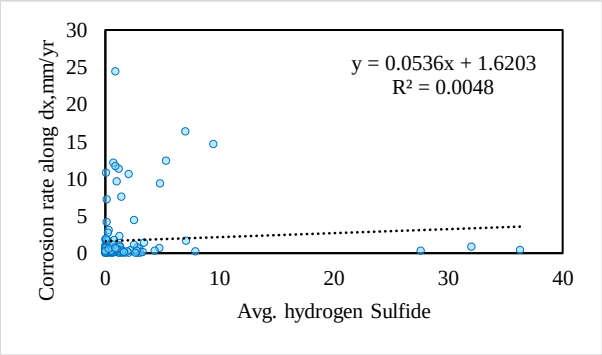
recorded lower hydrogen sulfide concentrations could be the faster adsorption of hydrogen sulfide onto the manhole's walls, resulting in lower concentrations in the gas phase.

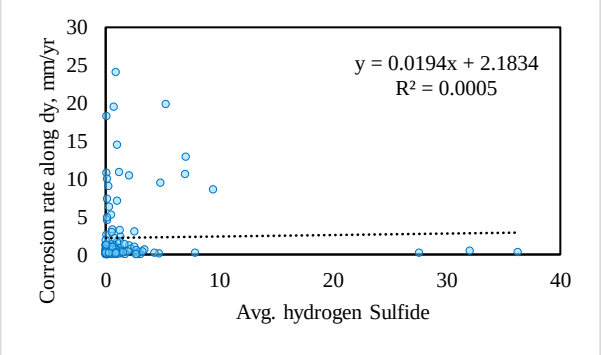
The above discussion clearly outlies that there might be two reasons for certain designs such as obtuse, medium drop showing higher corrosion rates; they are:

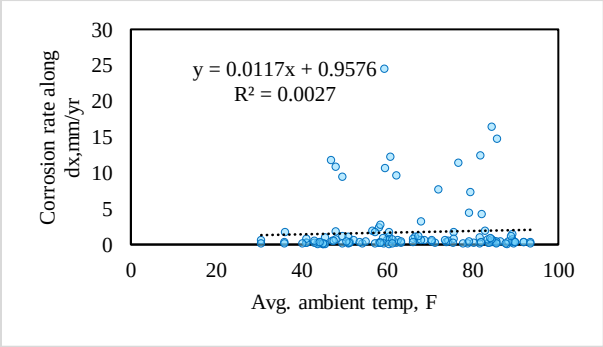
a. Manholes with less turbulent designs (such as high drop, hydraulic jump, acute bend) provide a more stable environment to allow for microbial growth thus allowing continuous and steady corrosion.
b. Manholes' surfaces which are exposed to wastewater due to drop or turbulent flow might not have been entirely considered due to our assumption that lower sections are highly corroded.

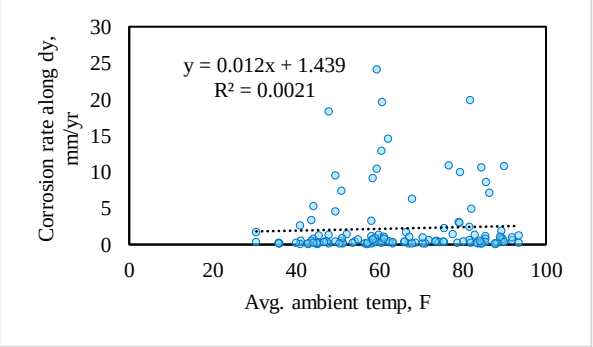
4.6.4 EFFECT OF GAS AND LIQUID PARAMETERS ON CORROSION RATE

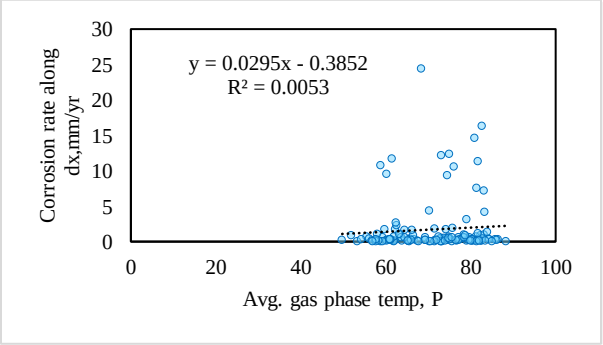
Fig 4.32 shows that none of the gas or liquid factors show any strong correlation with that of corrosion rates. The highest observed correlation is only of 2% between the BOD and corrosion rate along the dx, sulfate along the dy, and dissolved oxygen along the dx. The reason for no strong correlation could be because many factors control rate of corrosion as clear from literature review. The process of corrosion takes place in 4 stages[37, 39, 41,50] with the involvement of different microbes in each stage: the sulfate in wastewater is converted to sulfide, which then volatilizes into the gas phase as hydrogen sulfide, which then adsorbs onto the concrete surface where the microbes convert it to sulfuric acid to cause corrosion. In addition, pH of the concrete declines with advancement into next stage and the decline in pH hastens the corrosion process. Also, microbial activity depends on presence of necessary electron donors and temperature; thus, a direct relation between these parameters might may not be visible, but there might a combined effect of these factors, which the MLR model will help to identify, as MLR helps to identify non-linear relationships (such as log or inverse). Another factor that might be affecting correlation is the assumption that bottom 20% is the most corroded, and considering remaining 75% as the original manhole diameter, as that might hold truth for all the manholes (upper sections of the manhole might be more corroded, or might have gone through structural damage during their lifetime due to loading).

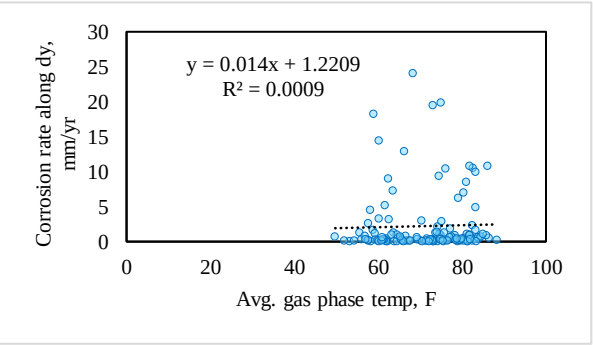

y = 0.0536x + 1.6203
R² = 0.0048
Corrosion rate along dx,mm/yr
Avg. hydrogen Sulfide

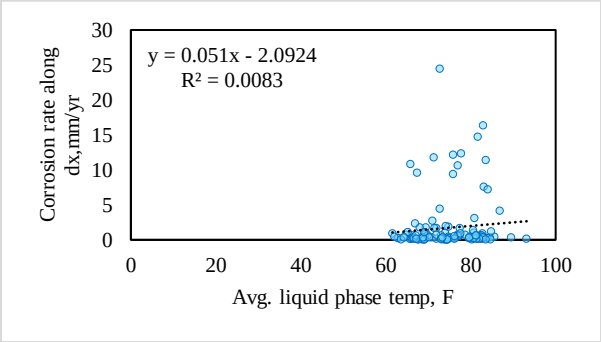

y = 0.0194x + 2.1834
R² = 0.0005
Corrosion rate along dy, mm/yr
Avg. hydrogen Sulfide

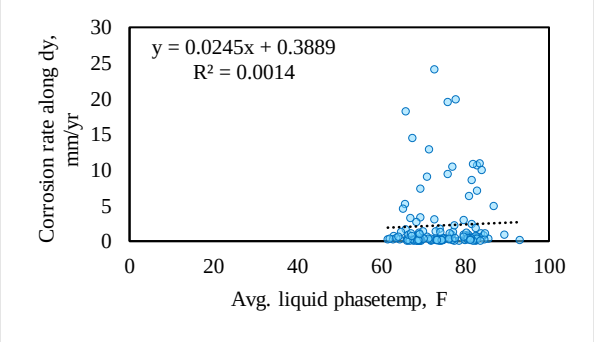

y = 0.0117x + 0.9576
R² = 0.0027
Corrosion rate along dx,mm/yr
Avg. ambient temp, F

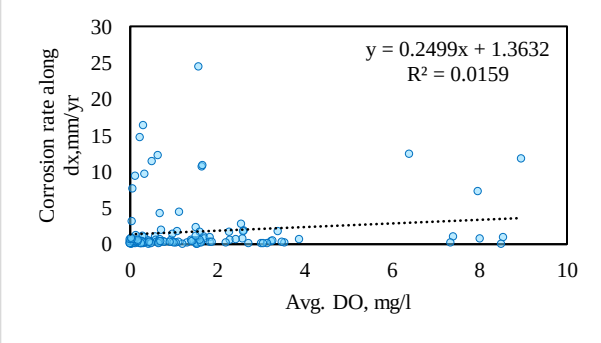

y = 0.012x + 1.439
R² = 0.0021
Corrosion rate along dy, mm/yr
Avg. ambient temp, F

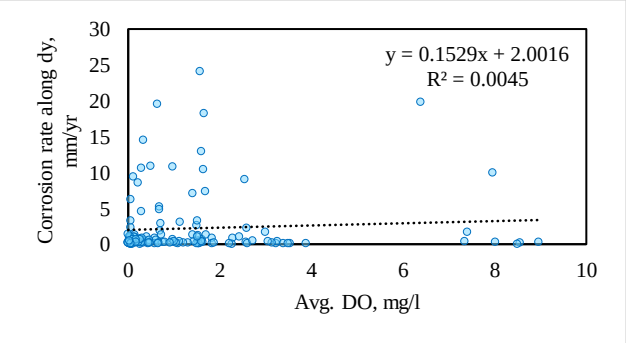

y = 0.0295x - 0.3852
R² = 0.0053
Corrosion rate along dx,mm/yr
Avg. gas phase temp, P

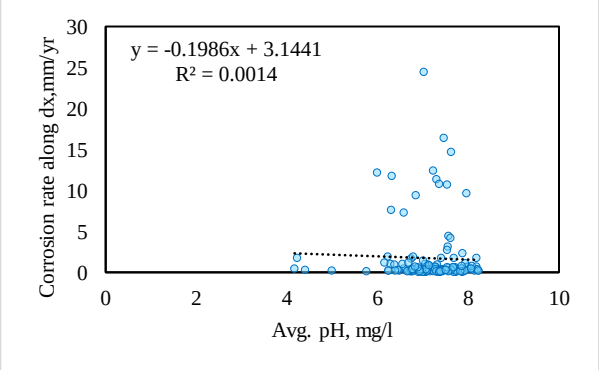

y = 0.014x + 1.2209
R² = 0.0009
Corrosion rate along dy, mm/yr
Avg. gas phase temp, F

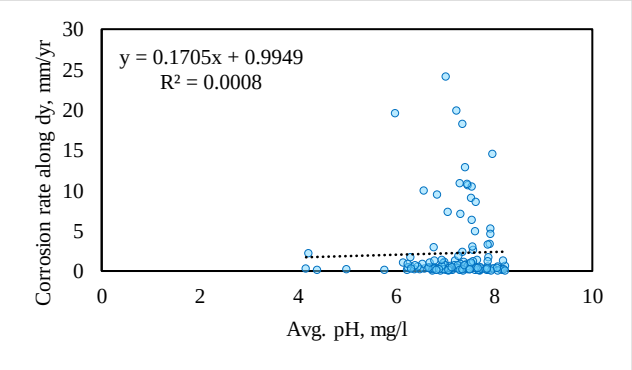

y = 0.051x - 2.0924
R² = 0.0083
Corrosion rate along dx,mm/yr
Avg. liquid phase temp, F


y = 0.0245x + 0.3889
R² = 0.0014
Corrosion rate along dy, mm/yr
Avg. liquid phasetemp, F


y = 0.2499x + 1.3632
R² = 0.0159
Corrosion rate along dx,mm/yr
Avg. DO, mg/l


y = 0.1529x + 2.0016
R² = 0.0045
Corrosion rate along dy, mm/yr
Avg. DO, mg/l


y = -0.1986x + 3.1441
R² = 0.0014
Corrosion rate along dx,mm/yr
Avg. pH, mg/l


y = 0.1705x + 0.9949
R² = 0.0008
Corrosion rate along dy, mm/yr
Avg. pH, mg/l

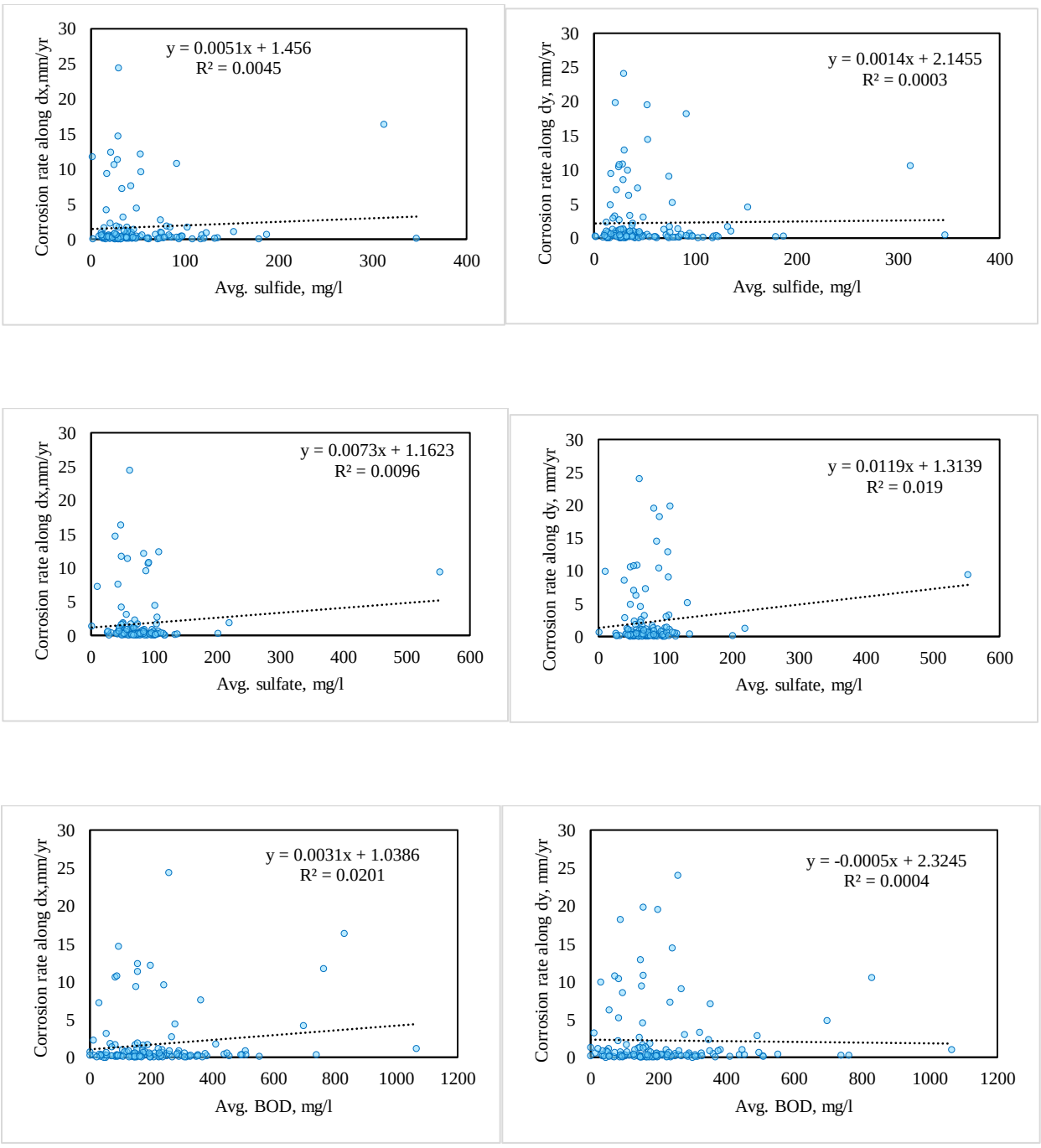


Fig 4.32 Corrosion rate vs gas and liquid phase parameters along the dx and dy axis

In conclusion, we can infer that:

a. Designs and corrosion rate seem to be related for the manholes adjacent to high H_2S manholes, which displayed continuous high hydrogen sulfide concentrations.
b. Gas and liquid parameters do not show strong association with the corrosion rate individually; however stronger correlation may be seen if all the factors are considered together.
c. Since corrosion is a complex process, factors not measured, such as microbial community, may play a vital role in determining corrosion rate.
d. Hydrogen sulfide concentrations show a strong correlation with temperature; however, the same is not seen with corrosion.

CHAPTER 5

MULTIPLE LINEAR REGRESSION MODELS FOR PREDICTING HYDROGEN SULFIDE CONCENTRATIONS AND CORROSION

From Chapter 4 results, we can infer that multiple factors contribute to hydrogen sulfide generation/release and corrosion rates. This chapter will discuss the development of prediction equations for hydrogen sulfide and corrosion using the multiple linear regression.

5.1 MODEL DEVELOPMENT PROCESS

The MLR model building process was conducted in the following stages:

1. Development of a preliminary model and check for model assumptions.
2. Transformation of variables to fix concerns of curvature, non-constant variance, and normality.
3. Preliminary model selection to determine significant variables.
4. Exploration of interaction terms and incorporation into the model.
5. Model selection using significant variables and interaction terms.
6. Validation of selected model.

For model building, we will be using the significance level of $\alpha = 0.1$. The model must satisfy the following assumptions to be considered for multiple linear regression:

1. Residuals are normally distributed.
2. Residuals have a constant variance.
3. Residuals satisfy the MLR form.
4. Residuals are uncorrelated.
5. No outliers are present.
6. Predictors are not highly correlated.

Various models were generated using the Stepwise method and Best Subset Selection method and compared based on the above assumptions. The model which performed the best in most of the categories was selected as the final model. The final model was then validated with the test set, and the MSPR, and Rsq(pred) values were found to confirm model prediction capabilities.

Because of the limited number of manholes present near lift stations in the city, data was insufficient to include this category in the MLR model. Also, most of the data collection belonged to non-rainy or less rainy days; thus, manholes sampled during precipitation days greater than 3 inches are not considered, as not enough data points were available for a prediction.

5.2 HYDROGEN SULFIDE GENERATION AND VOLATILIZATION MODEL

A multiple linear regression model is of the form as shown in Eqn. 5.1

$$Y = \beta_0 + \beta_1 x_1 \ldots . + \beta_{29} x_{29} + \varepsilon \ldots\ldots \text{Eqn. 5.1}$$

Where:

Y = the gas phase hydrogen sulfide concentration in the manhole in ppm (response variable),

β_o = intercept,

$\beta_1, \beta_2, \ldots . \beta_{29}$ = slopes for the respective predictor variables,

$x_1, x_2, \ldots . x_{29}$ are the predictor variables (see . Table 5.1),

ε = random error term. The slope demonstrates the relationship between the predictor variable and the rate of hydrogen sulfide (response variable).

A manhole can have either supercritical flow, subcritical flow, or a hydraulic jump but never all three; thus, to avoid multicollinearity, subcritical flow is not considered in model building. Similarly, a manhole can either have multiple inlets or bends (acute, right, obtuse, straight) but never both; thus, to avoid multicollinearity, acute angle bends are not considered in model building process.

Table 5.1 Predictor variables in the hydrogen sulfide model

Variable			Notation	Type of data
Flow Rate, MGD			x_1	Continuous
Velocity, ft/s			x_2	Continuous
Depth of Flow, ft			x_3	Continuous
Manhole Depth, ft			x_4	Continuous
Average Ambient Temperature, °F			x_5	Continuous
Precipitation inches			x_6	Continuous
Design Variables	Drop	>=2' – High drop	x_7	Counters
		>=0.2', <2' – Medium drop	x_8	Counters
		<=Std 0.1' – No drop	x_9	Counters
		Maximum drop height, ft	x_{10}	Continuous
	Pipe Size Change	Smaller to larger	x_{11}	Counters
		Larger to smaller	x_{12}	Counters
		Uniform	x_{13}	Counters
		Maximum Pipe size change, inches	x_{14}	Continuous
	Flow Type	Supercritical flow	x_{15}	Binary
		Hydraulic jump	x_{16}	Binary
	Bends	Multiple inlets	x_{17}	Binary
		80-110° Bend – Right bend	x_{18}	Binary
		>110° Bend – Obtuse bend	x_{19}	Binary
		No Bend 170° - 190° - Straight	x_{20}	Binary
	Adjacent to High H_2S manhole	Upstream to high H_2S manhole	x_{21}	Binary
		Downstream to high H_2S manhole	x_{22}	Binary
	Average Liquid Temp., °F		x_{23}	Continuous

Liquid-Phase Variable s	Average Liquid DO, mg/L	x_{24}	Continuous
	Average pH	x_{25}	Continuous
	Average Sulfide, mg/L	x_{26}	Continuous
	Average Sulfate, mg/L	x_{27}	Continuous
	Average BOD, mg/L	x_{28}	Continuous
Gas-Phase Variables	Average Temp., °F	x_{29}	Continuous

For the hydrogen sulfide generation and volatilization model, we divided the dataset into 2 sets: one for the training set, which consisted of 113 manholes, and test set, which consisted of 28 manholes. Test set dataset selection was done randomly with the use of software to avoid bias in the model. The model was built using the training set and validated for robustness with the test set.

5.2.1 PRELIMINARY ANALYSIS

For the preliminary analysis, models were developed using all the 29 predictor variables. **Data used for model development and validation available on request from the suthor.**

5.2.1.1 CORRELATION MATRIX

Before building the model, the correlation between the predictors and the response variable was evaluated, to identify any serious multicollinearity concerns between the predictors. Any correlation higher than 70% (0.7) between the predictors shows concerns of linear dependencies and may affect model accuracy. Higher collinearity between the response and predictor variable is better as it improves the accuracy of the model.

Correlation between response and predictor variables: The highest observed correlation was 44.7% between the gas phase hydrogen sulfide and the predictor variable manhole downstream of a high hydrogen sulfide manhole, demonstrating our previous assumption that manhole location does effect amount of hydrogen sulfide generation/release. The second highest correlation of 32.6% was seen between the liquid phase pH and the gas phase

hydrogen sulfide, which was expected, as pH regulates the solubility of the hydrogen sulfide in the wastewater. All the other predictor recorded less than 20% correlation with the response variable.

Correlation between the predictors: Flowrate and depth of flow showed a correlation of 93%, gas phase temperature with ambient temperature showed a correlation of 77%, gas phase temperature with liquid phase temperature showed a correlation of 92%, and high drop with maximum drop height showed a correlation of 78%. These high correlation between the temperature predictors seems reasonable as it was seen in Chapter 4 that ambient temperature affects the gas and liquid phase temperatures. All the other predictors recorded correlations less than 70% between each other. Since there is definitely multicollinearity between these variables, we need to check for variance inflation factor (VIF) in the final model.

5.2.1.2 FITTING THE PRELIMINARY MODEL

Table 5.2 shows the results of the preliminary model and Table 5.3 shows the Analysis of Variance table. Table 5.1 confirms that 27 predictors have P-value greater than 0.1, and 16 predictors have a VIF value of greater than 5, thus signifying multicollinearity concerns.

Table 5.2 Coefficient estimates for the predictor variables – preliminary hydrogen sulfide model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	13.7	8.04	1.7	0.092	
Flow_Rate	-0.047	0.363	-0.13	0.897	10.16
Velocity	0.812	0.48	1.69	0.094	3.08
Depth_of_flow	0.24	2.67	0.09	0.927	11.87
MH_Depth	-0.0725	0.0663	-1.09	0.277	1.97
W_Temp_A	0.0216	0.0364	0.59	0.554	3.94
Precipitation	-0.73	1.16	-0.63	0.532	1.46
G_Temp_A	-0.062	0.118	-0.52	0.604	13.11
L_Temp_A	0.058	0.151	0.38	0.702	12.33

L_DO_A	-0.02	0.172	-0.11	0.909	1.47
L_PH_A	-1.905	0.54	-3.53	0.001	1.69
L_Sulfide_A	0.0078	0.00772	1.01	0.315	1.39
L_Sulfate_A	0.0036	0.0135	0.27	0.79	1.62
L_BOD_A	0.00146	0.00218	0.67	0.503	1.29
High_Drop	-3.19	2.41	-1.33	0.189	8.69
Medium_Drop	-1.19	2.01	-0.59	0.555	11.71
No_Drop	-1.27	1.96	-0.65	0.519	13.87
Drop_max	0.108	0.166	0.65	0.514	5.15
Small_toLarge_PS	0.86	2.19	0.39	0.697	35.63
Large_to_Small_PS	0.08	2.75	0.03	0.978	5.96
Uniform	0.95	2.23	0.43	0.671	18.84
PS_Max_Neg	-0.005	0.0651	-0.08	0.939	3.12
Supercritical	0.01	1.22	0.01	0.994	1.83
Hydraulic_Jump	0.89	1.21	0.74	0.464	1.59
Right	-2.64	2.65	-1	0.322	9.65
Obtuse	-1.6	2.6	-0.62	0.54	11.01
Multiple_inlets	-1.67	2.87	-0.58	0.562	21.86
Straight	-1.01	2.63	-0.38	0.701	11.69
Adj_High_H2S_US	0.28	1.63	0.17	0.866	1.22
Adj_High_H2S_DS	6.19	1.43	4.32	0	1.45

Table 5.3 ANOVA for preliminary hydrogen sulfide model

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Regression	29	607.69	41.06%	607.695	20.955	1.99	0.008
Flow_Rate	1	1.96	0.13%	0.176	0.176	0.02	0.897
Velocity	1	14.78	1.00%	30.113	30.113	2.87	0.094
Depth_of_flow	1	0.87	0.06%	0.088	0.088	0.01	0.927
MH_Depth	1	9.28	0.63%	12.57	12.57	1.2	0.277
W_Temp_A	1	44.65	3.02%	3.715	3.715	0.35	0.554
Precipitation	1	3.08	0.21%	4.137	4.137	0.39	0.532
G_Temp_A	1	4.39	0.30%	2.853	2.853	0.27	0.604
L_Temp_A	1	0.88	0.06%	1.554	1.554	0.15	0.702
L_DO_A	1	0.4	0.03%	0.137	0.137	0.01	0.909
L_PH_A	1	209.84	14.18%	130.7	130.7	12.44	0.001
L_Sulfide_A	1	7.65	0.52%	10.732	10.732	1.02	0.315
L_Sulfate_A	1	0.64	0.04%	0.748	0.748	0.07	0.79
L_BOD_A	1	5.57	0.38%	4.76	4.76	0.45	0.503
High_Drop	1	13.77	0.93%	18.461	18.461	1.76	0.189
Medium_Drop	1	16.06	1.09%	3.701	3.701	0.35	0.555
No_Drop	1	10.31	0.70%	4.404	4.404	0.42	0.519
Drop_max	1	19.77	1.34%	4.507	4.507	0.43	0.514
Small_toLarge_PS	1	2.43	0.16%	1.605	1.605	0.15	0.697
Large_to_Small_PS	1	10.79	0.73%	0.008	0.008	0	0.978
Uniform	1	2.01	0.14%	1.908	1.908	0.18	0.671
PS_Max_Neg	1	0.73	0.05%	0.061	0.061	0.01	0.939
Supercritical	1	7.16	0.48%	0.001	0.001	0	0.994
Hydraulic_Jump	1	8.54	0.58%	5.686	5.686	0.54	0.464
Right	1	0.42	0.03%	10.433	10.433	0.99	0.322

Obtuse	1	0.31	0.02%	3.976	3.976	0.38	0.54
Multiple_inlets	1	5.42	0.37%	3.57	3.57	0.34	0.562
Straight	1	0.97	0.07%	1.557	1.557	0.15	0.701
Adj_High_H2S_US	1	8.58	0.58%	0.299	0.299	0.03	0.866
Adj_High_H2S_DS	1	196.45	13.27%	196.45	196.45	18.69	0
Error	83	872.33	58.94%	872.325	10.51		
Total	112	1480.02	100.00%				

5.2.1.3 PRELIMINARY MODEL – ASSUMPTIONS CHECK

Before the preliminary model can be considered for further analysis, the model must satisfy the model assumptions. The first assumption that was checked was if the model form was reasonable. As clear from Fig 5.1 that the model is not normal and has a long tail on the top. Also, the spread of the fitted versus the residual plot was not clear; thus, there is a need for transformation on the response variable. No curvature was seen on predictor plots. Since the model failed the first 2 assumptions, transformation was carried out to fix the concerns. Even the residual plots of the predictor variable did not show a good spread.

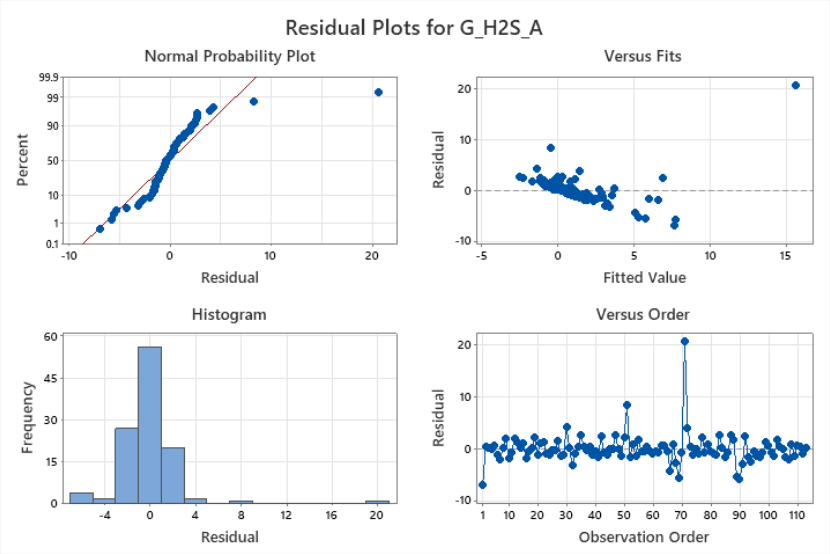


Fig 5.1 Residuals plots for preliminary hydrogen sulfide model

5.2.2 TRANSFORMED MODEL

Before proceeding further with the analysis, we need to fix the concerns for normality and the spread on the fitted vs residual plots. A compression of the response variable will result in a more linear relationship. Any transformation on Y will result affect the distribution of Y and distribution of the residuals. In our case the normality is not satisfied, so we decided to move forward with a transformation on Y.

Initially, weaker transformations were tried to see if they improved the model. The first attempt was to do a square root on the Y, and the model assumptions were again verified; however, the transformation did not fix the spread on the fitted versus the residuals, though the normality was satisfied, as seen from Fig 5.2.

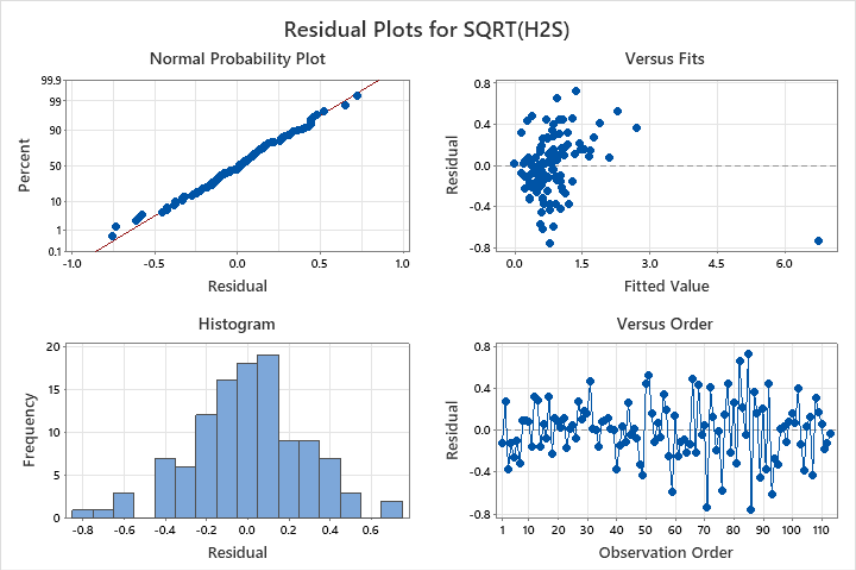


Fig 5.2 Residuals plots for SQRT(Y) transformation of hydrogen sulfide model

Since the SQRT(Y) transformation did not help with the spread, another transformation with logarithm (base10) was conducted. Since, few manholes had zero hydrogen sulfide concentrations, a constant 0.01 was added to the Y, and then a logarithm (base10) was calculated. The transformation improved the spread of the fitted versus the residuals, and the normality as well, as shown in Fig 5.3, and was considered for further analysis.

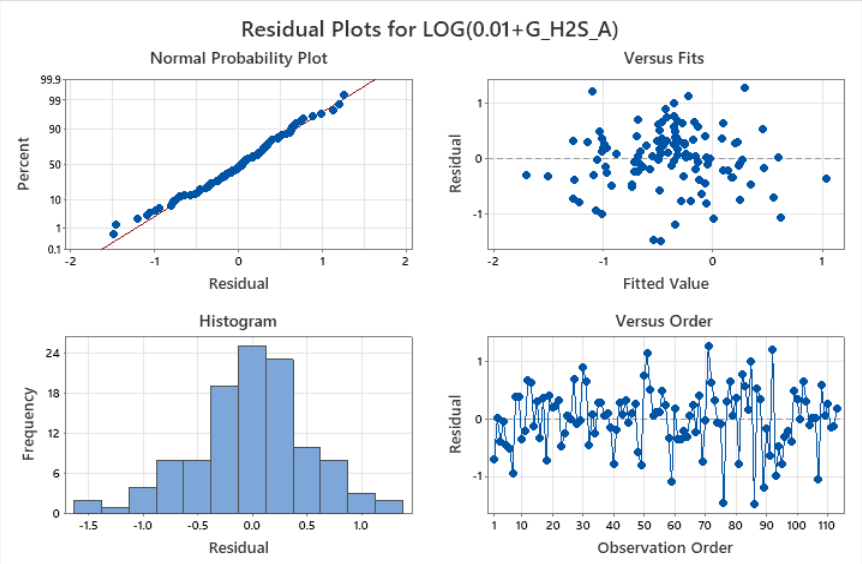


Fig 5.3 Residuals plots for $LOG_{10}(0.01+Y)$ transformation of hydrogen sulfide model

The coefficients of predictor variables and ANOVA table for the transformed model are provided in Table 5.4 and Table 5.5, respectively. Here also, we can see that 15 predictors were insignificant and 15 predictors had a high VIF>5.

Table 5.4 Coefficient estimates for the predictor variables – transformed hydrogen sulfide model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	3.43	1.53	2.25	0.027	
Flow_Rate	-0.0193	0.069	-0.28	0.781	10.16
Velocity	0.3614	0.0912	3.96	0	3.08
Depth_of_flow	-0.029	0.508	-0.06	0.954	11.87
MH_Depth	-0.0166	0.0126	-1.32	0.192	1.97
W_Temp_A	0.00203	0.00692	0.29	0.77	3.94
Precipitation	-0.536	0.221	-2.43	0.017	1.46
G_Temp_A	0.0484	0.0225	2.16	0.034	13.11
L_Temp_A	-0.0646	0.0287	-2.25	0.027	12.33

L_DO_A	-0.0152	0.0328	-0.46	0.643	1.47
L_PH_A	-0.38	0.103	-3.69	0	1.69
L_Sulfide_A	0.00066	0.00147	0.45	0.653	1.39
L_Sulfate_A	0.0016	0.00257	0.62	0.536	1.62
L_BOD_A	0.000575	0.000414	1.39	0.168	1.29
High_Drop	-1.288	0.458	-2.81	0.006	8.69
Medium_Drop	-0.685	0.382	-1.79	0.076	11.71
No_Drop	-0.667	0.373	-1.79	0.078	13.87
Drop_max	0.0658	0.0315	2.09	0.04	5.15
Small_toLarge_PS	0.445	0.416	1.07	0.288	35.63
Large_to_Small_PS	0.665	0.523	1.27	0.207	5.96
Uniform	0.414	0.424	0.98	0.332	18.84
PS_Max_Neg	0.0113	0.0124	0.91	0.364	3.12
Supercritical	-0.458	0.231	-1.98	0.051	1.83
Hydraulic_Jump	-0.558	0.229	-2.43	0.017	1.59
Right	-0.033	0.504	-0.06	0.949	9.65
Obtuse	0.055	0.495	0.11	0.912	11.01
Multiple_inlets	0.008	0.545	0.01	0.989	21.86
Straight	-0.4	0.501	-0.8	0.426	11.69
Adj_High_H2S_US	0.485	0.311	1.56	0.122	1.22
Adj_High_H2S_DS	0.126	0.272	0.46	0.644	1.45

Table 5.5 ANOVA for transformed hydrogen sulfide model

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Regression	29	26.468	45.64%	26.468	0.91269	2.4	0.001
Flow_Rate	1	3.4124	5.88%	0.0297	0.02968	0.08	0.781
Velocity	1	2.5346	4.37%	5.969	5.969	15.71	0
Depth_of_flow	1	0.3271	0.56%	0.0013	0.00128	0	0.954
MH_Depth	1	0.0097	0.02%	0.6582	0.65818	1.73	0.192
W_Temp_A	1	0.4478	0.77%	0.0328	0.03281	0.09	0.77
Precipitation	1	0.9297	1.60%	2.236	2.23596	5.89	0.017
G_Temp_A	1	0.0491	0.08%	1.764	1.76403	4.64	0.034
L_Temp_A	1	1.4186	2.45%	1.9259	1.92585	5.07	0.027
L_DO_A	1	0.0177	0.03%	0.0821	0.08209	0.22	0.643
L_PH_A	1	3.7448	6.46%	5.1859	5.18592	13.65	0
L_Sulfide_A	1	0.0051	0.01%	0.0772	0.07715	0.2	0.653
L_Sulfate_A	1	0.4045	0.70%	0.1468	0.14678	0.39	0.536
L_BOD_A	1	1.2956	2.23%	0.7337	0.73368	1.93	0.168
High_Drop	1	0.7633	1.32%	3.0023	3.00232	7.9	0.006
Medium_Drop	1	0.4217	0.73%	1.223	1.22304	3.22	0.076
No_Drop	1	1.7656	3.04%	1.2118	1.21178	3.19	0.078
Drop_max	1	1.3584	2.34%	1.6588	1.65877	4.37	0.04
Small_toLarge_PS	1	0.0208	0.04%	0.4346	0.43457	1.14	0.288
Large_to_Small_PS	1	0.0004	0.00%	0.6139	0.61387	1.62	0.207
Uniform	1	0.2044	0.35%	0.362	0.36204	0.95	0.332
PS_Max_Neg	1	0.0937	0.16%	0.317	0.31703	0.83	0.364
Supercritical	1	0.9144	1.58%	1.4863	1.48634	3.91	0.051
Hydraulic_Jump	1	3.1167	5.37%	2.2485	2.24848	5.92	0.017
Right	1	0.1893	0.33%	0.0016	0.00159	0	0.949

Obtuse	1	1.19	2.05%	0.0047	0.00465	0.01	0.912
Multiple_inlets	1	0.5321	0.92%	0.0001	0.00008	0	0.989
Straight	1	0.1689	0.29%	0.2427	0.24268	0.64	0.426
Adj_High_H2S_US	1	1.0499	1.81%	0.9263	0.92627	2.44	0.122
Adj_High_H2S_DS	1	0.082	0.14%	0.082	0.08196	0.22	0.644
Error	83	31.5274	54.36%	31.5274	0.37985		
Total	112	57.9954	100.00%				

5.2.2.1 TRANSFORMED MODEL – ASSUMPTIONS CHECK

Before the transformed model can be considered for further analysis, the model must satisfy the model assumptions.

a. **Residuals are normally distributed:** From Fig 5.3, we can see that the model is normal now. Also, normality test was done to validate normality assumption.

   Normality Test: Normality test is conducted to see if the residuals are normally distributed or not.

   Hypothesis: H_0: Normality is OK

   H_1: Normality is violated

   Test: Sample p-value < α, then we reject H_0.

   For the test, we use confidence level α = 0.1. From the software, we get the p-value = 0.214.

   Thus, p-value > α, so we fail to reject H_0. So, the normality is satisfied.

b. **Residuals have a constant variance:** From the fitted and residuals plot, we can infer that the model has constant variance. However, to confirm our assumption, we ran the modified Levene Test, and results for same are below.

   Modified Levene Test: For modified Levene Test, we conduct two-sample t-test on the absolute deviations of the residuals.

   i. We first divide our dataset into two sets of almost equal size. The median of the dataset is selected by using fitted values and it is found to be -0.35316, so we divided the dataset with $\hat{y}$ >-0.35316 and $\hat{y}$ <=-0.35316

ii. Next, we perform the F-Test to check if the variances are equal or not.

Hypothesis: H_0: Variances are equal

H_1: Variances are not equal

Test: If $p < \alpha$, then we reject H_0. For the test, we assume confidence level $\alpha = 0.1$. Results of the test are shown in Table 5.6.1.

Table 5.6.1 F-Test for equality of variances, transformed hydrogen sulfide model

Equality of Variances				
Method	**Num DF**	**Denom DF**	**p-value**	**α**
F Test	56	55	0.9919	0.1

So, our p (0.9919)> α(0.1); thus, we fail to reject H_0 and the variances are equal.

iii. After that we calculate the absolute deviations of the residuals around their group medians.

d1 = |e1 – median e1| and d2 = |e2 – median e2|

iv. Next, we perform the T-Test to see if the model has constant variance or not.

Hypothesis: H_0: Error variances are equal (Means of d1, d2 populations are equal)

H_1: Error Variances are not equal (Means are not equal)

Test: If $p < \alpha$ then we reject H_0. For the test, we assume confidence level $\alpha = 0.1$. Results of the test are shown in Table 5.6.2.

Table 5.6.2 T-Test for error variances, transformed hydrogen sulfide model

Variances	**DF**	**t value**	**p-value**
Equal	111	0.17684	**0.86**
Unequal	110.34	0.17674	0.86

Since the F-Test showed that variances are equal, we select the p-value from Equal T-Test. So, our p (0.86)> α(0.1); thus, we fail to reject H_0 and the error variances are equal. This is in accordance to e vs ŷ plot. Thus, our model satisfies constant variance assumption.

c. **Residuals satisfy the MLR form:** The predictor vs. the residual plots are better for the transformed model in comparison to the preliminary model. However, the residuals on the precipitation and maximum drop height predictors did not have a good spread, as most of the data points are zero, as seen in Fig 5.4. Also, the spread in the flowrate was clustered near zero, as most of the values in the datapoints were less 0.9, as seen in Fig 5.4. There was no curvature visible on any of the plots; thus, the residual satisfies the MLR form.

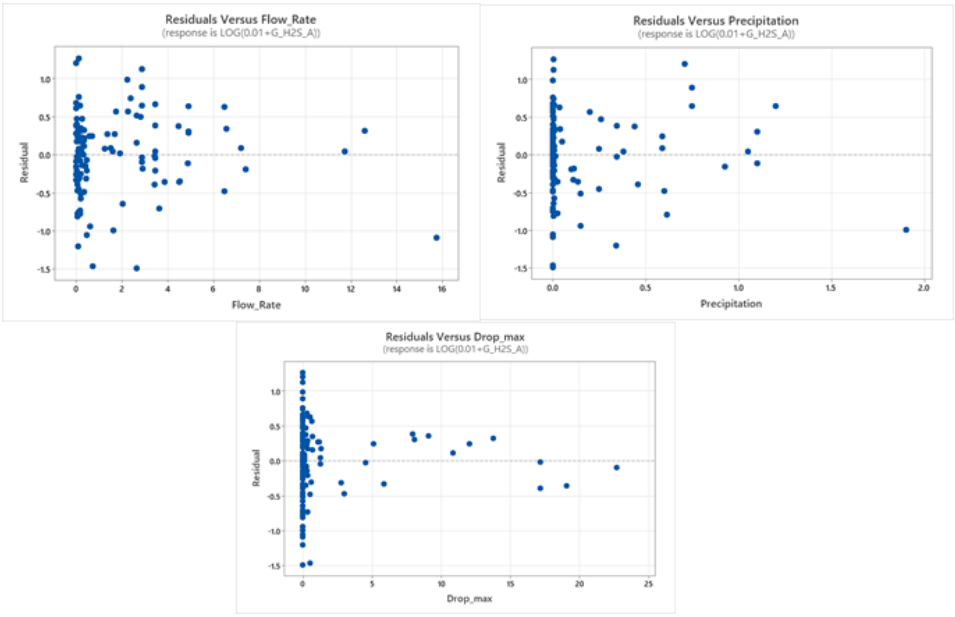


Fig 5.4 Predictor vs. Residuals plots for $LOG_{10}(0.01+Y)$ transformation, hydrogen sulfide model

d. **Residuals are uncorrelated:** From the Fig 5.3, time plot, we can confirm that residuals are uncorrelated for the transformed model.

e. **Check for outliers:** We need to check our data for any kind of outliers, and the influence they have on the ŷ values.

i. **X-Outliers:** To check for X-outliers, leverage values h_{ii} were found; if the h_{ii} is large, then observation *i* is outlying. Usually used guideline is if $h_{ii} > 2p/n$, then the observation *i* is outlying.

Here, $2p/n = 2(30)/113 = 0.5309$

The value h_{ii} is the diagonal element in matrix H. Matrix $H = X(X^TX)^{-1}X^T$. The values of h_{ii} which are greater than the cutoff value 0.5309 are for the observations 14, 46, 70, and 75, thus showing them as x-outliers. Table 5.7 shows the details on X-outliers.

Table 5.7 X- Outliers, transformed hydrogen sulfide model

Obs	h_{ii}	Cook's D	DFITS	Smaller to larger	Larger to smaller	Uniform	Right	Obtuse	Straight
14	0.5406	0	-0.3570	0.212	0.193	0.222	-0.0336	-0.041	-0.028
46	0.6081	0	0.2790	0.009	0.0101	0.0148	-0.232	-0.237	-0.237
70	0.6148	0	-0.0681	-0.016	-0.0136	-0.0156	-0.0014	-0.001	-0.002
75	0.6081	0	-0.2790	0.009	0.010	0.0148	0.205	0.208	0.202

ii. Y-Outliers: To check for Y-outliers, we use the studentized deleted residuals method. We calculate the t_i value by deleting each *i* observation in the dataset, where $t_i = d_i/(\text{sq.rt}(MSE_i\,(1-h_{ii}))$.

The usual guideline is if $|t_i| > t(1- \alpha/2n; n-p-1)$, then a particular observation is a y-outlier.

Assuming $\alpha = 0.1$, then our cutoff value is:

$t(1- \alpha/2n; n-p-1) = t(1-(0.1/2*113); 113-30-1) = 3.4132$

The value of t_i for each observation of $_i$ is provided in the inserted file. We see that none of the t_i values are higher than the cutoff value. So, we can confidently say that there is no y-outlier in our dataset.

iii. Outliers Influence on ŷ: The dataset does contain x-outliers; thus, we need to check the influence the x-outliers have on ŷ, individual LSEs, and combined LSEs. The effect of outliers is verified through 3 methods – Cook's Distance, DFFITS, and DFBETAS.

- **Cook's Distance:** This method verifies the combined impact of observation *i* on all the LSEs. The formula for the Cook's Distance D_i is:

  $D_i = (b\text{-}b_{(i)})^T X^T X(b\text{-}b_{(i)})/p(MSE) = [e_i^2/p(MSE)]*[h_{ii}/((1\text{-}h_{ii})^2)]$

  The values of D_i for various *i* observations are provided in the inserted file. The guideline is if $D_i >$ F(0.50; p,n-p), then the outlier does have influence.

  Here F(0.50; p, n-p) = F(0.50; 30, 113-30) = 0.99

  On comparing the D_i values with the cut off, we see that none of the observations exceed the cutoff values; thus, we can say that none of the outliers have influence on the combined LSES.

- **DFFITS:** This method checks the outliers' influence on the fitted values by computing $\hat{y}_{i(i)}$ for every *i* observation being omitted. The formula to calculate DFFITS is:

  $(DFFITS)_i = (\hat{y}_i - \hat{y}_{i(i)})/ sq.rt(MSE_i * h_{ii}) = ti*(sq.rt(h_{ii}/(1- h_{ii})))$

  The values of DFFITS for various *i* observations are provided in the inserted file. The guideline is if |DFFITS| > 2*sq.rt(p/n), then the outlier does have influence on fitted values.

  Here, 2*sq.rt(p/n) = 2*sq.rt(30/113) = 1.0305

  On comparing the DFFITS values with the cutoff, we see that none of the X-outliers exceed the cutoff, so outliers have no influence on the fitted values.

- **DFBETAS:** This method checks the outliers' influence on individual LSEs by computing $b_{k(i)}$ LSE for every *i* observation being omitted. The formula to calculate DFBETAS is:

  $(DFBETAS)_{k(i)} = b_k - b_{k(i)}/ sq.rt(MSE_i * C_{kk})$

  The values of DFBETAS for various *i* observations for each predictor variable, and the intercept, is provided in the inserted file. The guideline is if |DFBETAS| > 2/sq.rt(n), then the outlier does have influence.

  Here, 2/sq.rt(n) = 2/sq.rt(113) = 0.1881

  The DFBETAS values with the cutoff for each of the predictor and the interceptor were compared, to check for influence on outlier on individual LSEs. The DFBETAS values for the X-outliers are shown in the inserted file. The |DFBETAS| value for smaller to larger pipe size change, larger to

smaller pipe size change, and uniform pipe size in observation 14 shows an influence. |DFBETAS| value for right bend, obtuse bend, and straight in observation 46 and observation 75 shows an influence. The DBETAS values are shown in Table 5.7.

f. **Predictors not highly correlated:** Variance inflation factor is used for validating multicollinearity between the predictors. From the VIF value in Table 5.4, we can see that 15 predictors had a high VIF>5, thus effecting the model quality.

5.2.3 SELECTION OF SIGNIFICANT VARIABLES

From the transformed model, various predictors which are not required as they have very high VIF values. Thus, it becomes important to select only significant variables before we move further with our analysis. Stepwise regression and best subset selection method were conducted on the transformed model to gather a list of important predictor variables. To this new list of predictor variables, we will be incorporating the interaction terms.

a. **Stepwise method:** The stepwise method reduced the variables in the transformed model to 7, as presented in Table 5.8. All the predictors in the stepwise model are significant and the VIF values are also less than 5 so they will be considered in the further analysis.

Table 5.8 Output of stepwise model, transformed hydrogen sulfide model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.305	0.601	2.17	0.032	
Velocity	0.2646	0.0564	4.69	0	1.19
Precipitation	-0.34	0.184	-1.85	0.068	1.02
L_PH_A	-0.2787	0.085	-3.28	0.001	1.17
Small_toLarge_PS	-0.2661	0.0759	-3.5	0.001	1.2
Supercritical	-0.308	0.18	-1.71	0.09	1.13
Straight	-0.425	0.157	-2.7	0.008	1.17
Adj_High_H2S_US	0.531	0.283	1.88	0.063	1.02

b. **Best subset method:** The result for the best subset method is provided in the inserted file. The best models from the list were identified using the following metrics:

i. Highest 2 adjusted R-square values,

ii. Lowest 2 PRESS, Cp, AIC, and BIC values.

Using the above criteria, 5 potential models were identified and the list of significant predictors was developed. All the predictors from the 5 potential models were considered for the analysis irrespective of the VIF and the p-value to have a more comprehensive list of predictors in the model.

So, using the above method, we were able to shortlist 18 predictors out of the original 29 as important, and these predictors will now be used for model building after the addition of the interactions. The shortlisted variables are provided in Table 5.9, transformed hydrogen sulfide model.

Table 5.9 Selected predictor variables, transformed hydrogen sulfide model

Variables		
Velocity, ft/s		
Manhole Depth, ft		
Precipitation inches		
Design Variables	Drop	>=2' – High drop
		>=0.2', <2' – Medium drop
		<=Std 0.1' – No drop
		Maximum drop height, ft
	Pipe Size Change	Smaller to larger
		Larger to smaller
		Maximum Pipe size change, inches
	Flow Type	Supercritical flow
		Hydraulic jump
	Bends	No Bend 170° - 190° - Straight
	Adjacent to High H_2S manhole	Upstream to high H_2S manhole
Liquid-Phase Variables	Average Liquid Temp., °F	
	Average pH	
	Average BOD, mg/L	
Gas-Phase Variables	Average Temp., °F	

5.2.4 SELECTION OF INTERACTION TERMS

Two or more predictors can have an interaction effect on the response variable; thus, it becomes important to identify the interaction terms and consider them in the model. For the hydrogen sulfide model, the total number of possible interactions terms was 406. It is not feasible to study these many interactions; thus, knowledge from the literature was used to hand pick the interactions, and interactions which showed trends in the partial regression plots were selected.

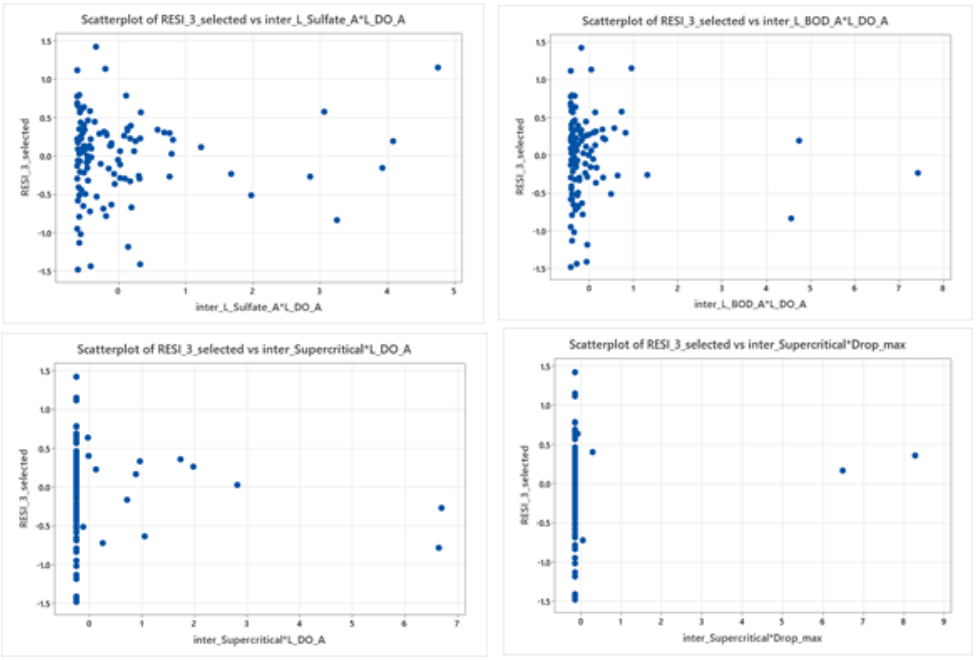
From our understanding of the literature, we expected that if designs or factors which contribute to hydrogen sulfide generation/release existed together in a manhole, then increased gas-phase hydrogen sulfide concentrations might be anticipated. Based on that assumption, the following 12 interactions were considered and partial plots for them were developed. Table 5.10 summarizes the selection reason for the interaction.

Table 5.10 Selected interactions, transformed hydrogen sulfide model

Interaction Considered	Reason
Sulfate*Dissolved oxygen BOD*Dissolved oxygen	High sulfate and BOD in wastewater can deplete the oxygen thus creating anerobic conditions. The assumption was that if a system already lacks oxygen, this relationship may have an enhanced effect on hydrogen sulfide generation/release.
Drop height*Dissolved oxygen Supercritical*Dissolved oxygen Hydraulic jump*Dissolved oxygen Multiple inlets*Dissolved oxygen	Drop height, supercritical flow, hydraulic jump, and multiple inlets would result in turbulence, thus incorporating oxygen into the wastewater, which may limit anaerobic conditions. However, they individually would allow for enhanced hydrogen sulfide release. Thus, a combined effect with dissolved oxygen may change the effect. Thus, they were considered.
Supercritical*Drop height Hydraulic*Drop height Multiple inlets*Hydraulic jump Multiple inlets*Supercritical	Supercritical flow, hydraulic jump, drop height, and multiple inlets contribute to enhanced hydrogen sulfide release. Thus, a manhole with a combined design may show even higher hydrogen sulfide release. Thus, they were considered.
Adjacent to high H_2S DS*DO	Manholes downstream from high hydrogen sulfide manholes may record high hydrogen sulfide levels, and if the wastewater dissolved oxygen is also low, then the hydrogen sulfide generated may be greater; thus, interactions of these terms were considered.
Adjacent to high H_2S DS*Sulfide	Manholes downstream from high hydrogen sulfide manholes may record high hydrogen sulfide levels; in addition, if the sulfide concentrations are high in the wastewater, then if may result in even greater hydrogen sulfide volatilization.

The interactions were standardized to reduce the concern of multicollinearity, which usually occurs with the addition of the interactions. Then the partial plots for the 12 interactions were developed, and if any linear trend was seen, then the interaction was incorporated into the transformed model. Out of the 12 considered, only 5 interactions

showed reasonable trends, as can be seen in the Fig. 5.5. The 5 interactions that showed trends are Sulfate*DO, BOD*DO, multiple inlets*DO, drop height*DO and supercritical flow*DO. However, interactions between supercritical flow*drop height, multiple inlets*hydraulic jump, multiple inlets*supercritical flow were considered even though they did not show any trends; the reason is because they are binary variables thus may not show trends. The combined effect of design may still result in greater hydrogen sulfide release. Thus, a total of 8 interactions were added to the model.



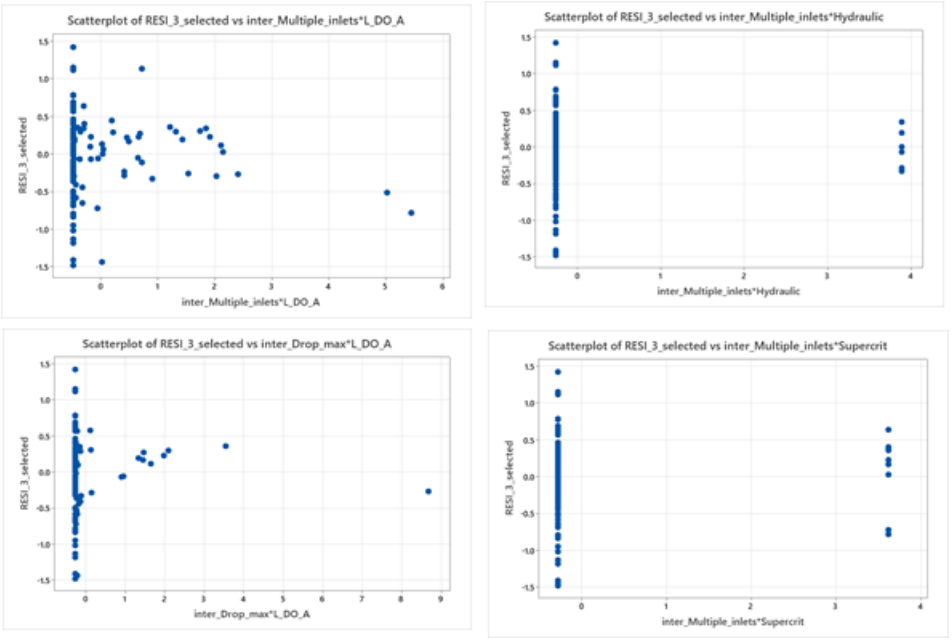


Fig 5.5 Partial regression plots for considered interactions, transformed hydrogen sulfide model

The new model with the 18 predictor variables and 8 interaction terms was considered for further analysis. The result of the new model is presented in Table 5.11. Further analysis was conducted using the model with interactions.

Table 5.11 Coefficients for the model with interactions, transformed hydrogen sulfide model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	3.09	1.21	2.55	0.012	
Velocity	0.364	0.0775	4.7	0	2.38
MH_Depth	-0.0187	0.0116	-1.62	0.11	1.78
Precipitation	-0.493	0.198	-2.49	0.015	1.26
G_Temp_A	0.0307	0.0193	1.59	0.115	10.33
L_Temp_A	-0.0375	0.0259	-1.45	0.151	10.7
L_PH_A	-0.414	0.1	-4.14	0	1.71
L_BOD_A	0.000695	0.00045	1.54	0.126	1.63

High_Drop	-0.843	0.325	-2.6	0.011	4.67
Medium_Drop	-0.279	0.177	-1.58	0.118	2.69
No_Drop	-0.199	0.161	-1.23	0.221	2.77
Drop_max	0.045	0.0358	1.26	0.212	7.1
Small_toLarge_PS	-0.023	0.139	-0.16	0.87	4.28
Large_to_Small_PS	0.15	0.24	0.62	0.534	1.35
PS_Max_Neg	0.0037	0.013	0.29	0.775	3.68
Supercritical	-0.366	0.3	-1.22	0.226	3.29
Hydraulic_Jump	-0.398	0.288	-1.38	0.17	2.68
Straight	-0.497	0.17	-2.92	0.004	1.44
Adj_High_H2S_US	0.57	0.319	1.78	0.078	1.37
inter_L_Sulfate_A*L_DO_A	0.1562	0.0946	1.65	0.102	2.84
inter_L_BOD_A*L_DO_A	-0.1312	0.0931	-1.41	0.162	2.75
inter_Drop_max*L_DO_A	0.0353	0.094	0.38	0.708	2.81
inter_Supercritical*L_DO_A	-0.0932	0.0836	-1.12	0.268	2.22
inter_Supercritical*Drop_max	0.0729	0.0844	0.86	0.39	2.27
inter_Multiple_inlets*L_DO_A	-0.0668	0.0853	-0.78	0.436	2.31
inter_Multiple_inlets*Supercrit	0.006	0.103	0.05	0.957	3.36
inter_Multiple_inlets*Hydraulic	-0.0223	0.0927	-0.24	0.81	2.73

5.2.5 MODEL SEARCH

The model search was conducted using the stepwise regression and best subset methods.

a. **Stepwise method:** The model was run in the stepwise method, and it generated a model containing 7 significant variables. The results of the model are presented in Table 5.12. All the predictors in the stepwise model are significant and the VIF values are also less than 5. Thus, this model will be considered as a candidate for final model selection.

Table 5.12 Output of stepwise model – model search, transformed hydrogen sulfide

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.354	0.592	2.29	0.024	
Velocity	0.2679	0.0559	4.79	0	1.18
Precipitation	-0.375	0.183	-2.04	0.044	1.03
L_PH_A	-0.2913	0.0837	-3.48	0.001	1.15
Small_toLarge_PS	-0.27	0.075	-3.6	0	1.18
Straight	-0.432	0.156	-2.76	0.007	1.17
Adj_High_H2S_US	0.541	0.281	1.93	0.056	1.01
inter_Supercritical*L_DO_A	-0.1223	0.0596	-2.05	0.043	1.08

b. **Best subset method:** The results for the best subset method are provided in the inserted file. The software provides different models containing different combinations of predictors and the goodness-of-fit metrics for each model. These goodness-of-fit metrics are then evaluated for the candidate model selection. The candidate model must have the metrics below:

i. Adjusted R-square should be the highest.

ii. PRESS, AIC, and BIC should be the lowest.

iii. C_p should be the lowest or closer to the number of predictors in the model.

Each model was evaluated on the above metrics and 5 candidate models were chosen. Results for candidate model selection are presented in Tables 5.13.

Table 5.13 Goodness-of-fit metrics of candidate models for transformed hydrogen sulfide, selected using best subset method*

Index	Vars	R-Sq	R-Sq (adj)	PRESS	Mallows Cp	AIC	BIC	Reason to Select/ Reject
M1	7	32.7	28.2	47	10.8	220.25	243.05	Selected for further discussion as similar to Stepwise
M2	9	36	30.4	47.6	9.7	219.54	246.93	Lowest AIC
M3	10	37.1	31	47.8	9.8	219.97	249.58	Reasonable AIC, CP, Adj Rsq
M4	12	39.7	32.4	44.4	9.7	220.46	254.36	Lowest PRESS
M5	13	41.3	33.6	45.6	9.1	220.03	255.99	Reasonable AIC, CP, Adj Rsq, and lowest Cp
M6	19	46.3	35.4	46.8	13	227.12	274.24	Highest Adj Rsq

*Top in each metric category shown in yellow; highlighted in blue is same as stepwise model

Regression was conducted to develop the 6 selected models and they were compared with each other for VIF and P-value to narrow the search to 2 to 3 strong models; it was inferred that only model M1 and M2 met the requirement for low VIF and all significant predictor variables. Coefficients and goodness-of-fit metrics for the 6 generated models are provided in Table 5.14. For Model M2, the predictor variable hydraulic jump had a p-value of 0.108, which is only slightly higher than $\alpha = 0.1$ and from an applicability point of view, it consists of an important design factor (hydraulic jump), and interaction term sulfate*DO; thus, this model seems like a good candidate.

Table 5.14 Candidate models for transformed hydrogen sulfide model, selected using best subset method

M1 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	1.354	0.592	2.29	0.024	
Velocity	0.2679	0.0559	4.79	0	1.18
Precipitation	-0.375	0.183	-2.04	0.044	1.03
L_PH_A	-0.2913	0.0837	-3.48	0.001	1.15
Small_toLarge_PS	-0.27	0.075	-3.6	0	1.18
Straight	-0.432	0.156	-2.76	0.007	1.17
Adj_High_H2S_US	0.541	0.281	1.93	0.056	1.01
inter_Supercritical*L_DO_A	-0.1223	0.0596	-2.05	0.043	1.08
M2 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	1.514	0.605	2.5	0.014	
Velocity	0.3099	0.0583	5.32	0	1.33
Precipitation	-0.444	0.184	-2.41	0.018	1.07
L_PH_A	-0.3251	0.0854	-3.8	0	1.23
Small_toLarge_PS	-0.2374	0.0753	-3.15	0.002	1.23
Straight	-0.376	0.156	-2.41	0.018	1.2
Adj_High_H2S_US	0.497	0.279	1.78	0.078	1.03
inter_Supercritical*L_DO_A	-0.1613	0.0615	-2.62	0.01	1.18
Hydraulic_Jump	-0.311	0.191	-1.62	0.108	1.17
inter_L_Sulfate_A*L_DO_A	0.1072	0.0612	1.75	0.083	1.17
M3 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	1.665	0.612	2.72	0.008	
Velocity	0.3117	0.0581	5.37	0	1.33

Precipitation	-0.434	0.183	-2.37	0.02	1.07
L_PH_A	-0.3449	0.0863	-4	0	1.27
Small_toLarge_PS	-0.2088	0.0779	-2.68	0.009	1.32
Straight	-0.4	0.157	-2.55	0.012	1.21
Adj_High_H2S_US	0.556	0.281	1.98	0.051	1.06
inter_Supercritical*L_DO_A	-0.1634	0.0612	-2.67	0.009	1.19
Hydraulic_Jump	-0.307	0.191	-1.61	0.11	1.17
inter_L_Sulfate_A*L_DO_A	0.1084	0.0609	1.78	0.078	1.17
High_Drop	-0.227	0.165	-1.38	0.172	1.2
M4 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	3.05	1.02	3	0.003	
Velocity	0.2431	0.0563	4.31	0	1.28
Precipitation	-0.419	0.183	-2.29	0.024	1.09
L_PH_A	-0.3374	0.0876	-3.85	0	1.33
Straight	-0.459	0.156	-2.94	0.004	1.23
inter_Supercritical*L_DO_A	-0.1551	0.0584	-2.66	0.009	1.1
High_Drop	-0.845	0.265	-3.19	0.002	3.17
G_Temp_A	0.0344	0.016	2.15	0.034	7.21
L_Temp_A	-0.0462	0.0213	-2.18	0.032	7.34
Medium_Drop	-0.306	0.124	-2.46	0.015	1.35
No_Drop	-0.273	0.109	-2.5	0.014	1.29
inter_Multiple_inlets*Hydraulic	-0.113	0.0595	-1.9	0.06	1.14
Drop_max	0.049	0.0229	2.14	0.035	2.96
M5 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	3.14	1.04	3.01	0.003	

Velocity	0.2574	0.0565	4.56	0	1.3
Precipitation	-0.42	0.184	-2.28	0.025	1.12
L_PH_A	-0.3549	0.089	-3.99	0	1.4
Straight	-0.414	0.156	-2.66	0.009	1.25
inter_Supercritical*L_DO_A	-0.1551	0.0581	-2.67	0.009	1.11
High_Drop	-0.859	0.263	-3.27	0.001	3.16
G_Temp_A	0.0363	0.0158	2.3	0.023	7.14
L_Temp_A	-0.0475	0.0211	-2.25	0.027	7.38
Medium_Drop	-0.324	0.123	-2.64	0.01	1.34
No_Drop	-0.279	0.107	-2.59	0.011	1.27
Drop_max	0.0419	0.023	1.82	0.071	3.03
Adj_High_H2S_US	0.485	0.28	1.73	0.087	1.09
Hydraulic_Jump	-0.354	0.187	-1.9	0.061	1.17
M6 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	2.78	1.12	2.49	0.015	
Velocity	0.3507	0.0657	5.34	0	1.81
Precipitation	-0.448	0.185	-2.41	0.018	1.17
L_PH_A	-0.3887	0.0906	-4.29	0	1.49
Straight	-0.45	0.158	-2.84	0.006	1.33
inter_Supercritical*L_DO_A	-0.1242	0.0744	-1.67	0.098	1.87
High_Drop	-0.908	0.266	-3.42	0.001	3.32
G_Temp_A	0.0286	0.0169	1.7	0.093	8.38
L_Temp_A	-0.0322	0.0229	-1.41	0.163	8.94
Medium_Drop	-0.339	0.123	-2.75	0.007	1.39
No_Drop	-0.277	0.111	-2.51	0.014	1.38
Drop_max	0.0455	0.0242	1.88	0.063	3.44

Adj_High_H2S_US	0.606	0.286	2.12	0.037	1.17
Hydraulic_Jump	-0.432	0.187	-2.31	0.023	1.2
MH_Depth	-0.0197	0.0109	-1.81	0.073	1.66
L_BOD_A	0.000735	0.000428	1.72	0.089	1.56
Supercritical	-0.312	0.231	-1.35	0.18	2.08
inter_L_Sulfate_A*L_DO_A	0.1622	0.0851	1.91	0.06	2.45
inter_L_BOD_A*L_DO_A	-0.1303	0.087	-1.5	0.138	2.56
inter_Supercritical*Drop_max	0.0773	0.0649	1.19	0.237	1.42

5.2.6 MODEL SELECTION

From the above discussion, we select M1 and M2 as the 2 strong candidate models for further analysis. In this section, we will compare the M1 and M2 models based on the model assumptions and the prediction on the test set. The model which performs better in all or most categories will be selected as the final model.

5.2.6.1 M1 MODEL

The results for the M1 model are shown in Table 5.15. From Table 5.15, we can infer that all the predictor variables are significant, as the P-value is less than 0.1 and multicollinearity is not a concern, as VIF is less than 5. Also, the model has R^2 of 32.7% and Adjusted R^2 of 28.2%.

Table 5.15 Coefficients, summary, and ANOVA for M1 model for hydrogen sulfide

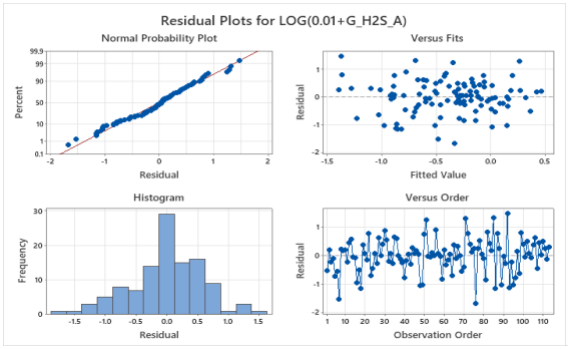
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.354	0.592	2.29	0.024	
Velocity	0.2679	0.0559	4.79	0	1.18
Precipitation	-0.375	0.183	-2.04	0.044	1.03
L_PH_A	-0.2913	0.0837	-3.48	0.001	1.15
Small_toLarge_PS	-0.27	0.075	-3.6	0	1.18
Straight	-0.432	0.156	-2.76	0.007	1.17
Adj_High_H2S_US	0.541	0.281	1.93	0.056	1.01
inter_Supercritical*L_DO_A	-0.1223	0.0596	-2.05	0.043	1.08

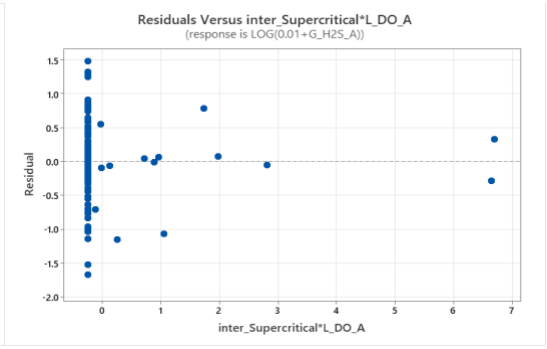
Model Summary					
R-sq	R-sq (adj)	PRESS	R-sq (pred)	AICc	BIC
32.73%	28.25%	46.9689	19.01%	220.25	243.05

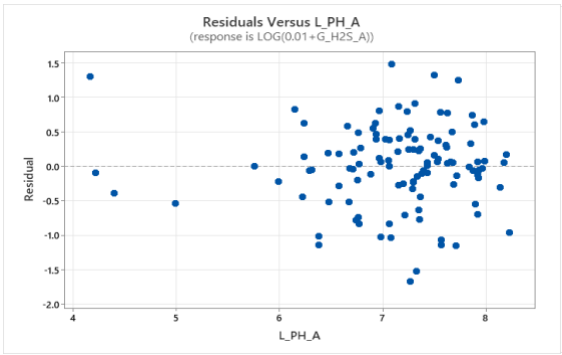
ANOVA							
Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Regression	7	18.984	32.73%	18.984	2.712	7.3	0
Velocity	1	5.46	9.41%	8.523	8.5231	22.94	0
Precipitation	1	1.099	1.90%	1.55	1.5499	4.17	0.044
L_PH_A	1	3.354	5.78%	4.496	4.496	12.1	0.001
Small_toLarge_PS	1	3.252	5.61%	4.814	4.8143	12.96	0
Straight	1	2.704	4.66%	2.827	2.8266	7.61	0.007
Adj_High_H2S_US	1	1.553	2.68%	1.382	1.3822	3.72	0.056
inter_Supercritical*L_DO_A	1	1.562	2.69%	1.562	1.5618	4.2	0.043
Error	105	39.011	67.27%	39.011	0.3715		
Total	112	57.995	100.00%				

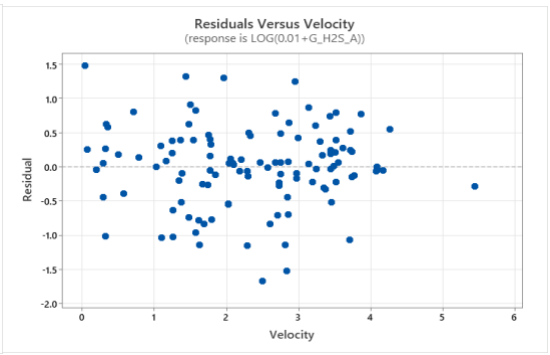
5.2.6.1.1 M1 MODEL ASSUMPTIONS

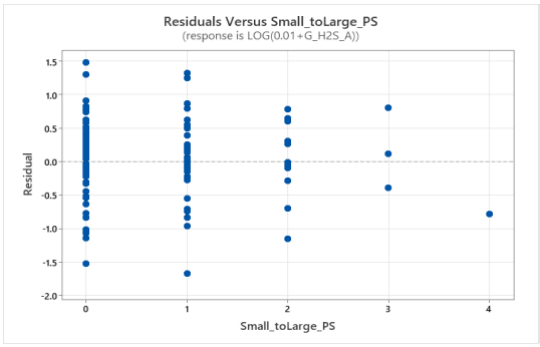
The model must satisfy the model assumptions to be considered. Fig 5.6 shows the residual plots for the M1 model.

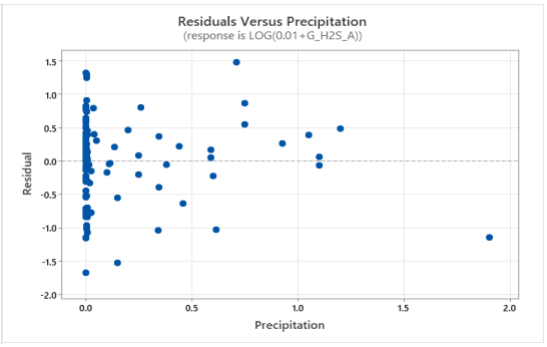












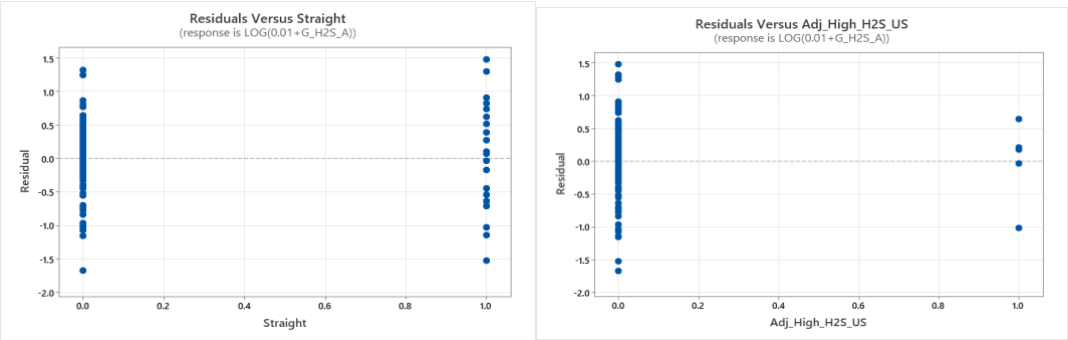


Fig 5.6 Normality plot, residual plot and time plot for M1 for hydrogen sulfide

a. **Residuals are normally distributed:** From Fig 5.6, we can predict that the model is normal. Also, the normality test was done to confirm the normality assumption.

Normality Test: The normality test is conducted to see if the residuals are normally distributed or not.

Hypothesis: H_0: Normality is OK

H_1: Normality is violated

Test: Sample p-value, then we reject H_0.

For the test, we assume confidence level $\alpha = 0.1$. From the software, we obtain p-value = 0.277. Thus, p-value > α, so we fail to reject H_0. So, the normality is satisfied.

b. **Residuals have a constant variance:** From the fitted and residuals plot, we can infer that the model has constant variance. However, to confirm our assumption, we conducted the modified Levene Test, and results for same are below.

Modified Levene Test: For the modified Levene test, we conduct a two-sample t-test on the absolute deviations of the residuals.

i. We first divide our dataset into two sets of almost equal size. The median of the dataset is selected by using fitted values and it is found to be -0.35316, so we divided the dataset into $\hat{y}$ >-0.31697 and $\hat{y}$ <=-0.31697.

ii. Next, we perform the F-Test to check if the variances are equal or not.

Hypothesis: H_0: Variances are equal.

H_1: Variances are not equal.

Test: If p< α, then we reject H_0. For the test, we assume confidence level α = 0.1. Results of the test are shown in Table 5.16.1.

Table 5.16.1 F-Test for equality of variances, M1 hydrogen sulfide model

Equality of Variances				
Method	**Num DF**	**Denom DF**	**p-value**	**α**
F Test	56	55	0.000839	0.1

So, our p (0.000839)< α(0.1); thus, we reject H_0 and the variances are not equal.

iii. After that we calculate the absolute deviations of the residuals around their group medians.

d1 = |e1 – median e1| and d2 = |e2 – median e2|

iv. Next, we perform the T-Test to see if the model has constant variance or not.

Hypothesis: H_0: Error Variances are equal (Means of d1, d2 populations are equal)

H_1: Error Variances are not equal (Means are not equal)

Test: If p< α, then we reject H_0. For the test, we assume confidence level α = 0.1. Results of the test are shown in Table 5.16.2.

Table 5.16.2 T-Test for error variances, M1 hydrogen sulfide model

Variances	**DF**	**t value**	**p-value**
Equal	111	-3.0618	0.00276
Unequal	95.899	-3.0514	**0.002946**

Since the F-Test showed that variances are equal, we select the p-value from the unequal t-test. So, our p (0.002946)< α(0.1); thus, we reject H_0 and the error variances are not equal. This is in contrast to the e vs ŷ plot. Thus, our model fails the constant variance assumption.

c. **Residuals satisfy the MLR form:** The residuals vs. predictor plots had no curvature and showed a good spread, thus satisfying the MLR form. However, the residuals on the precipitation and maximum drop height predictors did not have a good spread, as most of the data points are zero, as seen in Fig 5.6.

d. **Residuals are uncorrelated:** From the Fig 5.6 time plot, we can confirm that residuals are uncorrelated for the M1 model.

e. **Check for outliers:** We need to check for outliers, and the influence they have on the $\hat{y}$ values.

i. **X-Outliers:** To check for X-outliers, leverage values h_{ii} were found, and if the h_{ii} is large, then observation *i* is outlying. Usually used guideline is if $h_{ii} > 2p/n$, then the observation *i* is outlying.

Here, $2p/n = 0.1416$.

The value h_{ii} is the diagonal elements in the matrix H. Matrix $H = X(X^TX)^{-1}X^T$. The values of h_{ii} which are greater than the cutoff value 0.1416 are for 12, thus showing them as x-outliers. Table 5.17 shows the details on the X-outliers.

Table 5.17 X- Outliers for M1 model for hydrogen sulfide

Obs	**HI**	**Cook's D**	**DFITS**	**Intercept**	**Veloc-ity**	**Precipi-tation**	**pH**	**Small to large**	**Straight**	**Adj_High_H2S_US**	**Supercritical*L_DO_A**
9	0.224	0.000	0.179	0.030	-0.045	0.001	-0.018	-0.006	-0.022	0.163	0.012
10	0.220	0.010	0.209	0.008	0.058	0.010	-0.021	0.008	-0.017	0.197	-0.013
35	0.235	0.060	0.675	-0.169	0.072	0.000	0.134	0.206	0.024	0.593	-0.050
49	0.252	0.160	-1.131	-0.300	0.315	0.003	0.190	0.326	0.242	-0.989	-0.111
52	0.256	0.000	-0.030	0.006	-0.005	0.001	-0.005	0.000	-0.011	-0.026	0.001
66	0.171	0.000	-0.076	-0.066	-0.009	0.013	0.066	-0.018	0.006	0.003	0.001
67	0.475	0.060	0.701	0.007	-0.230	0.022	0.067	-0.156	-0.065	0.012	0.675

70	0.190	0.010	-0.341	-0.227	0.024	-0.037	0.228	-0.171	-0.003	0.016	-0.003
71	0.200	0.180	1.216	1.033	0.270	-0.250	-1.054	-0.144	0.371	-0.029	0.018
91	0.461	0.040	-0.590	-0.122	-0.159	0.017	0.155	-0.088	-0.022	-0.015	-0.495
93	0.307	0.280	-1.519	-0.070	-0.089	-1.365	0.146	-0.009	-0.390	-0.061	-0.058
96	0.168	0.050	-0.628	0.046	0.011	0.051	-0.011	-0.592	-0.142	0.067	0.081

ii. Y-Outliers: To check for Y-outliers, we use the studentized deleted residuals method. We calculate the t_i value by deleting each *i* observation in the dataset, where $t_i = d_i/(sq.rt(MSE_i\,(1-h_{ii}))$.

The usual guideline is if $|t_i| > t(1- \alpha/2n; n-p-1)$, then a particular observation is a y-outlier.

Assuming $\alpha = 0.1$, then our cutoff value is:

$t(1- \alpha/2n; n-p-1) = 3.4237$

The value of t_i for each observation of $_i$ is provided in the inserted file. We see that none of the t_i values are higher than the cutoff value. So, we can confidently say that there is no y-outlier in our dataset.

iii. Outliers' Influence on ŷ: The dataset does contain x-outliers; thus, we need to check the influence x-outliers have on ŷ, individual LSEs, and combined LSEs. The effect of outliers is verified through 3 methods – Cook's Distance, DFFITS, and DFBETAS.

- **Cook's Distance:** This method verifies the combined impact of observations *i* on all the LSEs.

  The formula for the Cook's Distance D_i is:

  $D_i = (b-b_{(i)})^T X^T X(b-b_{(i)})/p(MSE) = [e_i^2/p(MSE)]*[h_{ii}/((1-h_{ii})^2)]$

  The values of D_i for various *i* observations are provided in the inserted file. The guideline is if $D_i >$ $F(0.50; p,n-p)$, then the outlier does have influence.

  Here $F(0.50; p, n-p) = 0.92$

  On comparing the D_i values with the cut off, we see that none of the observations exceed the cutoff values; thus, we can say that none of the outliers has influence on the combined LSES.

- **DFFITS:** This method checks the outliers' influence on the fitted values by computing $\hat{y}_{i(i)}$ for every *i* observation being omitted. The formula to calculate DFFITS is:

  $(DFFITS)_i = (\hat{y}_i - \hat{y}_{i(i)})/ sq.rt(MSE_i * h_{ii}) = ti*(sq.rt(h_{ii}/(1- h_{ii})))$

  The values of DFFITS for various *i* observations are provided in the inserted file. The guideline is if |DFFITS| > 2*sq.rt(p/n), then the outlier does have influence on fitted values.

  Here, 2*sq.rt(p/n) = 0.26607

  On comparing the DFFITS values with the cutoff, we see that 8 of the X-outliers exceed the cutoff, so outliers have some influence on the fitted values.

- **DFBETAS:** This method checks the outliers' influence on individual LSEs by computing $b_{k(i)}$ LSE for every *i* observation being omitted. The formula to calculate DFBETAS is:

  $(DFBETAS)_{k(i)} = b_k – b_{k(i)}/ sq.rt(MSE_i * C_{kk})$

  The values of DFBETAS for various *i* observations for each predictor variable, and the intercept, are provided in the inserted file. The guideline is if |DFBETAS| > 2/sq.rt(n), then the outlier does have influence.

  Here, 2/sq.rt(n) = 0.1881

  The DFBETAS values with the cutoff for each of the predictor and the interceptor were compared, to check for influence on outlier on individual LSEs. The DFBETAS values of the X-outliers are shown in the inserted file. The |DFBETAS| value exceeded the cutoff for the following variables:

  a. Intercept for observation 49, 70, 71.
  b. Velocity for observation 49,67, 71.
  c. Precipitation for observation 71, 93.
  d. pH for observation 49,70,71,91,93.
  e. Smaller to larger pipe size change for observation 35,49,96.
  f. Straight for observation 49,71,93.
  g. Upstream to a high hydrogen sulfide manhole for observation 35, 49.
  h. Supercritical*DO for observation 67, 91.

  The DBETAS values are shown in Table 5.7.

f. Predictors not highly correlated: Variance inflation factor is used for confirming multicollinearity among the predictors. From the VIF value in Table 5.15, we can see that none of the predictors has a VIF>5; thus, all the predictors are significant in the model.

5.2.6.1.2 M1 MODEL VALIDATION

Model validation was done by using the test dataset. 20% of the collected data was reserved for validation. Validation helps to estimate the model prediction capability when dataset different from that of training set is used. This helps to identify how close the model predicts to the original values of the test set. In this step, we predict the hydrogen sulfide generated in a manhole using the M1 model and compare it with the original values from the test set. Then we calculate the percentage of difference between the test set measured values and values calculated from the M1 model.

$$\%\ of\ difference = [Avg\left(\frac{\left(|Y - \hat{Y}|\right)}{Y^*} * 100\right)] = 33\%$$

Where:

Y^* = measured hydrogen sulfide concentration in test set

The percentage of difference between the predicted value and the test value was found to be 33%, this shows that modeled hydrogen sulfide values was an average of 33% above or below the field measured hydrogen sulfide values. This demonstrates that the model does a good job at identifying the designs which could be contributing to hydrogen sulfide release and predicting the hydrogen sulfide concentration in a manhole.

5.2.6.2 M2 MODEL

The results for the M2 model are shown in Table 5.18. From Table 5.18, we can infer that all the predictor variables are significant, as P-values are less than 0.1, except for hydraulic jump, which has a P-value of 0.108. However, the value is only 0.008 higher and thus acceptable. No multicollinearity is noticed, as VIF is less than 5. Also, the model has a R-square of 35.9% and Adjusted R-square of 30.4%.

Table 5.18 Coefficients, summary, and ANOVA for M2 model for hydrogen sulfide

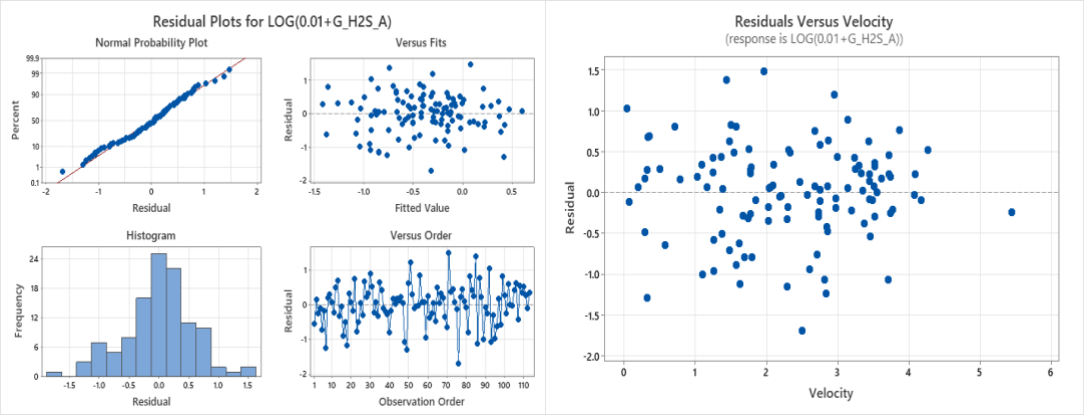
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.514	0.605	2.5	0.014	
Velocity	0.3099	0.0583	5.32	0	1.33
Precipitation	-0.444	0.184	-2.41	0.018	1.07
L_PH_A	-0.3251	0.0854	-3.8	0	1.23
Small_toLarge_PS	-0.2374	0.0753	-3.15	0.002	1.23
Straight	-0.376	0.156	-2.41	0.018	1.2
Adj_High_H2S_US	0.497	0.279	1.78	0.078	1.03
inter_Supercritical*L_DO_A	-0.1613	0.0615	-2.62	0.01	1.18
Hydraulic_Jump	-0.311	0.191	-1.62	0.108	1.17
inter_L_Sulfate_A*L_DO_A	0.1072	0.0612	1.75	0.083	1.17

Model Summary					
R-sq	**R-sq(adj)**	**PRESS**	**R-sq(pred)**	**AICc**	**BIC**
35.97%	30.38%	47.6214	17.89%	219.54	246.92

ANOVA							
Source	**DF**	**Seq SS**	**Contri-bution**	**Adj SS**	**Adj MS**	**F-Value**	**P-Value**
Regression	9	20.8628	35.9%	20.8628	2.3181	6.43	0
Velocity	1	5.4596	9.4%	10.1877	10.1877	28.26	0
Precipitation	1	1.0994	1.9%	2.0976	2.0976	5.82	0.018
L_PH_A	1	3.3545	5.8%	5.219	5.219	14.48	0
Small_toLarge_PS	1	3.2519	5.6%	3.5801	3.5801	9.93	0.002
Straight	1	2.7042	4.7%	2.0919	2.0919	5.8	0.018
Adj_High_H2S_US	1	1.5528	2.7%	1.1438	1.1438	3.17	0.078
inter_Supercritical*L_DO_A	1	1.5618	2.7%	2.48	2.48	6.88	0.01
Hydraulic_Jump	1	0.7714	1.3%	0.9503	0.9503	2.64	0.108
inter_L_Sulfate_A*L_DO_A	1	1.1071	1.9%	1.1071	1.1071	3.07	0.083
Error	103	37.1326	64.0%	37.1326	0.3605		
Total	112	57.9954	100.0%				

5.2.6.2.1 M2 MODEL ASSUMPTIONS

The model must satisfy the model assumptions to be considered. Fig 5.7 shows the residual plots for the M2 model.



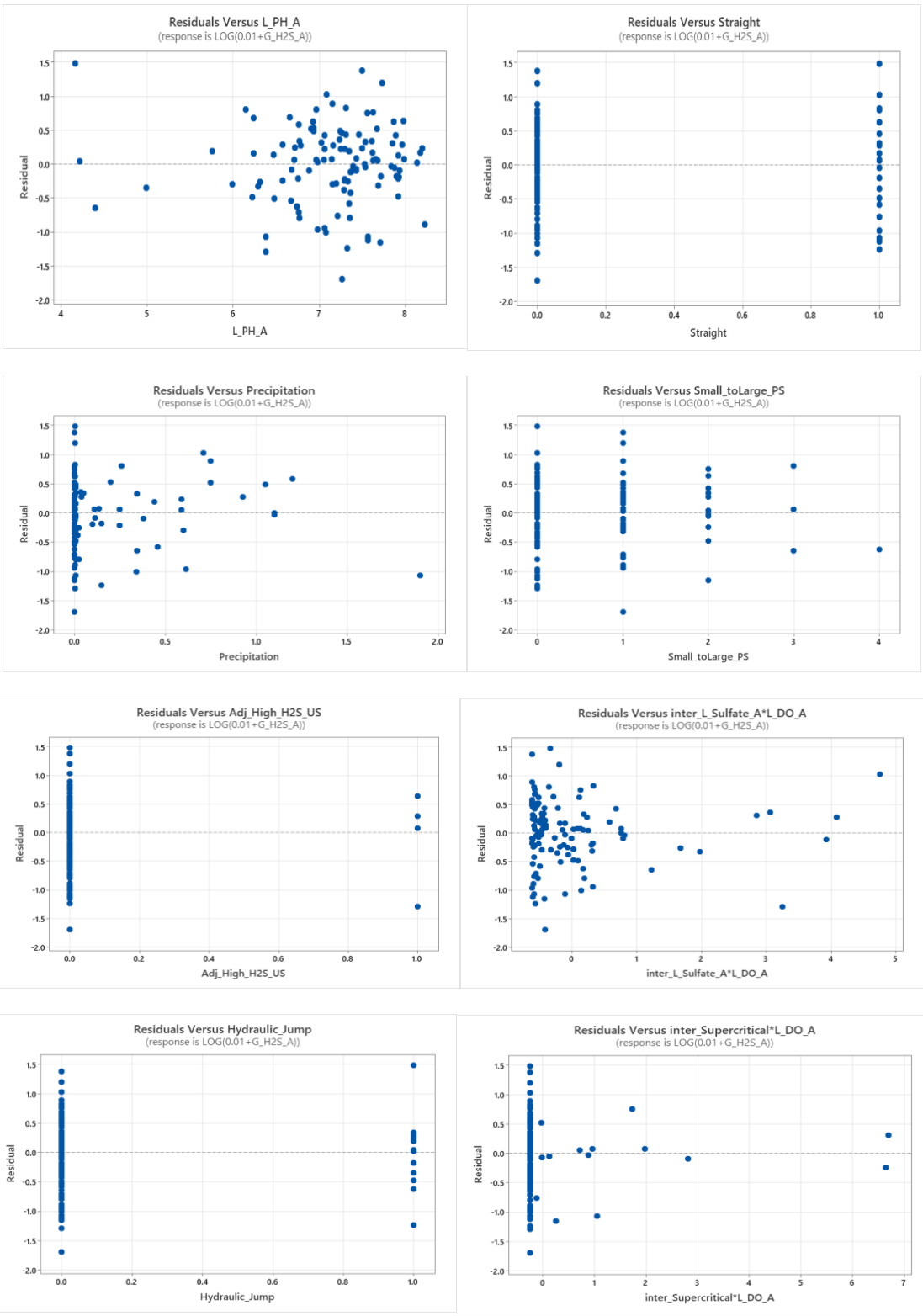


Fig 5.7 Normality plot, residual plot and time plot for M2 model for hydrogen sulfide

a. **Residuals are normally distributed:** From Fig 5.7, we can predict that the model is normal. Also, the normality test was conducted to confirm the normality assumption.

Normality Test: Normality test is conducted to see if the residuals are normally distributed or not.

Hypothesis: H_0: Normality is OK

H_1: Normality is violated

Test: Sample p-value<α, then we reject H_0.

For the test, we assume confidence level $\alpha = 0.1$. From the software, we get the p-value = 0.142. Thus, p-value >α, so we fail to reject H_0. So, the normality is satisfied.

b. **Residuals have a constant variance:** From the fitted and residuals plot, we can infer that the model has constant variance. However, to confirm our assumption we conducted the modified Levene Test, and results for same are below.

Modified Levene Test: For modified Levene test, we conduct a two-sample t-test on the absolute deviations of the residuals.

i. We first divide our dataset into two sets of almost equal size. The median of the dataset is selected by using fitted values and found to be -0.3912, so we divide the dataset into $\hat{y}$ >-0.3912 and $\hat{y}$ <=-0.3912 Next, we perform the F-Test to check if the variances are equal or not.

Hypothesis: H_0: Variances are equal.

H_1: Variances are not equal.

Test: If p< α, then we reject H_0. For the test, we assume confidence level $\alpha = 0.1$. Results of the test are shown in Table 5.19.1.

Table 5.19.1 F-Test for equality of variances for the M2 model for hydrogen sulfide

Equality of Variances				
Method	**Num DF**	**Denom DF**	**p-value**	**α**
F Test	56	55	0.4337	0.1

So, our p (0.4337)> α(0.1); thus, we reject H_0 and the variances are not equal.

ii. After that, we calculate the absolute deviations of the residuals around their group medians.

d1 = |e1 – median e1| and d2 = |e2 – median e2|

iii. Next, we perform the T-Test to see if the model has constant variance or not.

Hypothesis: H_0: Error Variances are equal (Means of d1, d2 populations are equal).

H_1: Error Variances are not equal (Means are not equal).

Test: If p< α then we reject H_0. For the test, we assume confidence level α = 0.1. Results of the test are shown in Table 5.19.2.

Table 5.19.2 T-Test for error variances for M2 model for hydrogen sulfide

Variances	DF	t value	p-value
Equal	111	-1.1485	**0.2532**
Unequal	110.34	-1.1484	0.2533

Since the F-Test showed that variances are equal, we select the p-value from equal T-Test. So, our p (0.2532)> α(0.1); thus, we reject H_0 and the error variances are not equal. This contrasts with the e vs ŷ plot. Thus, our model passes the constant variance assumption.

c. Residuals satisfy the MLR form: The residuals vs. predictor plots had no curvature and showed a good spread, thus satisfying the MLR form. However, the residuals on the precipitation plot did not have a good spread as most of the data points were zero, as seen in Fig 5.7.

d. Residuals are uncorrelated: From the Fig 5.7, time plot, we can confirm that residuals are uncorrelated for the M2 model.

e. Check for outliers: We need to check for outliers, and the influence they have on the ŷ values.

i. **X-Outliers:** To check for X-outliers, leverage values h_{ii} were found. If the h_{ii} is large, then observation *i* is outlying. The usually used guideline is if $h_{ii} > 2p/n$, then the observation *i* is outlying.

Here, $2p/n = 0.1769$.

The value h_{ii} is the diagonal elements in the matrix H. Matrix $H = X(X^TX)^{-1}X^T$. The values of h_{ii} which are greater than the cutoff value 0.1769 are for 14, thus showing them as x-outliers. Table 5.20 shows the details on the X-outliers.

Table 5.20 X- Outliers for M2 model for hydrogen sulfide

Obs	HI	Cook's D	DFITS	Intercept	Velocity	Precipitation	pH	Small_toLarge_PS	Straight	Adj_High_H2S_US	Supercritical *DO	Sulfate* DO	Hydraulic_Jump
9	0.252	0.010	0.318	0.065	-0.095	0.001	-0.039	-0.016	-0.039	0.286	0.043	-0.098	-0.031
10	0.230	0.000	0.079	0.005	0.025	0.001	-0.011	0.006	-0.003	0.071	-0.009	0.011	-0.013
35	0.240	0.050	0.687	-0.138	0.063	-0.012	0.108	0.210	0.032	0.602	-0.041	-0.058	-0.082
49	0.301	0.280	-1.729	-0.401	0.193	0.079	0.302	0.343	0.256	-1.288	0.056	-0.682	0.218
52	0.326	0.020	0.401	-0.115	0.011	0.020	0.102	-0.037	0.099	0.301	0.023	-0.062	0.181
67	0.488	0.050	0.706	-0.019	-0.203	0.030	0.084	-0.155	-0.069	-0.003	0.626	0.082	0.073
70	0.224	0.040	-0.650	-0.424	-0.036	-0.019	0.439	-0.344	-0.047	0.039	0.060	-0.162	0.212
71	0.224	0.230	1.554	1.089	0.195	-0.201	-1.100	-0.265	0.358	-0.037	0.114	-0.158	0.500
75	0.182	0.000	-0.101	0.013	0.021	0.010	-0.020	-0.014	0.002	0.019	0.021	-0.083	0.009
91	0.473	0.030	-0.517	-0.116	-0.115	0.017	0.140	-0.073	-0.019	-0.024	-0.424	0.068	0.035
92	0.299	0.180	1.358	-0.173	-0.190	0.352	0.197	0.087	0.470	-0.178	-0.189	1.062	-0.095
93	0.311	0.210	-1.466	-0.073	-0.033	-1.299	0.132	0.012	-0.356	-0.078	-0.098	0.159	-0.022
96	0.212	0.040	-0.607	0.110	0.035	-0.002	-0.080	-0.461	-0.083	0.075	0.049	-0.039	-0.270
100	0.254	0.010	0.308	-0.042	0.012	-0.001	0.048	-0.060	-0.056	-0.051	-0.049	0.227	0.162

ii. Y-Outliers: To check for Y-outliers, we use the studentized deleted residuals method. We calculate the t_i value by deleting each *i* observation in the dataset, where $t_i = d_i/(sq.rt(MSE_i\,(1-h_{ii}))$.

The usual guideline is if $|t_i| > t(1- \alpha/2n; n-p-1)$, then a particular observation is a y-outlier.

Assuming $\alpha = 0.1$, then our cutoff value is:

$t(1- \alpha/2n; n-p-1) = 3.4257$

The value of t_i for each observation of $_i$ is provided in the inserted file. We see that none of the t_i values are higher than the cutoff value. So, we can confidently say that there is no y-outlier in our dataset.

iii. Outliers' Influence on ŷ: The dataset does contain x-outliers; thus, we need to check the influence that the x-outliers have on ŷ, individual LSEs, and combined LSEs. The effect of outliers is verified through 3 methods – Cook's Distance, DFFITS, and DFBETAS.

- **Cook's Distance:** This method verifies the combined impact of observation *i* on all the LSEs. The formula for the Cook's Distance D_i is:

  $D_i = (b-b_{(i)})^T X^T X(b-b_{(i)})/p(MSE) = [e_i^2/p(MSE)]*[h_{ii}/((1-h_{ii})^2)]$

  The values of D_i for various *i* observations are provided in the inserted file. The guideline is if $D_i > F(0.50; p,n-p)$, then the outlier does have influence.

  Here $F(0.50; p, n-p) = 0.94$.

  On comparing the D_i values with the cut off, we see that none of the observations exceeds the cutoff values; thus, we can say that none of the outliers has influence on the combined LSES.

- **DFFITS:** This method checks the outliers' influence on the fitted values by computing $\hat{y}_{i(i)}$ for every *i* observation being omitted. The formula to calculate DFFITS is:

  $(DFFITS)_i = (\hat{y}_i - \hat{y}_{i(i)})/ sq.rt(MSE_i * h_{ii}) = ti*(sq.rt(h_{ii}/(1- h_{ii})))$

  The values of DFFITS for various *i* observations are provided in the inserted file. The guideline is if $|DFFITS| > 2*sq.rt(p/n)$, then the outlier does have influence on fitted values.

  Here, $2*sq.rt(p/n) = 0.5949$

On comparing the DFFITS values with the cutoff, we see that observations 35, 49, 67, 70, 71, 91, 92, 93, and 96 exceed the cutoff, so the X-outliers have some influence on the fitted values.

- **DFBETAS:** This method checks the outliers' influence on individual LSEs by computing $b_{k(i)}$ LSE for every *i* observation being omitted. The formula to calculate DFBETAS is:

  $(DFBETAS)_{k(i)} = b_k - b_{k(i)} / sq.rt(MSE_i * C_{kk})$

  The values of DFBETAS for various *i* observations for each predictor variable, and the intercept, are provided in the inserted file. The guideline is if |DFBETAS| > 2/sq.rt(n), then the outlier does have influence.

  Here, 2/sq.rt(n) = 0.1881

  The DFBETAS values with the cutoff for each of the predictors and the interceptor were compared, to check for influence of outliers on individual LSEs. The DFBETAS value of the X-outlier shown in the inserted file. The |DFBETAS| value exceeded the cutoff for the following variables.:

  i. Intercept for observations 49, 70, 71.
  ii. Velocity for observations 49,67, 71, 92.
  iii. Precipitation for observations 71, 92, 93.
  iv. pH for observations 35, 49,70,71,96.
  v. Smaller to larger pipe size change for observations 35,49,70, 71, 96.
  vi. Straight for observations 49,71,92, 93.
  vii. Upstream to a high hydrogen sulfide manhole for observations 9, 35, 49, 52, 92.
  viii. Supercritical*DO for observations 67, 91, 92.
  ix. Sulfate*DO for observations 49, 92, 100.
  x. Hydraulic jump for observations 49, 70, 71, 96.

  The DBETAS values are shown in Table 5.20.

f. Predictors not highly correlated: Variance inflation factor is used for evaluating multicollinearity between the predictors. From the VIF value in Table 5.15, we can see that none of the predictors has a VIF>5; thus, all the predictors are significant in the model.

5.2.6.2.2 M2 MODEL VALIDATION

Model validation was done by using the test dataset. 20% of the collected data was reserved for validation. Validation helps to estimate the model prediction capability when dataset different from that of training set is used. This helps to identify how close the model predicts to the original values of the test set. In this step, we predict the hydrogen sulfide generated in a manhole using the M2 model and compare it with the original values from the test set. Then we calculate the percentage of difference between the test set measured values and values calculated from the M2 model.

$$\%\ of\ difference = [Avg\left(\frac{(|Y - \hat{Y}|)}{Y^*} * 100\right)] = 42.98\%$$

Where:

Y^* = measured hydrogen sulfide concentration in test set

The percentage of difference between the predicted value and the test value was found to be 42.98%, this shows that modeled hydrogen sulfide values was an average of 42.98% above or below the field measured hydrogen sulfide values. This demonstrates that the model does a good job at identifying the designs which could be contributing to hydrogen sulfide release and predicting the hydrogen sulfide concentration in a manhole.

5.2.6.3 M1 and M2 MODEL COMPARISON

This section compares the model M1 and M2 for model assumptions, outliers, and goodness of fit parameters, to make the final model selection. Table 5.21 provides details on the metric comments for model M1 and M2. From Table 5.21, we can clearly see that both the M1 and M2 model are very similar; however, M2 model performed better overall, since it had a better R-square, Adjusted R-square, AIC, Cp values. The M2 model had predictor variable hydraulic jump, which had a slightly higher p-value; however, it was not too great to reject the model. Also,

both the model had similar percentage of difference in predicting the hydrogen sulfide concentrations for the test set with M1 showing % of difference as 33% and M2 model showing a value of 43%. In contrast, the M1 model completely failed the modified Levene's test, which is an important to confirm the variance between the residuals is constant. Non – constant variance will result in incorrect prediction. Thus, the M2 model was selected as the final hydrogen sulfide model.

Table 5.21 M1 and M2 hydrogen sulfide model comparison

Index	M1	M2	Comments
Correlation between predictors	Highest between predictors 31%	Highest between predictors 31%	Both the models perform equally well, as the correlation between the predictors is less than 70%.
Correlation of predictors with Y	Highest 33%	Highest 33%	Both models show similar result, thus both seem acceptable.
Residual analysis	No clear curvature	No clear curvature	Both models perform well.
Constant variance - plot	Satisfied	Satisfied	Both models perform well.
Levene Test	Failed	Passed	M1 fails the modified Levene test, and an important assumption; thus, M2 is a better model.
VIF	Highest 1.18	Highest 1.23	Both models have VIF less than 5 so both are good.
P-value	Highest 0.056	Highest 0.108	For M2, one of the predictors has a p-value slightly higher; however, it is not too high, thus acceptable.
Normality	Satisfied	Satisfied	Both models perform equally well.
Normality Test	Satisfied	Satisfied	Both models perform equally well.
Rsq	32.73%	35.97%	M2 performs better.
Adj R sq	28.25%	30.38%	M2 performs better.

Cp	10.8	9.7	M2 performs better as the Cp value should be lesser for a good model.
PRESS (Train)	46.9689	47.6214	M1 performs better since the PRESS value should be lesser for a good model.
AIC	220.25	219.54	M2 performs better since the AIC value should be lesser for a good model.
BIC	243.05	246.92	M1 performs better since the AIC value should be lesser for a good model.
High hii	12	15	M2 model is sensitive to outliers; thus, M1 model performed better.
Y-outlier	None	None	Both models performed equally well.
Outlier influence	8	10	M2 model is sensitive to outliers; thus, M1 model performed better.
% of difference	33%	42.98%	M1 model had a lower % of difference.
Rsq Pred (test)	0.01	0.03	M2 model had a better prediction.

5.2.7 MODEL INTERPRETATION

From the above analysis, we were able to develop a hydrogen sulfide model. The model to predict hydrogen sulfide concentrations in manholes is presented below:

$$Log_{10}(0.01 + Y) = 1.514 + 0.3099V - 0.444Pr - 0.3251pH - 0.2374PS_{sl} - 0.376St + 0.497US_{H2S} - 0.1613(Super * DO) - 0.311Flow_{hydrualic} + 0.1072(Sulfate * DO)$$ ……Eqn. 5.2

Where:

Y = Gas phase hydrogen sulfide concentration in the manhole, ppm

V = Velocity of wastewater, ft/s

Pr = Precipitation, inches

pH = pH of the wastewater

PS_{sl} = Change in pipe size diameter (inlet pipe smaller than outlet pipe) (binary)

St = Absence of bends and multiple inlets (binary)

US_{H2S} = Upstream of a high hydrogen sulfide manhole (binary)

Supercritical*DO = Interaction between supercritical flow and dissolved oxygen (mg/l)

$Flow_{hydraulic}$ = Presence of hydraulic jump (binary)

Sulfate*DO = Interaction between the liquid sulfate concentration and dissolved oxygen (mg/l)

The model predicts manhole hydrogen sulfide concentration in increases with wastewater velocity (which produces more turbulence) and if the manhole is upstream of a high H_2S manhole. The H_2S concentration decreases with precipitation, wastewater pH, inlet pipe being smaller than the outlet pipe, absence of bends (straight flow) and hydraulic jump. Precipitation would absorb gas-phase H_2S, decreasing its concentration; however, this could also mean an increased corrosion because the rainwater on the manhole surface provides environment for faster microbial activity. As pH increases, more H_2S is present in dissociated form, so it is less likely to volatilize. Absence of bends and multiple inlets means less turbulent flow, which would decrease volatilization of H_2S. The model suggests that increasing the number of inlets pipe smaller than outlet may decrease the hydrogen sulfide concentration. The reason that a pipe size change or presence of hydraulic jump decreases hydrogen sulfide concentration is unclear. However, it is interesting to note that both pipe size change and hydraulic jump reduce the hydrogen sulfide concentrations. The flow in manhole with inlet pipe smaller than outlet pipe will experience flow conditions similar to a hydraulic jump, as the flow velocity will reduce with increase in pipe diameter.

An interesting result of the model was that hydrogen sulfide concentration decreased with increased interaction between the supercritical flow and dissolved oxygen, which sounds reasonable as supercritical flow can incorporate dissolved oxygen into the wastewater via agitation. This would decrease formation of H_2S in the liquid phase. The turbulence associated with supercritical flow would also transfer a large portion of H_2S to the gas phase once formed, but the impact of incorporating oxygen appears to dominate.

As seen in Eq. 5.2, H_2S concentration decreases with interaction between supercritical flow and dissolved oxygen, and increases with interaction between sulfate and dissolved oxygen The interaction between sulfate and dissolved

oxygen implies that as dissolved oxygen decreases, sulfate decreases (two negative terms multiplied together produce a positive term), which makes sense. As dissolve oxygen decreases, microorganisms will use chemically-bound oxygen for respiration, which decreases sulfate concentrations. The sulfate converts to sulfide, which increases hydrogen sulfide generation and volatilization in the manhole's headspace.

The model predicts the hydrogen sulfide values 42.98% above or below the test set; thus, it was able to identify major variables responsible for hydrogen sulfide generation and volatilization.

The primary importance of the study is to identify the key variables that primarily contribute towards the hydrogen sulfide generation and volatilization in manholes. As clear from the literature, there are various sewer system design characteristics and wastewater characteristics that play a role in regulating the system, and thus this research identifies major variables that can be controlled or regulated to reduce hydrogen sulfide generation and volatilization in manholes or measured regularly to understand the condition and life expectancy of the manholes. Also, comparison of the Equation 5.2 with the equations for hydrogen sulfide prediction in sewer pipes demonstrates that the sewer pipes equations do not apply to manholes, the reason being manhole design dictates the predictors that will control hydrogen sulfide generation and volatilization in manhole headspace; thus, for a holistic view, incorporation of design is important. As discussed in Chapter 2 Section 2.5, we can see that none of the previously developed equations consider difference in sewer system design as a factor. Also, the design of the manhole is fundamentally different from that of the sewer system because a manhole has a vertical design which is not submerged in wastewater but is exposed to hydrogen sulfide and has good concentrations of oxygen, as it near the ground surface. In sewer pipes equations such as 1977 Pomeroy-Parkhurst equation, Pomeroy Z formula, and Thistlewayte's equation, a submerged hydraulic radius is an important factor; this factor does not apply to manholes. The only study that considered both submerged and unsubmerged sections of the pipe was by Jiang et al. [47],;however in this study only 3 different temperatures and 2 relative humidity levels were considered. Also, designs such as high drop, or multiple inlets may not have the same effect in sewer pipe because of non-existent of the vertical structure in sewer pipes.

5.3 CORROSION MODEL

A multiple linear regression model of the form shown in Eqn. 5.3 was used:

$$Y = \beta_0 + \beta_1 x_1 \ldots . + \beta_{31} x_{31} + \varepsilon \ldots\ldots \text{Eqn. 5.3}$$

Where:

Y = the corrosion rate in mm/yr (response variable),

β_o = intercept,

$\beta_1, \beta_2, \ldots . \beta_{31}$ = slopes for the respective predictor variables,

$x_1, x_2, \ldots x_{31}$ are the predictor variables (see Table 5.22),

ε = random error term.

Table 5.22 Predictor variables in the corrosion model

Variables			**Notation**	**Type of data**
Flow Rate, MGD			x_1	Continuous
Velocity, ft/s			x_2	Continuous
Depth of Flow, ft			x_3	Continuous
Manhole Depth, ft			x_4	Continuous
Average Ambient Temperature, °F			x_5	Continuous
Precipitation inches			x_6	Continuous
Design Variables	Drop	>=2' – High drop	x_7	Counters
		>=0.2', <2' – Medium drop	x_8	Counters
		<=Std 0.1' – No drop	x_9	Counters
		Maximum drop height, ft	x_{10}	Continuous
	Pipe Size Change	Smaller to larger	x_{11}	Counters
		Larger to smaller	x_{12}	Counters
		Uniform	x_{13}	Counters

		Maximum Pipe size change, inches	x_{14}	Continuous
	Flow Type	Supercritical flow	x_{15}	Binary
		Hydraulic jump	x_{16}	Binary
	Bends	Multiple inlets	x_{17}	Binary
		80-110° Bend – Right bend	x_{18}	Binary
		>110° Bend – Obtuse bend	x_{19}	Binary
		No Bend 170° - 190° - Straight	x_{20}	Binary
	Adjacent to High H_2S manhole	Upstream to high H_2S manhole	x_{21}	Binary
		Downstream to high H_2S manhole	x_{22}	Binary
Liquid-Phase Variables	Average Liquid Temp., °F		x_{23}	Continuous
	Average Liquid DO, mg/L		x_{24}	Continuous
	Average pH		x_{25}	Continuous
	Average Sulfide, mg/L		x_{26}	Continuous
	Average Sulfate, mg/L		x_{27}	Continuous
	Average BOD, mg/L		x_{28}	Continuous
Gas-Phase Variables	Average Temp., °F		x_{29}	Continuous
Manhole Age			x_{30}	Continuous
Average hydrogen sulfide concentrations, ppm			x_{31}	Continuous

The slope demonstrates the relationship between the predictor variable and the rate of hydrogen sulfide (response variable). A manhole can have either supercritical flow, subcritical flow, or a hydraulic jump but never all three; thus, to avoid multicollinearity, subcritical flow is not considered in model building. Similarly, a manhole can either have multiple inlets or bends (acute, right, obtuse, straight) but never both; thus, to avoid multicollinearity, acute angle bends are not considered in model building process.

For the corrosion model, we divided the dataset into 2 sets: one for the training set, which consisted of 101 manholes, and test set, which consisted of 25 manholes. Test set dataset selection was done randomly with the use of software to avoid bias in the model. The model was built using the training set and validated for robustness with the test set.

5.3.1 PRELIMINARY ANALYSIS

For the preliminary analysis, the model was developed using all the 31 predictor variables. Data used for model development and validation available on request from the author.

5.3.1.1 CORRELATION MATRIX

Before building the model, the correlation between the predictors and the response variable was evaluated, to identify any serious multicollinearity concerns between the predictors. Any correlation higher than 70% (0.7) between the predictors shows concerns of linear dependencies and may affect model accuracy. Higher collinearity between the response and predictor variables is better as it improves the accuracy of the model.

Correlation between response and predictor variables: The highest observed correlation was 27.4% between the rate of corrosion and the predictor variable velocity, which could be because velocity was a significant factor in hydrogen sulfide model. Since velocity regulates hydrogen sulfide generation, it indirectly also determines the rate of corrosion. The second highest correlation of 25.9% was seen between absence of a drop and corrosion rate, which supports our assumption of Chapter 4 that a manhole environment not much disturbed by turbulence may allow for more microbial activity, resulting in a faster corrosion rate. All the other predictors recorded less than 20% of correlation with the response variable.

Correlation between the predictors: Flowrate and depth of flow showed a correlation of 93%, gas phase temperature with ambient weather showed a correlation of 81.6%, gas phase temperature with liquid phase temperature showed a correlation of 93.6%, ambient weather with liquid phase temperature showed a correlation of 76.3%, and high drop with maximum drop height showed a correlation of 70%. These high correlations between the temperature predictors seems reasonable, as it was seen in Chapter 4 that weather effects the gas and liquid phase

temperatures. All the other predictors recorded correlations less than 70% between each other. Since there is definitely multicollinearity between variables, we need to check for variance inflation factor (VIF) in the final model.

5.3.1.2 FITTING THE PRELIMINARY MODEL

Table 5.23 shows the results of the preliminary model and Table 5.24 shows the Analysis of Variance table. Table 5.23 confirms that 30 predictors have P-value greater than 0.1, and 14 predictors have a VIF value of greater than 5, thus signifying multicollinearity concerns.

Table 5.23 Estimates of coefficients for the predictor variables – preliminary corrosion rate model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-12.2	12.3	-0.99	0.328	
Flow_Rate	0.054	0.491	0.11	0.912	12.37
Velocity	1.843	0.647	2.85	0.006	3.71
Depth_of_flow	-1.39	3.66	-0.38	0.706	14.31
MH_Depth	-0.0677	0.0929	-0.73	0.469	2.01
Age_H2S	-0.0441	0.0286	-1.54	0.128	1.86
W_Temp_A	-0.0061	0.0483	-0.13	0.9	4.71
Precipitation	1.26	1.37	0.92	0.361	1.37
G_Temp_A	0.039	0.166	0.23	0.816	16.41
G_H2S_A	-0.029	0.184	-0.16	0.877	2.17
L_Temp_A	0.07	0.212	0.33	0.742	14.95
L_DO_A	0.623	0.26	2.4	0.019	1.43
L_PH_A	-0.034	0.849	-0.04	0.968	2.33
L_Sulfide_A	0.00921	0.00962	0.96	0.342	1.56
L_Sulfate_A	-0.0019	0.0166	-0.12	0.908	1.83
L_BOD_A	0.0019	0.00242	0.79	0.435	1.38

High_Drop	1.46	4.94	0.29	0.769	17.82
Medium_Drop	1.66	4.35	0.38	0.705	31.21
No_Drop	0.28	4.46	0.06	0.951	49.46
Drop_max	-0.272	0.268	-1.02	0.313	3.4
Small_toLarge_PS	1	4.71	0.21	0.833	97.63
Large_to_Small_PS	-1.1	5.43	-0.2	0.841	12.76
Uniform	1.08	4.83	0.22	0.823	57.98
PS_Max_Negative	-0.0731	0.0905	-0.81	0.422	3
Supercritical	-2.55	1.69	-1.51	0.136	2
Hydraulic_Jump	0.58	1.8	0.32	0.748	1.93
Right	1.31	4.4	0.3	0.767	16.44
Obtuse	4.4	4.36	1.01	0.316	19.46
Multiple_inlets	-0.93	4.79	-0.19	0.847	37.45
Straight	3.82	4.33	0.88	0.38	22.73
Adj_High_H2S_US	0.57	2.08	0.27	0.786	1.88
Adj_High_H2S_DS	1.77	2.31	0.76	0.447	2.01

Table 5.24 ANOVA for preliminary corrosion rate model

Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	**F-Value**	**P-Value**
Regression	31	658.98	38.81%	658.98	21.257	1.41	0.118
Flow_Rate	1	60.4	3.56%	0.18	0.183	0.01	0.912
Velocity	1	71.73	4.22%	122.04	122.043	8.11	0.006
Depth_of_flow	1	0.14	0.01%	2.16	2.164	0.14	0.706
MH_Depth	1	18.77	1.11%	8	7.995	0.53	0.469
Age_H2S	1	56.39	3.32%	35.66	35.662	2.37	0.128
W_Temp_A	1	7.11	0.42%	0.24	0.238	0.02	0.9
Precipitation	1	3.25	0.19%	12.73	12.729	0.85	0.361

G_Temp_A	1	3.27	0.19%	0.83	0.826	0.05	0.816
G_H2S_A	1	5.07	0.30%	0.36	0.365	0.02	0.877
L_Temp_A	1	9.93	0.58%	1.64	1.645	0.11	0.742
L_DO_A	1	46.61	2.75%	86.89	86.888	5.77	0.019
L_PH_A	1	4.02	0.24%	0.02	0.024	0	0.968
L_Sulfide_A	1	17.52	1.03%	13.8	13.805	0.92	0.342
L_Sulfate_A	1	3.37	0.20%	0.2	0.203	0.01	0.908
L_BOD_A	1	10.94	0.64%	9.29	9.29	0.62	0.435
High_Drop	1	28.11	1.66%	1.31	1.309	0.09	0.769
Medium_Drop	1	13.91	0.82%	2.18	2.179	0.14	0.705
No_Drop	1	36.07	2.12%	0.06	0.058	0	0.951
Drop_max	1	26.45	1.56%	15.55	15.551	1.03	0.313
Small_toLarge_PS	1	0.4	0.02%	0.67	0.673	0.04	0.833
Large_to_Small_PS	1	51.4	3.03%	0.61	0.614	0.04	0.841
Uniform	1	2.08	0.12%	0.76	0.756	0.05	0.823
PS_Max_Negative	1	7.4	0.44%	9.82	9.821	0.65	0.422
Supercritical	1	37.04	2.18%	34.29	34.295	2.28	0.136
Hydraulic_Jump	1	0.49	0.03%	1.57	1.573	0.1	0.748
Right	1	45.34	2.67%	1.34	1.335	0.09	0.767
Obtuse	1	17.86	1.05%	15.34	15.343	1.02	0.316
Multiple_inlets	1	52.06	3.07%	0.56	0.565	0.04	0.847
Straight	1	11.95	0.70%	11.73	11.732	0.78	0.38
Adj_High_H2S_US	1	1.1	0.06%	1.11	1.115	0.07	0.786
Adj_High_H2S_DS	1	8.79	0.52%	8.79	8.79	0.58	0.447
Error	69	1038.9	61.19%	1038.9	15.056		
Total	100	1697.88	100.00%				

5.3.1.3 PRELIMINARY MODEL – ASSUMPTIONS CHECK

Before the preliminary model can be considered for further analysis, the model must satisfy the model assumptions. The first assumption that was validated if the model form was reasonable. As clear from Fig 5.8, the model is not normal and has a long tail on the top. Also, the spread of the fitted versus the residual plot was not clear; there is a need for transformation on the response variable. No curvature was seen on predictor plots. Since the model failed the first 2 assumptions, transformation was carried out to fix the concerns. Even the residual plots of the predictor variable did not show a good spread.

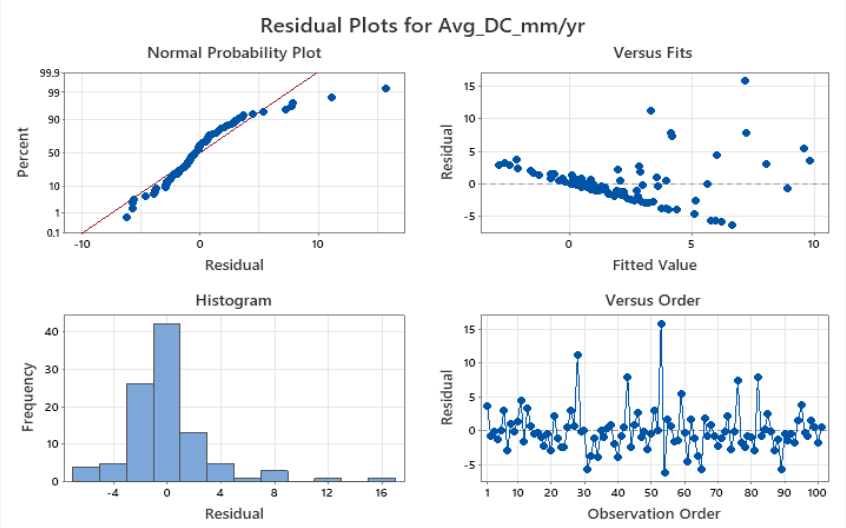


Fig 5.8 Residuals plots for the preliminary model

5.3.2 TRANSFORMED MODEL

Before proceeding further with the analysis, we need to fix the concerns for normality and the spread on the fitted vs residual plots. A compression of the response variable will result in a more linear relationship. Any transformation on Y will result affect the distribution of Y and distribution of the residuals. In our case the normality is not satisfied, so we decided to move forward with a transformation on Y.

However, the transformation on the Y variable did not satisfy the non-constant variance assumption. Thus, to solve this concern, first transformations were carried out on the predictor variables and then when no more improvement could be made on the predictor vs residual plots, transformation on Y was conducted.

a. **Predictor transformation:** Transformation on a categorical variable is not useful since it is binary and thus will have no effect on the residuals. Thus, they were not considered for transformation. Most of the other predictors showed a good spread; however, sulfide, dissolved oxygen, precipitation, gas phase hydrogen sulfide, and flow rate did not show a good spread in the data, as shown in Fig 5.9.

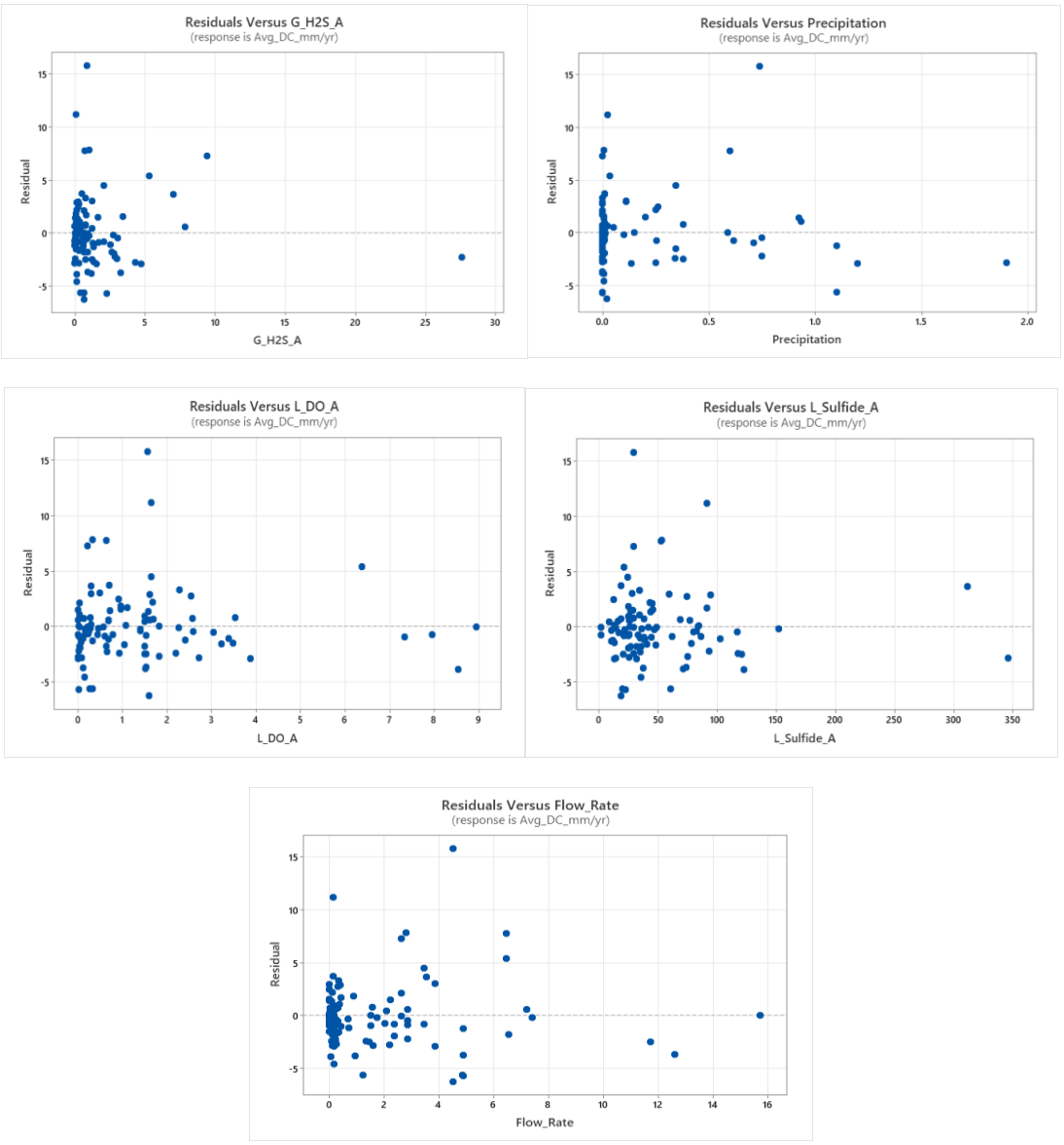


Fig 5.9 Residuals vs predictor variables

Various transformations were conducted on these predictors; however, the following transformations showed the best result.

i. Log (0.01+hydrogen sulfide) transformation on gas phase hydrogen sulfide.

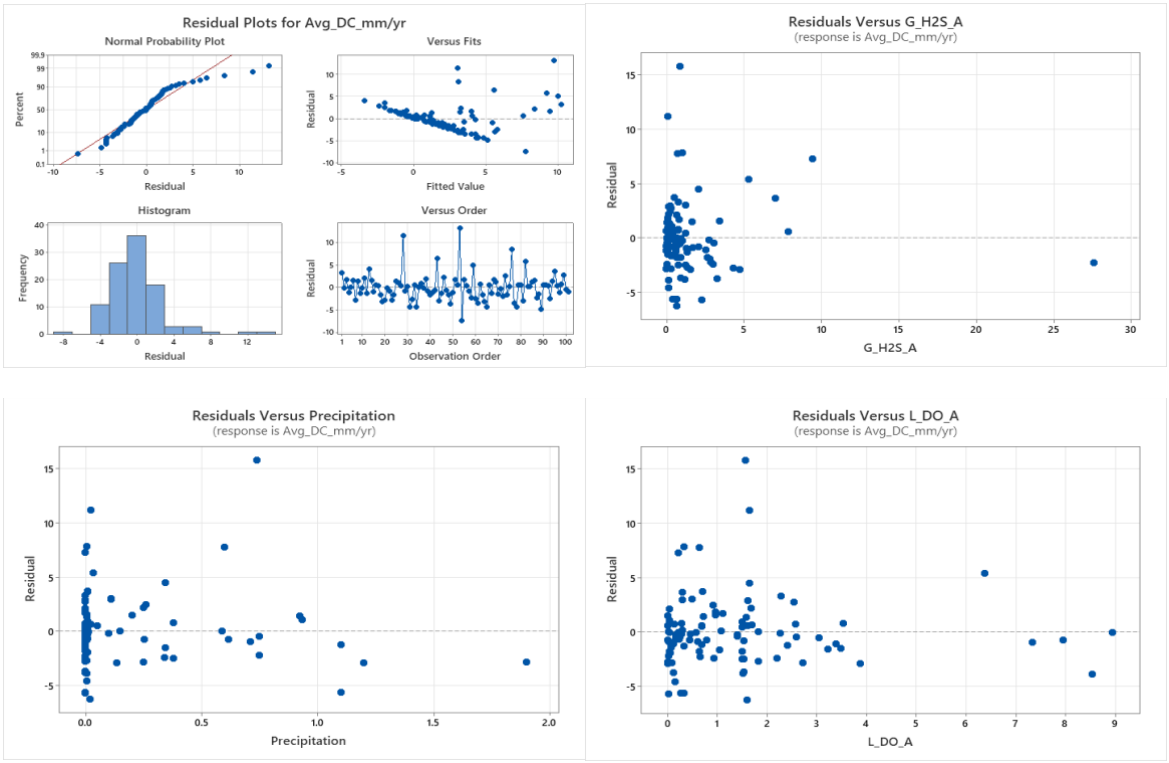
ii. Log (0.01+precipitation) transformation on precipitation.

iii. Log (0.1+DO) transformation on dissolved oxygen.

iv. Log (0.01+sulfide) transformation on sulfide.

v. Log (0.1+flowrate) transformation on flowrate.

The above transformations relatively improved the spread of the predictors, as clear from Fig 5.10. However, the spread of the fitted vs the residuals was still not good. Thus, a transformation on the response variable was conducted to fix the spread.



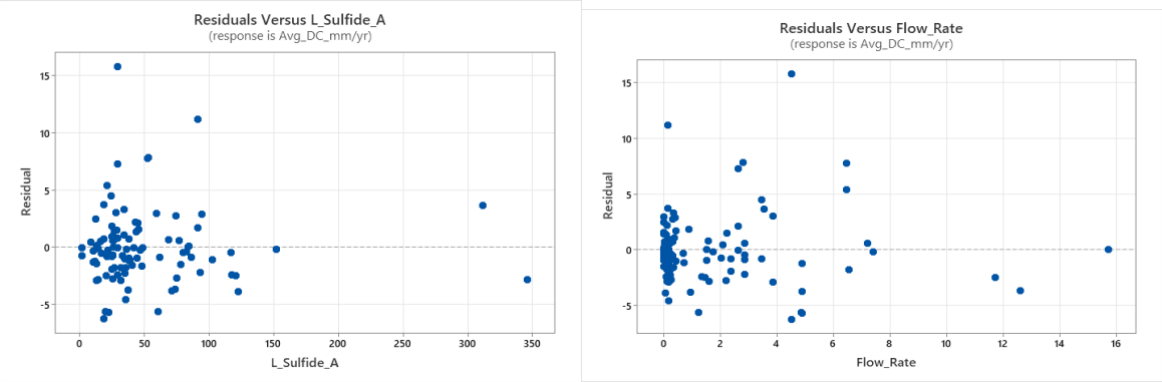


Fig 5.10 Residuals vs transformed predictor variables

Initially, weaker transformations were tried to see if they improved the model. The first attempt was to do a square root on the Y, and the model assumptions were again verified; however, the transformation did not fix the spread on the fitted versus the residuals. Since the SQRT(Y) transformation did not help with the spread, another transformation with logarithm(base10) was conducted. Since a few manholes had zero hydrogen sulfide concentrations, a constant 0.01 was added to the Y, and then a logarithm(base10) was calculated. The transformation improved the spread of the fitted versus the residuals, and the normality as well, as shown in the below Fig 5.11, and was considered for further analysis.

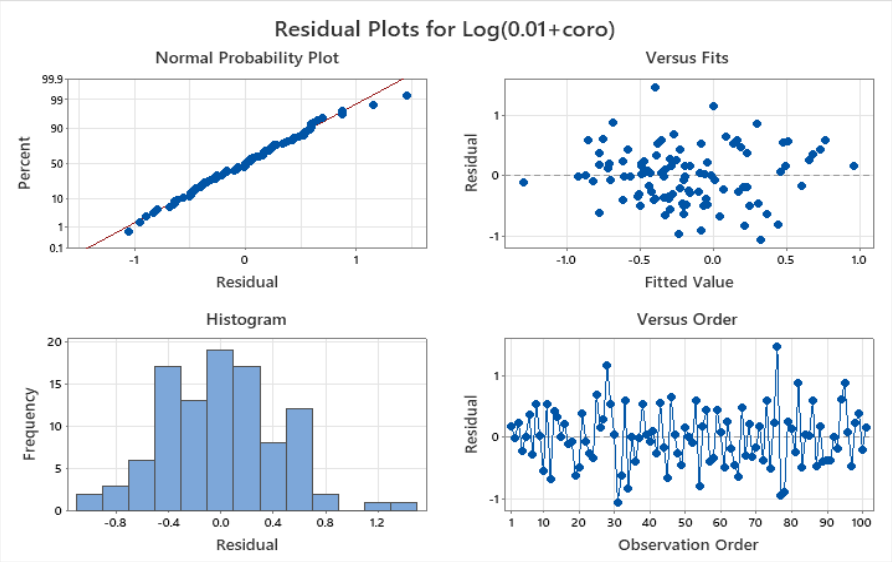


Fig 5.11 Residuals plots for $LOG_{10}(0.01+Y)$ transformation

The coefficients of predictor variables and ANOVA table for the transformed model are provided in Table 5.25, and Table 5.26 respectively. Here also, we can see that 29 predictors were insignificant, and 15 predictors had a high VIF>5.

Table 5.25 Estimates of coefficients for the predictor variables – transformed corrosion rate model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-1.83	2.02	-0.9	0.37	
Velocity	-0.047	0.198	-0.24	0.812	16.88
Depth_of_flow	-0.952	0.64	-1.49	0.141	21.22
MH_Depth	-0.0172	0.0135	-1.27	0.208	2.06
W_Temp_A	0.00496	0.00662	0.75	0.456	4.29
G_Temp_A	-0.0367	0.0248	-1.48	0.143	17.67
L_Temp_A	0.0634	0.0313	2.02	0.047	15.92
L_PH_A	0.042	0.117	0.36	0.719	2.16
High_Drop	0.727	0.718	1.01	0.315	18.36
Medium_Drop	0.893	0.64	1.39	0.168	32.87

No_Drop	0.803	0.656	1.22	0.225	52.17
Small_toLarge_PS	-0.633	0.7	-0.9	0.369	104.65
Large_to_Small_PS	-1.178	0.807	-1.46	0.149	13.71
Uniform	-0.581	0.72	-0.81	0.422	62.74
Hydraulic_Jump	0.318	0.259	1.23	0.224	1.96
Right	-0.204	0.634	-0.32	0.749	16.61
Obtuse	0.26	0.635	0.41	0.684	20.1
Multiple_inlets	-0.335	0.709	-0.47	0.639	39.87
Straight	0.353	0.631	0.56	0.577	23.52
Adj_High_H2S_US	0.128	0.272	0.47	0.64	1.56
Adj_High_H2S_DS	0.095	0.309	0.31	0.759	1.74
Log(0.01+Precip)	0.1312	0.0946	1.39	0.17	1.51
Supercritical	0.022	0.288	0.08	0.939	2.83
PS_Max_Negative	-0.0182	0.0125	-1.45	0.15	2.77
Age_H2S	-0.01161	0.00394	-2.95	0.004	1.72
L_BOD_A	0.000448	0.000348	1.29	0.202	1.38
Drop_max	-0.0381	0.0377	-1.01	0.315	3.27
L_Sulfate_A	-0.00042	0.00234	-0.18	0.859	1.77
LOG(0.1+DO)	0.455	0.139	3.27	0.002	1.68
Log(0.1+Flowrate)	0.891	0.677	1.31	0.193	57.48
LOG(0.1+H2S)	-0.029	0.147	-0.2	0.845	1.89
LOG(0.01+Sulfide)	0.027	0.178	0.15	0.879	1.32

Table 5.26 ANOVA for transformed corrosion rate model

Source	DF	Seq SS	Contribution	Adj SS	Adj MS	F-Value	P-Value
Regression	31	17.9464	45.66%	17.9464	0.57892	1.87	0.016
Velocity	1	0.7456	1.90%	0.0176	0.01759	0.06	0.812
Depth_of_flow	1	0.1791	0.46%	0.6858	0.68583	2.22	0.141
MH_Depth	1	0.2875	0.73%	0.4994	0.49937	1.61	0.208
W_Temp_A	1	0.4016	1.02%	0.1741	0.1741	0.56	0.456
G_Temp_A	1	0.418	1.06%	0.6782	0.67825	2.19	0.143
L_Temp_A	1	0.4939	1.26%	1.2652	1.26522	4.09	0.047
L_PH_A	1	0.041	0.10%	0.0404	0.04044	0.13	0.719
High_Drop	1	0.0915	0.23%	0.3169	0.3169	1.02	0.315
Medium_Drop	1	0.3708	0.94%	0.6018	0.60182	1.94	0.168
No_Drop	1	0.1853	0.47%	0.463	0.46298	1.5	0.225
Small_toLarge_PS	1	0.0158	0.04%	0.2532	0.25323	0.82	0.369
Large_to_Small_PS	1	1.0693	2.72%	0.6588	0.65877	2.13	0.149
Uniform	1	0.6801	1.73%	0.2019	0.20188	0.65	0.422
Hydraulic_Jump	1	0.1807	0.46%	0.4667	0.46674	1.51	0.224
Right	1	1.8522	4.71%	0.0319	0.03189	0.1	0.749
Obtuse	1	0.1507	0.38%	0.0519	0.05186	0.17	0.684
Multiple_inlets	1	0.4329	1.10%	0.0689	0.06888	0.22	0.639
Straight	1	0.2116	0.54%	0.0972	0.09717	0.31	0.577
Adj_High_H2S_US	1	0.389	0.99%	0.0681	0.06812	0.22	0.64
Adj_High_H2S_DS	1	0.0001	0.00%	0.0294	0.02938	0.09	0.759
Log(0.01+Precip)	1	0.7061	1.80%	0.5959	0.59591	1.93	0.17
Supercritical	1	0.3832	0.97%	0.0018	0.0018	0.01	0.939
PS_Max_Negative	1	0.4662	1.19%	0.6551	0.6551	2.12	0.15
Age_H2S	1	4.1562	10.58%	2.6865	2.68648	8.68	0.004

L_BOD_A	1	0.4552	1.16%	0.5145	0.51453	1.66	0.202
Drop_max	1	0.1573	0.40%	0.3165	0.31653	1.02	0.315
L_Sulfate_A	1	0.002	0.01%	0.0099	0.00985	0.03	0.859
LOG(0.1+DO)	1	2.857	7.27%	3.3147	3.31473	10.71	0.002
Log(0.1+Flowrate)	1	0.5472	1.39%	0.5351	0.53509	1.73	0.193
LOG(0.1+H2S)	1	0.0121	0.03%	0.0119	0.01192	0.04	0.845
LOG(0.01+Sulfide)	1	0.0072	0.02%	0.0072	0.00722	0.02	0.879
Error	69	21.3546	54.34%	21.3546	0.30949		
Total	100	39.301	100.00%				

5.3.2.1 TRANSFORMED MODEL – ASSUMPTIONS CHECK

Before the transformed model can be considered for further analysis, the model must satisfy the model assumptions.

a. **Residuals are normally distributed:** From Fig 5.11, we can predict that the model is normal now. Also, the normality test was done to validate normality assumption.

Normality Test: The normality test is conducted to see if the residuals are normally distributed or not.

Hypothesis: H_0: Normality is OK

H_1: Normality is violated

Test: Sample correlation p-value < α, then we reject H_0.

For the test, we assume confidence level $\alpha = 0.1$. From the software, we get the p-value = 0.677.

p-value > 0.1

Thus, we fail to reject H_0. So, the normality is satisfied.

b. **Residuals have a constant variance:** From the fitted and residuals plot, we can infer that the model has constant variance. However, to confirm our assumption we ran the modified Levene Test, and results for same are below.

Modified Levene Test: For modified Levene Test, we conduct a two-sample t-test on the absolute deviations of the residuals.

i. We first divide our dataset into two sets of almost equal size. The median of the dataset is selected by using fitted values; it was found to be -0.224159, so we divided the dataset with $\hat{y}$ >-0.224159 and $\hat{y}$ <=-0.224159

ii. Next, we perform the F-Test to check if the variances are equal or not.

Hypothesis: H_0: Variances are equal.

H_1: Variances are not equal.

Test: If p< α then we reject H_0. For the test, we assume confidence level α = 0.1. Results of the test are shown in Table 5.27.1.

Table 5.27.1 F-Test for equality of variances, transformed corrosion rate model

Equality of Variances				
Method	**Num DF**	**Denom DF**	**p-value**	**α**
F Test	49	50	0.346	0.1

So, our p (0.346)> α(0.1), thus we fail to reject H_0 and the variances are equal.

iii. After that we calculate the absolute deviations of the residuals around their group medians.

d1 = |e1 – median e1| and d2 = |e2 – median e2|

iv. Next, we perform the T-Test to see if the model has constant variance or not.

Hypothesis: H_0: Error Variances are equal (Means of d1, d2 populations are equal)

H_1: Error Variances are not equal (Means are not equal)

Test: If p< α then we reject H_0. For the test, we assume confidence level α = 0.1. Results of rhe test are shown in Table 5.27.2.

Table 5.27.2 T-Test for error variances, transformed corrosion model

Variances	DF	t value	p-value
Equal	99	1.477	**0.143**
Unequal	99	1.477	0.143

Since the F-Test showed that variances are equal, we select the p-value from Equal T-Test. So, our p (0.143)> α (0.1), thus we fail to reject H_0 and the error variances are equal. This is in accordance with the e vs ŷ plot. Thus, our model satisfies the constant variance assumption.

c. Residuals satisfy the MLR form: The predictor vs. residual plot are better for the transformed model in comparison to the preliminary model, as discussed above. There was no curvature visible on any of the plots; thus, the residual satisfies the MLR form.

d. Residuals are uncorrelated: From the Fig 5.11 time plot, we can confirm that residuals are uncorrelated for the transformed model.

e. Check for outliers: We need to verify our model for any kind of outliers, and the influence they have on the ŷ values.

i. X-Outlier: To check for X-outlier, leverage values h_{ii} were found, and if the h_{ii} is large, then observation *i* is outlying. The usually used guideline is if $h_{ii} > 2p/n$, then the observation *i* is outlying. Here, 2p/n = 0.63366

The value h_{ii} is the diagonal elements in the matrix H. Matrix $H = X(X^TX)^{-1}X^T$. The values of h_{ii} which are greater than the cutoff value 0.63366 are for the observation 5,7,35,58, thus showing them as X-outliers. Table 5.28 shows the details on X-outliers.

Table 5.28 X- Outliers, transformed corrosion rate model

Obs	HI	Cook's D	DFITS
5	1.000	*	*
7	0.732	0.070	-1.545
35	1.000	*	*
58	0.660	0.060	-1.438

ii. Y-Outliers: To check for Y-outliers, we use the studentized deleted residuals method. We calculate the t_i value by deleting each *i* observation in the dataset, where $t_i = d_i/(\text{sq.rt}(MSE_i\,(1-h_{ii}))$.

The usual guideline is if $|t_i| > t(1- \alpha/2n; n-p-1)$, then a particular observation is a y-outlier.

Assuming $\alpha = 0.1$, then our cutoff value is:

$t(1- \alpha/2n; n-p-1) = 3.4666$

The value of t_i for each observation of $_i$ is provided in the inserted file. We see that none of the t_i values are higher than the cutoff value, and for 2 of the observations the values could not be calculated as the leverage was 1.

iii. Outliers Influence on ŷ: Dataset does contain x-outlier thus we need to check the influence the x-outlier has on ŷ, individual LSEs, and combined LSEs. The effect of outliers is verified through 3 methods – Cook's Distance, DFFITS, and DFBETAS.

- **Cook's Distance:** It verifies the combined impact of observation *i* on all the LSEs. The formula for the Cook's Distance D_i is:

  $D_i = (b-b_{(i)})^T X^T X(b-b_{(i)})/p(MSE) = [e_i^2/p(MSE)]*[h_{ii}/((1-h_{ii})^2)]$

  The values of D_i for various *i* observations are provided in inserted file. The guideline is if $D_i > F(0.50; p,n-p)$, then the outlier does have influence.

  Here $F(0.50; p, n-p) = 0.99$

On comparing the D_i values with the cut off we see that none of the observations exceed the cutoff values; thus, we can say that none of the outliers have influence on the combined LSES.

- **DFFITS:** It checks the outliers influence on the fitted values by computing $\hat{y}_{i(i)}$ for every *i* observation being omitted. The formula to calculate DFFITS is:

  $(DFFITS)_i = (\hat{y}_i - \hat{y}_{i(i)})/ sq.rt(MSE_i * h_{ii}) = ti*(sq.rt(h_{ii}/(1- h_{ii})))$

  The values of DFFITS for various *i* observations is provided in the inserted file. The guideline is if |DFFITS| > 2*sq.rt(p/n), then the outlier does have influence on fitted values.

  Here, 2*sq.rt(p/n) = 1.1257

  On comparing the DFFITS values with the cutoff we see that none of the X-outliers exceed the cutoff, so outliers have no influence on the fitted values.

- **DFBETAS:** It checks the outliers influence on individual LSEs by computing $b_{k(i)}$ LSE for every *i* observation being omitted. The formula to calculate DFBETAS is:

  $(DFBETAS)_{k(i)} = b_k – b_{k(i)}/ sq.rt(MSE_i * C_{kk})$

  The values of DFBETAS for various *i* observations for each predictor variable and the intercept are provided in the inserted file. The guideline is if |DFBETAS| > 2/sq.rt(n), then the outlier does have influence.

  Here, 2/sq.rt(n) = 0.199

  The DFBETAS values with the cutoff for each of the predictor and the interceptor were compared, to check for influence on outliers on individual LSEs. The DFBETAS value of the X-outlier shown in the inserted file. The |DFBETAS| value showed influence for observation 7 on intercept, maximum drop, larger to smaller pipe size change, uniform flow, maximum change in pipe size, hydraulic jump, upstream to a high hydrogen sulfide manhole, and downstream to a high hydrogen sulfide manhole. The |DFBETAS| value showed influence for observation 58 on manhole depth, age, ambient temperature, high drop, maximum drop height, and smaller to larger pipe size change.

f. **Predictors not highly correlated:** Variance inflation factor is used for validating multicollinearity between the predictors. From the VIF value in Table 5.25 we can see that 15 predictors had a high VIF>5, thus effecting the model quality.

5.3.3 SELECTION OF SIGNIFICANT VARIABLES

From the transformed model, various predictors are not required, as they have very high VIF values. Thus, it becomes important to select only significant variables before we move further with our analysis. Stepwise regression and best subset selection method were conducted on the transformed model to gather a list of important predictor variables. To this new list of predictor variables, we will be incorporating the interaction terms.

Since the software cannot compute best subsets on predictor variables greater than 30, the gas phase temperature and liquid phase temperature variables were separated (as they have high VIF and correlation values), and 2 models were built (model A not containing the liquid phase temperature predictor variable, and model B not containing the gas phase temperature). The predictor output from both models were used to develop a significant predictor list.

a. **Stepwise method:** The stepwise method reduced the variables in the transformed model to 3 and generated the same output for both the models, as presented in Table 5.29. All the predictors in the stepwise model are significant and the VIF values are also less than 5, so they will be considered in the further analysis.

Table 5.29 Output of stepwise model, transformed corrosion rate model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-0.072	0.147	-0.49	0.623	
Velocity	0.1208	0.0515	2.34	0.021	1.05
Right	-0.347	0.163	-2.13	0.036	1.01
Age_H2S	-0.01212	0.00322	-3.76	0	1.05

b. **Best subset method:** The result for the best subset method is provided in the inserted file. The best models from the list were identified using the following metrics:

iii. Highest 2 adjusted R-square value

iv. Lowest 2 PRESS, Cp, AIC, and BIC values

Using the above criteria, 2 potential models; model A and model B were identified, and the list of significant predictors was developed. All the predictors from the 2 potential models were considered for the analysis irrespective of the VIF and the p-value to have a more comprehensive list of predictors in the model.

So, using the above method, we were able to shortlist 21 predictors out of the original 31 as important, and these predictors will now be used for model building after the addition of the interactions. The shortlisted variables are provided in Table 5.30.

Table 5.30 Selected predictor variables, transformed corrosion model

Variables		
Flowrate, MGD		
Velocity, ft/s		
Manhole age, years		
Depth of flow, ft		
Manhole Depth, ft		
Precipitation inches		
Ambient temperature, F		
Design Variables	Drop	>=0.2', <2' – Medium drop
		Maximum drop height, ft
	Pipe size change	Larger to smaller

		Maximum Pipe size change, inches
	Bends	No Bend 170° - 190° - Straight
		Right angle bend (80° - 110°)
		Obtuse angle bend (120° - 170°)
		Multiple inlets
	Adjacent to High H_2S manhole	Upstream to high H_2S manhole
Liquid-Phase Variables	Average Liquid Temp., °F	
	Average Liquid DO, mg/l	
	Average BOD, mg/L	
	Average Sulfate, mg/l	
Gas-Phase Variables	Average Temp., °F	

5.3.4 SELECTION OF INTERACTION TERMS

Two or more predictors can have an interaction effect on the response variable; thus, it becomes important to identify the interaction terms and consider them in the model. For the corrosion rate model, the total number of possible interactions terms was 465. It is not feasible to study these many interactions; thus, the knowledge from the literature was used to hand pick the interactions, and interactions which showed trends in the partial regression plots were selected.

From our understanding of the literature, we expected that if designs or factors which contribute to corrosion existed together in a manhole, then increased corrosion rates might be anticipated. Based on that assumption, the following

4 interactions were considered and partial plots for them were developed. Table 5.31 summarizes the selection reason for the interaction.

Table 5.31 Selected interactions, transformed corrosion model

Interaction Considered	Reason
Gas phase hydrogen sulfide*manhole age	According to literature, the age of concrete effects the diffusion rate of hydrogen sulfide into the concrete, which may enhance the corrosion rate. Thus, the interaction was considered.
Sulfide*gas phase hydrogen sulfide	Higher sulfide concentrations in the wastewater can result in higher hydrogen sulfide volatilization into the manhole's headspace, thus greater rates of corrosion. Thus, the interaction was considered.
Sulfate*gas phase hydrogen sulfide	Higher sulfate concentrations will allow faster sulfide production and thus increased hydrogen sulfide volatilization and may result in increased corrosion rate. Thus, the interaction was considered.
Gas phase temp*manhole age	For older manholes, higher temperature has shown to increase hydrogen sulfide diffusion, thus faster corrosion. Thus, the interaction was considered.

The interactions were standardized to reduce the concern of multicollinearity, which usually occurs with the addition of interactions. Then the partial plots for the 4 interactions were developed, and if any linear trend was seen, the interaction was incorporated into the transformed model. Out of the 4 considered, only 1 interaction showed reasonable trends, as can be seen in Fig. 5.12. The interaction that showed some trend is Sulfate*hydrogen sulfide. However, the remaining 2 interactions were still incorporated as they showed importance from a literature standpoint. Thus, a total of 4 interactions were added to the model.

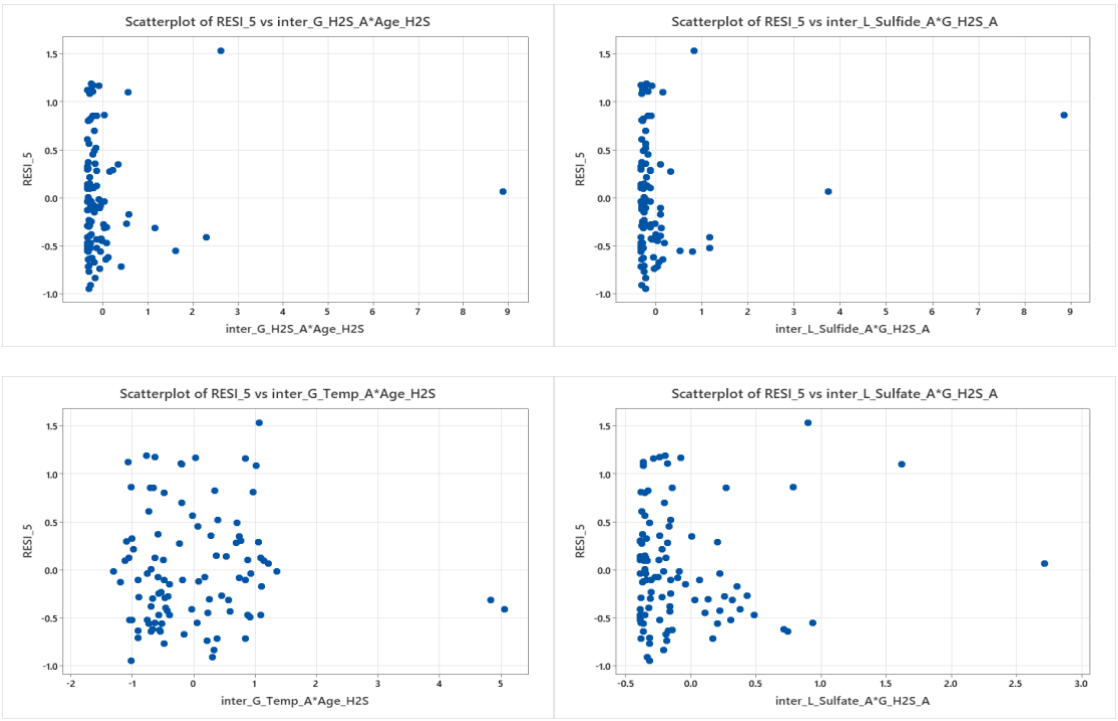


Fig 5.12 Partial regression plots for considered interactions

The new model with the 21 predictor variables and 4 interaction terms was considered for further analysis. The result of the new model is presented in the Table 5.32. Further analysis was conducted using the model with interactions.

Table 5.32 Coefficients for the model with interactions, transformed corrosion rate model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-0.82	2.47	-0.33	0.741	
Velocity	-0.051	0.146	-0.35	0.728	9.59
Depth_of_flow	-1.083	0.587	-1.84	0.069	18.74
MH_Depth	-0.0216	0.0132	-1.64	0.106	2.06
W_Temp_A	0.00529	0.00653	0.81	0.421	4.38
Medium_Drop	0.155	0.137	1.13	0.261	1.57

Large_to_Small_PS	-0.494	0.25	-1.98	0.052	1.37
Right	-0.23	0.611	-0.38	0.708	16.11
Obtuse	0.255	0.601	0.42	0.672	18.87
Multiple_inlets	-0.183	0.616	-0.3	0.767	31.48
Straight	0.309	0.598	0.52	0.607	22.17
Adj_High_H2S_US	0.292	0.274	1.07	0.29	1.65
Log(0.01+Precip)	0.1212	0.0846	1.43	0.156	1.27
PS_Max_Negative	-0.0163	0.0116	-1.4	0.165	2.51
Age_H2S	-0.0115	0.0341	-0.34	0.736	134.45
L_BOD_A	0.000403	0.000349	1.15	0.252	1.46
Drop_max	-0.0418	0.0271	-1.54	0.127	1.77
L_Sulfate_A	-0.00039	0.00226	-0.17	0.862	1.73
LOG(0.1+DO)	0.464	0.132	3.52	0.001	1.58
Log(0.1+Flowrate)	0.988	0.574	1.72	0.089	43.26
G_Temp_A	-0.0292	0.0292	-1	0.32	25.72
L_Temp_A	0.0524	0.0287	1.83	0.071	13.94
inter_G_Temp_A*Age_H2S	-0.007	0.646	-0.01	0.991	143.32
inter_G_H2S_A*Age_H2S	0.07	0.12	0.59	0.559	4.95
inter_L_Sulfide_A*G_H2S_A	-0.0092	0.0893	-0.1	0.918	2.75
inter_L_Sulfate_A*G_H2S_A	-0.155	0.275	-0.56	0.576	5.31

5.3.5 MODEL SEARCH

The model search was conducted using the stepwise regression and best subset methods.

a. **Stepwise method:** The model was run in the stepwise method, and it generated a model containing 3 significant variables. The results of the model are presented in the Table 5.33. All the predictors in the stepwise model are

significant and the VIF values are also less than 5. Thus, this model will be considered as a candidate for final model selection.

Table 5.33 Output of stepwise model – model search, transformed corrosion rate model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-0.072	0.147	-0.49	0.623	
Velocity	0.1208	0.0515	2.34	0.021	1.05
Right	-0.347	0.163	-2.13	0.036	1.01
Age_H2S	-0.01212	0.00322	-3.76	0	1.05

b. **Best subset method:** The results for the best subset method are provided in the inserted file. The software provides different models containing different combinations of predictors and the goodness-of-fit metrics for each model. These goodness-of-fit metrics are then evaluated for the candidate model selection. The candidate model must have the following metrics:

i. Adjusted R-square should be the highest.

ii. PRESS, AIC, and BIC should be the lowest.

iii. Cp should be the lowest or close to the number of predictors in the model.

Each model was evaluated on the above metrics and 5 candidate models were chosen. Results for candidate model selection are presented in Tables 5.34.

Table 5.34 Goodness-of-fit metrics for candidate models for transformed corrosion rate model, selected using best subset method*

Index	Vars	R-Sq	R-Sq (adj)	PRESS	R-Sq (pred)	Mallows Cp	AICc	BIC	Comments
M1	3	16.9	14.3	35.4	9.8	17.5	183.226	195.67	Selected for further discussion as similar to Stepwise regression
M2	9	35.1	28.7	30.6	22.1	5.3	172.622	198.422	Lowest CP, lowest AIC, okay PRESS - accepted
M3	11	37.9	30.3	30.5	22.5	5.5	173.293	203.105	Lowest PRESS, Okay CP - accepted
M4	16	42.6	31.6	32.6	16.9	9.4	179.612	218.343	Highest Adj Rsq

*Top in each metric category shown in yellow; highlighted in blue is same as stepwise model

Regression was done to develop the 4 selected models and they were compared with each other for VIF and P-value to narrow the search to 2 to 3 strong models. It was found that only model M1 and M2 met the requirement for low VIF and all significant predictor variables. The results for the 4 candidate model are presented below in Table 5.35.

Table 5.35 Candidate models for transformed corrosion model, selected using best subset method

M1 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	-0.072	0.147	-0.49	0.623	
Right	-0.347	0.163	-2.13	0.036	1.01
Velocity	0.1208	0.0515	2.34	0.021	1.05
Age_H2S	-0.01212	0.00322	-3.76	0	1.05
M2 model					

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-2.683	0.654	-4.1	0	
Large_to_Small_PS	-0.582	0.213	-2.73	0.008	1.06
L_BOD_A	0.000521	0.000291	1.79	0.076	1.06
Drop_max	-0.052	0.0226	-2.3	0.023	1.3
LOG(0.1+DO)	0.362	0.115	3.14	0.002	1.28
L_Temp_A	0.03105	0.00843	3.68	0	1.27
Multiple_inlets	-0.23	0.123	-1.86	0.066	1.33
Right	-0.53	0.162	-3.28	0.001	1.19
inter_G_Temp_A*Age_H2S	-0.2169	0.0584	-3.71	0	1.24
Velocity	0.1611	0.0481	3.35	0.001	1.1
M3 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**
Constant	-1.539	0.676	-2.28	0.025	
Depth_of_flow	-0.971	0.331	-2.94	0.004	6.4
MH_Depth	-0.0174	0.0114	-1.53	0.128	1.65
Large_to_Small_PS	-0.555	0.211	-2.64	0.01	1.05
L_BOD_A	0.000512	0.000289	1.77	0.08	1.08
Drop_max	-0.047	0.023	-2.05	0.043	1.37
LOG(0.1+DO)	0.458	0.119	3.84	0	1.4
Log(0.1+Flowrate)	0.868	0.231	3.76	0	7.53
L_Temp_A	0.03276	0.00851	3.85	0	1.32
Multiple_inlets	-0.288	0.126	-2.29	0.024	1.41
Right	-0.504	0.16	-3.15	0.002	1.19
inter_G_Temp_A*Age_H2S	-0.2245	0.0587	-3.82	0	1.28
M4 model					
Term	**Coef**	**SE Coef**	**T-Value**	**P-Value**	**VIF**

Constant	-1.421	0.808	-1.76	0.082	
Depth_of_flow	-0.879	0.341	-2.58	0.012	6.93
MH_Depth	-0.021	0.0115	-1.84	0.07	1.71
Medium_Drop	0.155	0.117	1.32	0.189	1.26
Large_to_Small_PS	-0.509	0.226	-2.26	0.026	1.23
Obtuse	0.428	0.159	2.69	0.009	1.45
Straight	0.5	0.146	3.42	0.001	1.45
Adj_High_H2S_US	0.262	0.217	1.21	0.23	1.14
Log(0.01+Precip)	0.1116	0.0769	1.45	0.151	1.15
PS_Max_Negative	-0.0149	0.00892	-1.67	0.099	1.63
Age_H2S	-0.01195	0.00323	-3.7	0	1.33
L_BOD_A	0.000385	0.000297	1.3	0.199	1.16
Drop_max	-0.0447	0.0229	-1.95	0.055	1.39
LOG(0.1+DO)	0.453	0.12	3.77	0	1.44
Log(0.1+Flowrate)	0.762	0.24	3.18	0.002	8.28
G_Temp_A	-0.0193	0.0177	-1.09	0.279	10.41
L_Temp_A	0.0496	0.0239	2.08	0.041	10.63

5.3.6 MODEL SELECTION

From the above discussion, we select M1 and M2 as the 2 strong candidate models for further analysis. In this section we will compare the M1 and M2 models based on the model assumptions and the prediction using the test set. The model which performs better in all or most categories will be selected as the final model.

5.3.6.1 M1 MODEL

The results for the M1 model are shown in the Table 5.36. From Table 5.36, we can infer that all the predictor variables are significant, as the P-value is less than 0.1 and multicollinearity is not a concern, as VIF is less than 5. Also, the model has a R^2 of 42.5% and Adjusted R^2 of 31.6%.

Table 5.36 Coefficients, summary, and ANOVA for M1 model

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-0.072	0.147	-0.49	0.623	
Velocity	0.1208	0.0515	2.34	0.021	1.05
Right	-0.347	0.163	-2.13	0.036	1.01
Age_H2S	-0.01212	0.00322	-3.76	0	1.05

Model Summary					
R-sq	**R-sq(adj)**	**PRESS**	**R-sq(pred)**	**AICc**	**BIC**
16.90%	14.33%	35.44	9.82%	183.23	195.67

Analysis of Variance							
Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	**F-Value**	**P-Value**
Regression	3	6.6428	16.90%	6.643	2.2143	6.58	0
Velocity	1	0.7456	1.90%	1.849	1.849	5.49	0.021
Right	1	1.1293	2.87%	1.521	1.5209	4.52	0.036
Age_H2S	1	4.7679	12.13%	4.768	4.7679	14.16	0
Error	97	32.6582	83.10%	32.658	0.3367		
Total	100	39.301	100.00%				

5.3.6.1.1 M1 MODEL ASSUMPTIONS

The model must satisfy the model assumptions to be considered. Fig 5.13 shows the residual plot for the M1 model.

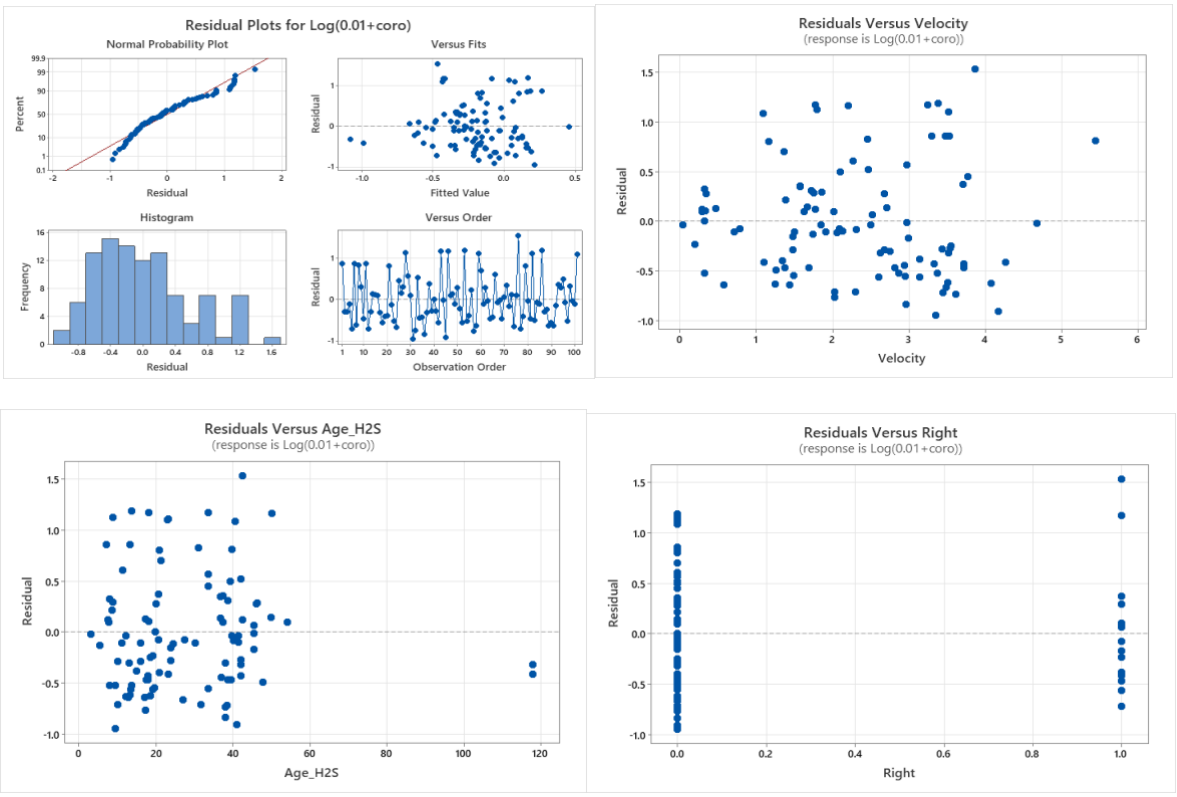


Fig 5.13 Normality plot, Residual plot and time plot

a. Residuals are normally distributed: From Fig 5.13, we can see that the model is normal. Also, a normality test was done to validate normality assumption.

Normality Test: The normality test is conducted to see if the residuals are normally distributed or not.

Hypothesis: H_0: Normality is OK.

H_1: Normality is violated.

Test: Sample p-value < α, then we reject H_0.

For the test, we assume confidence level α = 0.1. From the software, we obtain the p-value = 0.0058.

p-value < α

Thus, p-value < α, so we reject H_0. So, the normality is not satisfied.

b. Residuals have a constant variance: From the fitted and residuals plot, we can infer that the model has constant variance. However, to confirm our assumption, we conducted the modified Levene Test, and results for same are below.

Modified Levene Test: For the modified Levene test, we conducta two-sample t-test on the absolute deviations of the residuals.

i. We first divide our dataset into two sets of almost equal size. The median of the dataset is selected by using fitted values; it was found to be -0.22415, so we divided the dataset with $\hat{y}$ >-0.22415 and $\hat{y}$ <=-0.22415.

ii. Next, we perform the F-Test to check if the variances are equal or not.

Hypothesis: H_0: Variances are equal.

H_1: Variances are not equal.

Test: If p< α then we reject H_0. For the test, we assume confidence level α = 0.1. Results of test shown in Table 5.37.1.

Table 5.37.1 F-Test for equality of variances, M1 corrosion rate model

Equality of Variances				
Method	**Num DF**	**Denom DF**	**p-value**	**α**
F Test	49	50	0.0799	0.1

So, our p (0.0799)< α(0.1), thus we reject H_0 and the variances are not equal.

iii. After that we calculate the absolute deviations of the residuals around their group medians.

d1 = |e1 – median e1| and d2 = |e2 – median e2|

iv. Next, we perform the T-Test to see if the model has constant variance or not.

Hypothesis: H_0: Error Variances are equal (Means of d1, d2 populations are equal)

H_1: Error Variances are not equal (Means are not equal)

Test: If p< α then we reject H_0. For the test, we assume confidence level α = 0.1. Results of the test are shown in Table 5.37.2.

Table 5.37.2 T-Test for error variances, M1 corrosion rate model

Variances	DF	t value	p-value
Equal	99	1.887	0.0621
Unequal	99	1.887	**0.0617**

Since the F-Test showed unequal variances, we select the p-value from unequal T-Test. So, our p (0.0617)< α(0.1); thus, we reject H_0 and the error variances are not equal. This contrasts with e vs $\hat{y}$ plot. Thus, our model fails the constant variance assumption.

c. Residuals satisfy the MLR form: The residuals vs . plots had no curvature and showed a good spread, thus satisfying the MLR form. However, the residuals on the precipitation and maximum drop height predictors did not have a good spread, as most of the data points are zero, as seen in Fig 5.13.

d. Residuals are uncorrelated: From the Fig 5.13 time plot, we can confirm that residuals are uncorrelated for the M1 model.

e. Check for outliers: We need to verify our model for any kind of outliers, and the influence they have on the $\hat{y}$ values.

i. X-Outliers: To check for X-outliers, leverage values h_{ii} were found. I the h_{ii} is large, then observation *i* is outlying. The guideline usually used is if $h_{ii} > 2p/n$, then the observation *i* is outlying.

Here, 2p/n = 0.079207

The value h_{ii} is the diagonal elements in the matrix H. Matrix $H = X(X^TX)^{-1}X^T$. The values of h_{ii} which are greater than the cutoff value 0.079207 are for 12 observations, thus showing them as X-outliers. Table 5.38 shows the details on X-outliers.

Table 5.38 X- Outlier for M1 model for corrosion rate

Obs	HI	Cook's D	DFITS	Intercept	Age_H2S	Right	Velocity
16	0.100	0.000	0.063	0.028	0.001	0.049	-0.035
17	0.247	0.030	-0.355	0.179	-0.339	0.006	-0.001
19	0.251	0.060	-0.469	0.276	-0.429	0.012	-0.064
38	0.081	0.010	0.198	-0.045	-0.033	0.159	0.082
44	0.089	0.000	-0.008	0.001	0.005	0.002	-0.007
51	0.105	0.010	-0.141	-0.064	-0.005	-0.107	0.084
64	0.083	0.020	-0.252	0.052	0.055	-0.199	-0.107
70	0.079	0.000	0.037	-0.007	0.015	0.032	-0.002
72	0.080	0.000	-0.088	0.026	-0.032	-0.076	-0.008
76	0.087	0.180	0.883	-0.388	0.210	0.710	0.319
78	0.080	0.010	-0.216	-0.063	-0.012	-0.185	0.088
79	0.086	0.050	0.453	-0.262	0.003	-0.082	0.410

ii. Y-Outliers: To check for Y-outliers, we use the studentized deleted residuals method. We calculate the t_i value by deleting each *i* observation in the dataset, where $t_i = d_i/(\text{sq.rt}(MSE_i\,(1-h_{ii}))$.

The usual guideline is if $|t_i| > t(1- \alpha/2n; n-p-1)$, then a particular observation is a y-outlier.

Assuming $\alpha = 0.1$, then our cutoff value is:

$t(1- \alpha/2n; n-p-1) = 3.4209$

The value of t_i for each observation of $_i$ is provided in the inserted file. We see that none of the t_i values are higher than the cutoff value. So, we can confidently say that there is no y-outlier in our dataset.

iii. Outliers Influence on ŷ: The dataset does contain x-outliers; thus, we need to check the influence x-outliers have on ŷ, individual LSEs, and combined LSEs. The effect of outliers is verified through 3 methods – Cook's Distance, DFFITS, and DFBETAS.

- **Cook's Distance:** This method verifies the combined impact of observations *i* on all the LSEs.

  The formula for the Cook's Distance D_i is:

  $D_i = (b-b_{(i)})^T X^T X(b-b_{(i)})/p(MSE) = [e_i^2/p(MSE)]*[h_{ii}/((1-h_{ii})^2)]$

  The values of D_i for various *i* observations are provided in the inserted file. The guideline is if $D_i >$ F(0.50; p,n-p), then the outlier does have influence.

  Here F(0.50; p, n-p) = 0.85

  On comparing the D_i values with the cut off, we see that none of the observations exceed the cutoff values; thus, we can say that none of the outliers has influence on the combined LSES.

- **DFFITS:** This method checks the outliers' influence on the fitted values by computing $\hat{y}_{i(i)}$ for every *i* observation being omitted. The formula to calculate DFFITS is:

  $(DFFITS)_i = (\hat{y}_i - \hat{y}_{i(i)})/ sq.rt(MSE_i * h_{ii}) = ti*(sq.rt(h_{ii}/(1- h_{ii})))$

  The values of DFFITS for various *i* observations are provided in the inserted file. The guideline is if |DFFITS| > 2*sq.rt(p/n), then the outlier does have influence on fitted values.

  Here, 2*sq.rt(p/n) = 0.39801

  On comparing the DFFITS values with the cutoff, we see that 2 of the X-outliers exceed the cutoff, so outliers have some influence on the fitted values.

- **DFBETAS:** This method checks the outliers influence on individual LSEs by computing $b_{k(i)}$ LSE for every *i* observation being omitted. The formula to calculate DFBETAS is:

  $(DFBETAS)_{k(i)} = b_k – b_{k(i)}/ sq.rt(MSE_i * C_{kk})$

  The values of DFBETAS for various *i* observations for each predictor variable and the intercept, is provided in the inserted file. The guideline is if |DFBETAS| > 2/sq.rt(n), then the outlier does have influence.

  Here, 2/sq.rt(n) = 0.199

The DFBETAS values with the cutoff for each of the predictor and the interceptor were compared, to check for influence on outlier on individual LSEs. The DFBETAS value of the X-outliers are shown in the inserted file. The |DFBETAS| value exceeded the cutoff for the following variables.

i. Intercept for observation 19,79.

j. Age of manhole for observation 19.

k. Velocity for observation 79.

The DBETAS values are shown in the Table 5.50.

f. Predictors not highly correlated: Variance inflation factor is used for identifying multicollinearity between the predictors. From the VIF value in Table 5.36, we can see that none of the predictors has a VIF>5; thus, all the predictors are significant in the model.

5.3.6.1.2 M1 MODEL VALIDATION

Model validation was done by using the test dataset. 20% of the collected data was reserved for validation. Validation helps to estimate the model prediction capability when dataset different from that of training set is used. This helps to identify how close the model predicts to the original values of the test set. In this step, we predict the corrosion rate in a manhole using the M1 model and compare it with the original values from the test set. Then we calculate the percentage of difference between the test set measured values and values calculated from the M1 model.

$$\%\ of\ difference = [Avg\left(\frac{(|Y - \hat{Y}|)}{Y^*} * 100\right)] = 50.63\%$$

Where:

Y^* = measured corrosion rate in test set

The percentage of difference between the predicted value and the test value was found to be 50.63%; this shows that modeled corrosion rate was an average of 50.63% above or below the measured field corrosion rate values. This

demonstrates that the model does a good job at identifying the designs and variables which may be responsible for corrosion.

5.3.6.2 M2 MODEL

The results for the M2 model are shown in Table 5.39. From Table 5.39, we can infer that all the predictor variables are significant, as P-values are less than 0.1. No multicollinearity is noticed, as VIF is less than 5. Also, the model has a R-square of 35.08% and Adjusted R-square of 28.66%.

Table 5.39 Coefficients, summary, and ANOVA for M2 model for corrosion rate

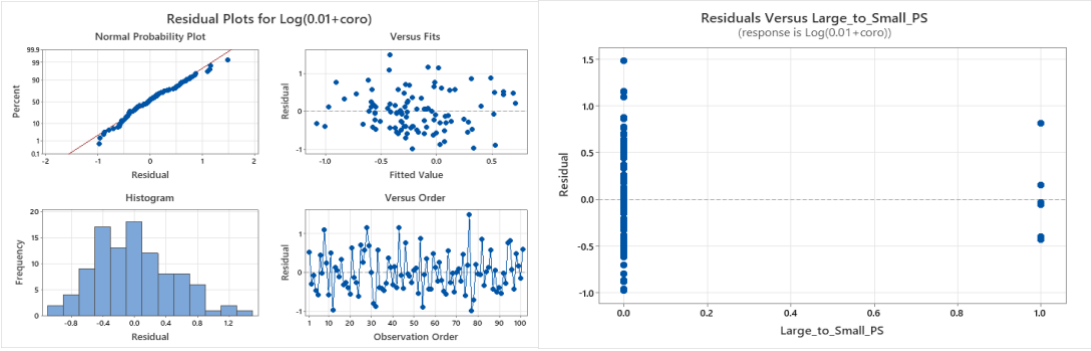
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-2.683	0.654	-4.1	0	
Velocity	0.1611	0.0481	3.35	0.001	1.1
Right	-0.53	0.162	-3.28	0.001	1.19
Large_to_Small_PS	-0.582	0.213	-2.73	0.008	1.06
L_BOD_A	0.000521	0.000291	1.79	0.076	1.06
Drop_max	-0.052	0.0226	-2.3	0.023	1.3
LOG(0.1+DO)	0.362	0.115	3.14	0.002	1.28
L_Temp_A	0.03105	0.00843	3.68	0	1.27
Multiple_inlets	-0.23	0.123	-1.86	0.066	1.33
inter_G_Temp_A*Age_H2S	-0.2169	0.0584	-3.71	0	1.24

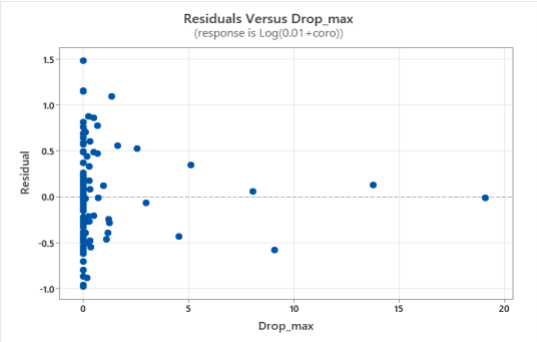
Model Summary					
R-sq	**R-sq(adj)**	**PRESS**	**R-sq(pred)**	**AICc**	**BIC**
35.08%	28.66%	30.6048	22.13%	172.62	198.42

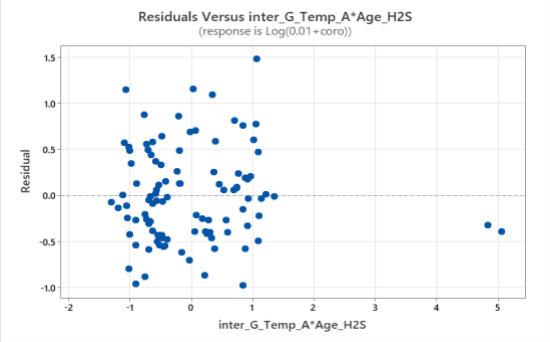
Analysis of Variance							
Source	**DF**	**Seq SS**	**Contribution**	**Adj SS**	**Adj MS**	**F-Value**	**P-Value**
Regression	9	13.7881	35.08%	13.7881	1.532	5.46	0
Velocity	1	0.7456	1.90%	3.1442	3.1442	11.21	0.001
Right	1	1.1293	2.87%	3.0197	3.0197	10.77	0.001
Large_to_Small_PS	1	0.7731	1.97%	2.0858	2.0858	7.44	0.008
L_BOD_A	1	0.4573	1.16%	0.9018	0.9018	3.22	0.076
Drop_max	1	0.2907	0.74%	1.4887	1.4887	5.31	0.023
LOG(0.1+DO)	1	2.8935	7.36%	2.7643	2.7643	9.86	0.002
L_Temp_A	1	2.0231	5.15%	3.7995	3.7995	13.55	0
Multiple_inlets	1	1.6095	4.10%	0.9732	0.9732	3.47	0.066
inter_G_Temp_A*Age_H2S	1	3.866	9.84%	3.866	3.866	13.79	0
Error	91	25.5129	64.92%	25.5129	0.2804		
Total	100	39.301	100.00%				

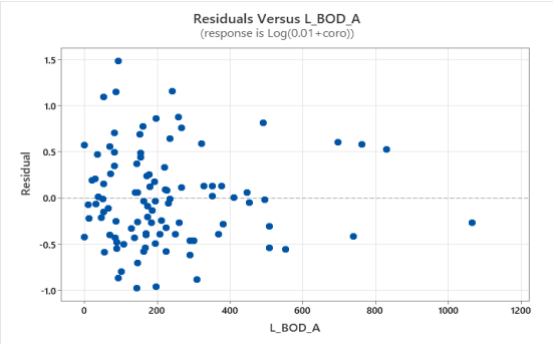
5.3.6.2.1 M2 MODEL ASSUMPTIONS

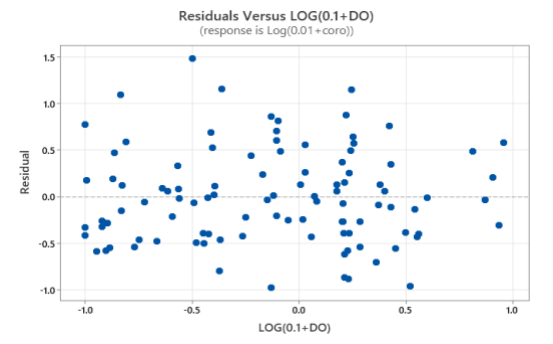
The model must satisfy the model assumptions to be considered. Fig 5.14 shows the residual plot for the M2 model.

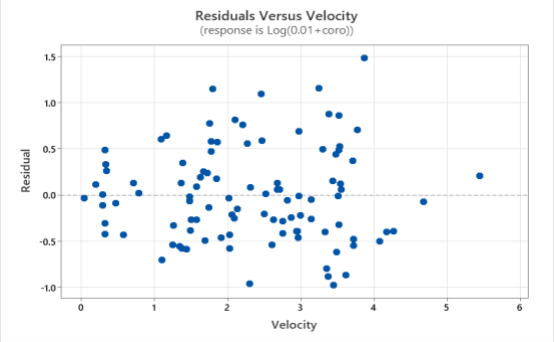


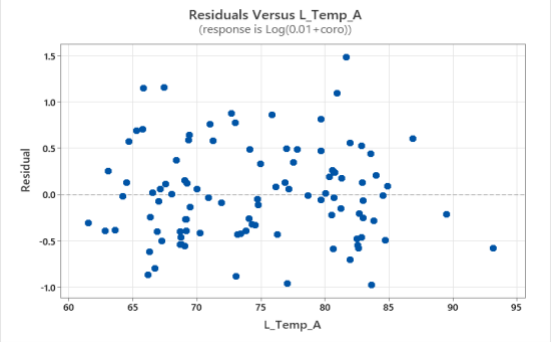

Residuals Versus Drop_max
(response is Log(0.01+coro))
Residual
Drop_max

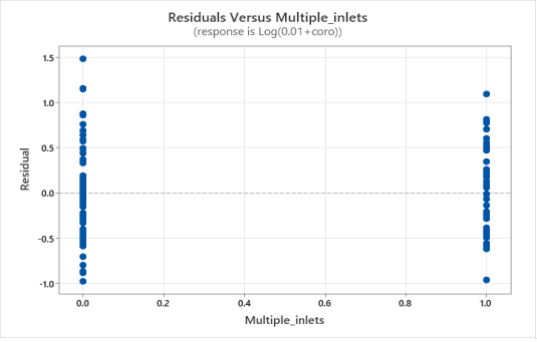

Residuals Versus inter_G_Temp_A*Age_H2S
(response is Log(0.01+coro))
Residual
inter_G_Temp_A*Age_H2S

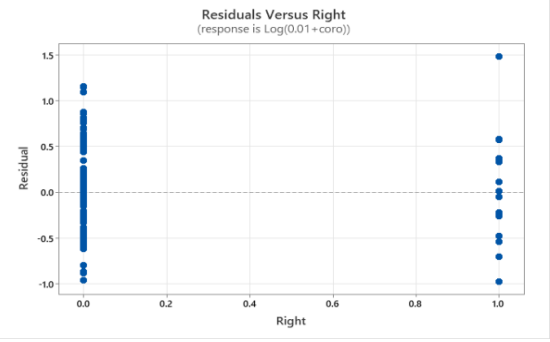

Residuals Versus L_BOD_A
(response is Log(0.01+coro))
Residual
L_BOD_A


Residuals Versus LOG(0.1+DO)
(response is Log(0.01+coro))
Residual
LOG(0.1+DO)


Residuals Versus Velocity
(response is Log(0.01+coro))
Residual
Velocity


Residuals Versus L_Temp_A
(response is Log(0.01+coro))
Residual
L_Temp_A


Residuals Versus Multiple_inlets
(response is Log(0.01+coro))
Residual
Multiple_inlets


Residuals Versus Right
(response is Log(0.01+coro))
Residual
Right

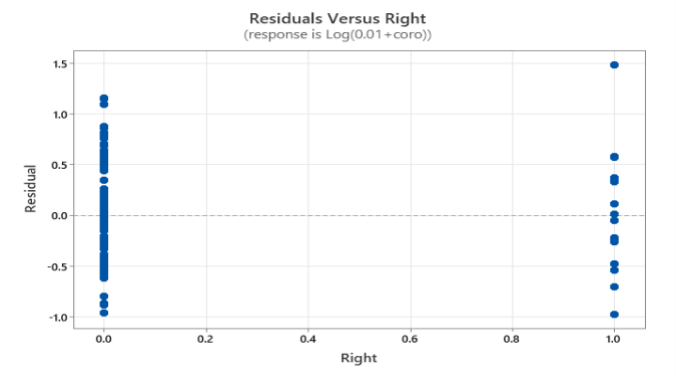


Fig 5.14 Normality plot, Residual plot and time plot

a. Residuals are normally distributed: From Fig 5.14, we can discern that the model is normal. Also, a normality test was conducted to confirm the normality assumption.

Normality Test: Normality test is conducted to see if the residuals are normally distributed or not.

Hypothesis: H_0: Normality is OK.

H_1: Normality is violated.

Test: Sample p-value < α, then we reject H_0.

For the test, we assume confidence level α = 0.1. From the software, we get the p-value = 0.214.

p-value > α

Thus, p-value > α, so we fail to reject H_0. So, the normality is satisfied.

b. Residuals have a constant variance: From the fitted and residuals plots, we can infer that the model has constant variance. However, to confirm our assumption, we conducted the modified Levene Test, and results for same are below.

Modified Levene Test: For modified Levene test, we conduct two-sample t-test on the absolute deviations of the residuals.

i. We first divide our dataset into two sets of almost equal size. The median of the dataset is selected by using fitted values. It was found to be 1.4429, so we divided the dataset with $\hat{y}$ >1.4429 and $\hat{y}$ <=1.4429. Next, we perform the F-Test to check if the variances are equal or not.

Hypothesis: H_0: Variances are equal.

H_1: Variances are not equal.

Test: If p< α then we reject H_0. For the test, we assume confidence level α = 0.1. Results of the test shown in Table 5.40.1.

Table 5.40.1 F-Test for equality of variances for the M2 model for corrosion rate

Equality of Variances				
Method	**Num DF**	**Denom DF**	**p-value**	**α**
F Test	49	50	1.4429	0.1

So, our p (1.4429)> α(0.1); thus, we reject H_0 and the variances are not equal.

ii. After that, we calculate the absolute deviations of the residuals around their group medians.

d1 = |e1 – median e1| and d2 = |e2 – median e2|

iii. Next, we perform the T-Test to see if the model has constant variance or not.

Hypothesis: H_0: Error Variances are equal (Means of d1, d2 populations are equal)

H_1: Error Variances are not equal (Means are not equal)

Test: If p< α then we reject H_0. For the test, we assume confidence level α = 0.1. Results of test shown in Table 5.40.2.

Table 5.40.2 T-Test for error variances for M2 model for corrosion rate

Variances	**DF**	**t value**	**p-value**
Equal	99	1.6585	**0.10037**
Unequal	98.217	1.6585	0.1006

Since the F-Test showed that variances are equal, we select the p-value from the equal T-Test. So, our p (0.10037)> α(0.1); thus, we reject H_0 and the error variances are not equal. This is in accordance with e vs ŷ plot. Thus, our model passes the constant variance assumption.

c. Residuals satisfy the MLR form: The residuals vs predictor plots had no curvature and showed a good spread, thus satisfying the MLR form, as seen in Fig. 5.14.

d. Residuals are uncorrelated: From the Fig 5.14 time plot, we can confirm that residuals are uncorrelated for the M1 model.

e. Check for outliers: We need to check our model for any kind of outliers, and the influence they have on the ŷ values.

i. X-Outliers: To check for X-outliers, leverage values h_{ii} were found. If the h_{ii} is large, then observation *i* is outlying. The guideline usually used is if $h_{ii} > 2p/n$, then the observation *i* is outlying.

Here, 2p/n = 0.198

The value h_{ii} is the diagonal elements in the matrix H. Matrix $H = X(X^TX)^{-1}X^T$. The values of h_{ii} which are greater than the cutoff value 0.198 are for 11 observation, thus showing them as X-outliers. Table 5.41 shows the details on X-outliers.

Table 5.41 X-outliers for M2 model for corrosion model

Obs	Fit	SE Fit	Resid	Std Resid	Del Resid	HI	Cook's D	DFITS
1	0.603	0.248	0.525	1.12	1.12	0.218568	0.04	0.59455
7	-0.413	0.375	-0.011	-0.03	-0.03	0.501307	0	-0.03027
13	-0.283	0.266	0.128	0.28	0.28	0.252977	0	0.16135
17	-1.074	0.282	-0.317	-0.71	-0.7	0.284232	0.02	-0.44401
19	-1.003	0.29	-0.39	-0.88	-0.88	0.299541	0.03	-0.57451
45	-0.571	0.247	-0.401	-0.86	-0.85	0.217388	0.02	-0.4501
55	-0.55	0.255	-0.048	-0.1	-0.1	0.231487	0	-0.05699
79	0.708	0.251	0.208	0.45	0.44	0.225349	0.01	0.23993
80	-0.552	0.275	-0.03	-0.07	-0.07	0.268962	0	-0.04034
86	0.168	0.239	0.582	1.23	1.24	0.203257	0.04	0.62428
93	0.286	0.256	-0.267	-0.58	-0.57	0.232894	0.01	-0.31652

ii. **Y-Outliers:** To check for Y-outlier, we use the studentized deleted residuals method. We calculate the t_i value by deleting each *i* observation in the dataset, where $t_i = d_i/(\text{sq.rt}(MSE_i\,(1-h_{ii}))$.

The usual guideline is if $|t_i| > t(1- \alpha/2n; n-p-1)$, then a particular observation is a y-outlier.

Assuming $\alpha = 0.1$, then our cutoff value is:

$t(1- \alpha/2n; n-p-1) = 3.4283$

The value of t_i for each observation of $_i$ is provided in the inserted file. We see that none of the t_i values are higher than the cutoff value. So, we can confidently say that there is no y-outlier in our dataset.

iii. **Outliers Influence on ŷ:** The dataset does contain x-outliers; thus, we need to check the influence the x-outlier has on ŷ, individual LSEs, and combined LSEs. The effect of outliers is verified through 3 methods – Cook's Distance, DFFITS, and DFBETAS.

- **Cook's Distance:** This method verifies the combined impact of observation *i* on all the LSEs. The formula for the Cook's Distance D_i is:

  $D_i = (b-b_{(i)})^T X^T X(b-b_{(i)})/p(MSE) = [e_i^2/p(MSE)]*[h_{ii}/((1-h_{ii})^2)]$

  The values of D_i for various *i* observations are provided in inserted file. The guideline is if $D_i >$ F(0.50; p,n-p), then the outlier does have influence.

  Here F(0.50; p, n-p) = 0.94

  On comparing the D_i values with the cut off, we see that none of the observations exceeds the cutoff values thus, we can say that none of the outliers has influence on the combined LSES.

- **DFFITS:** This method checks the outliers influence on the fitted values by computing $\hat{y}_{i(i)}$ for every *i* observation being omitted. The formula to calculate DFFITS is:

  $(DFFITS)_i = (\hat{y}_i - \hat{y}_{i(i)})/ sq.rt(MSE_i * h_{ii}) = ti*(sq.rt(h_{ii}/(1- h_{ii})))$

  The values of DFFITS for various *i* observations is provided in the inserted file. The guideline is that if |DFFITS| > 2*sq.rt(p/n), then the outlier does have influence on fitted values.

  Here, 2*sq.rt(p/n) = 0.6293

  On comparing the DFFITS values with the cutoff, we see that observation 93 exceeds the cutoff, so the X-outliers have some influence on the fitted values.

- **DFBETAS:** This method checks the outliers' influence on individual LSEs by computing $b_{k(i)}$ LSE for every *i* observation being omitted. The formula to calculate DFBETAS is:

  $(DFBETAS)_{k(i)} = b_k – b_{k(i)}/ sq.rt(MSE_i * C_{kk})$

  The values of DFBETAS for various *i* observations for each predictor variable and the intercept are provided in the inserted file. The guideline is if |DFBETAS| > 2/sq.rt(n), then the outlier does have influence.

Here, 2/sq.rt(n) = 0.199

The DFBETAS values with the cutoff for each of the predictor and the interceptor were compared, to check for influence of outliers on individual LSEs. The DFBETAS value of the X-outlier shown in the inserted file. The |DFBETAS| value for BOD of observation 93 exceeded the cutoff. The DBETAS values are shown in Table 5.42.

Table 5.42 DFBETAS for M2 corrosion model

Obs	Intercept	Velocity	L_Temp_A	DO	L_BOD_A	Drop_max	Large_to_Small_PS	Right	Multiple_inlets	inter_G_Temp_A *Age_H2S
1	-0.314	0.246	0.235	-0.085	0.475	-0.074	-0.045	-0.009	0.133	-0.255
7	-0.002	-0.001	0.002	0.000	0.002	-0.028	-0.001	0.000	0.005	-0.002
13	0.009	0.000	-0.011	-0.008	0.014	0.143	0.005	0.004	-0.011	-0.006
17	-0.103	0.004	0.089	0.021	0.006	-0.083	0.015	0.049	0.135	-0.398
19	-0.083	-0.078	0.103	-0.065	-0.083	-0.006	0.029	-0.012	-0.062	-0.500
45	-0.096	-0.115	0.121	-0.060	0.089	0.041	-0.340	-0.013	-0.149	-0.030
55	0.004	-0.008	0.000	0.006	-0.017	0.000	-0.043	-0.029	-0.004	0.009
79	-0.116	0.146	0.107	0.157	-0.021	-0.089	-0.054	-0.036	0.037	0.023
80	0.002	0.018	-0.007	-0.016	0.005	-0.001	-0.025	0.006	0.010	-0.014
86	-0.099	0.010	0.070	0.269	0.411	-0.099	-0.094	0.306	-0.004	-0.001
93	0.032	0.001	-0.020	-0.047	-0.287	0.027	0.036	0.033	0.067	-0.038

f. **Predictors not highly correlated:** Variance inflation factor is used for identifying multicollinearity between the predictors. From the VIF value in Table 5.39, we can see that none of the predictors has a VIF>5; thus, all the predictors are significant in the model.

5.3.6.2.2 M2 MODEL VALIDATION

Model validation was done by using the test dataset. 20% of the collected data was reserved for validation. Validation helps to estimate the model prediction capability when dataset different from that of training set is used. This helps to identify how close the model predicts to the original values of the test set. In this step, we predict the corrosion rate in a manhole using the M2 model and compare it with the original values from the test set. Then we calculate the percentage of difference between the test set measured values and values calculated from the M2 model.

$$\%\ of\ difference = [Avg\left(\frac{(|Y - \hat{Y}|)}{Y^*} * 100\right)] = 30.42\%$$

Where:

Y^* = measured corrosion rate in test set

The percentage of difference between the predicted value and the test value was found to be 30.42%; this shows that modeled corrosion rate was an average of 30.42% above or below the measured field corrosion rate values. This demonstrates that the model does a good job at identifying the designs and variables which may be responsible for corrosion.

5.3.6.3 M1 and M2 MODEL COMPARISON

This section compares the model M1 and M2 for model assumptions, outliers, and goodness of fit parameters to make the final model selection. Table 5.43 provides details on the comparison. From Table 5.43, we can clearly see that M2 model performed significantly better than M1 model, since it had a better R-square, Adjusted R-square, AIC, Cp, BIC, PRESS, and R-square$_{(pred)}$ values. In contrast, the M1 model completely failed the modified Levene's test and the normality test. Non – constant variance will result in incorrect prediction; a non-normal model is not fit

for multiple linear regression. Also, M2 model had a % of difference of only 30.42% in comparison to M1 which had % of error of 50.63%. Thus, the M2 model was selected as the final corrosion rate model.

Table 5.43 Model comparison

Index	M1	M2	Comments
Correlation between predictors	Highest between predictors 21%	Highest between predictors 35%	Both the models perform equally well, as the correlation between the predictors is less than 70%.
Correlation of predictors with Y	Highest 29%	Highest 28%	Both models show similar results, thus both seem acceptable.
Residual analysis	No clear curvature	No clear curvature	Both models perform well.
Constant variance - plot	Satisfied	Satisfied	Both models perform well.
Levene Test	Failed	Passed	M1 fails the modified Levene test, and an important assumption; thus M2 model is a better model.
VIF	Highest 1.05	Highest 1.33	Both models have VIF less than 5 so both are good.
P-value	Highest 0.056	Highest 0.076	Both models have all predictors as significant; thus, both are good.
Normality	Satisfied	Satisfied	Both models perform equally well
Normality Test	Failed	Satisfied	M1 fails the normality test, and an important assumption thus, M2 model is a better model.
Rsq	16.9%	35.1%	M2 performs better.
Adj R sq	14.33%	28.7%	M2 performs better.

Cp	17.5	5.3	M2 performs better, as the Cp value should be lesser for a good model.
PRESS (Train)	35.4	30.6	M2 performs better since the PRESS value should be lesser for a good model.
AIC	183.23	172.62	M2 performs better since the AIC value should be lesser for a good model.
BIC	195.67	198.42	M1 performs better since the BIC value should be lesser for a good model.
High h_{ii}	12	11	M1 model is sensitive to outliers; thus, M2 model performed better.
Y-outlier	None	None	Both models performed equally well.
Outlier influence	2	1	The M1 model is sensitive to outliers; thus, the M2 model performed better.
% of difference	50.63%	30.42%	Both failed to predict the test set.

5.3.7 MODEL INTERPRETATION

From the above analysis, we were able to develop a corrosion rate model. The model for corrosion in manhole is presented below:

$$Log_{10}(0.01 + Y) = -2.683 + 0.1611V - 0.53Bend_{right} - 0.582PS_{ls} + 0.000521BOD - 0.052Drop_{max} + 0.362\left(Log_{10}(0.1 + DO)\right) + 0.03105T_{liq} - 0.23Inlets_{multiple} - 0.2169(T_{gas} * Age) \quad \text{……Eqn. 5.4}$$

Where:

Y = Rate of corrosion, mm/yr

V = Velocity of wastewater, ft/s

$Bend_{right}$ = Presence of a right-angle bend

PS_{ls} = Change in pipe size diameter (inlet pipe larger than outlet pipe) (binary)

BOD = Wastewater BOD (mg/l)

$Drop_{max}$ = Maximum drop height (ft)

DO = Wastewater dissolved oxygen (mg/l)

T_{liq} = Wastewater temperature (F)

$Inlets_{multiple}$ = Presence of multiple inlets

T_{gas}*Age = Interaction between the manhole gas phase temperature and age (F*yrs)

The model shows that rate of corrosion increases with an increase in velocity, BOD, DO, and liquid temperature. As mentioned previously, increased velocity will allow for increased hydrogen volatilization into the manhole's headspace. Higher levels of BOD in the wastewater will result in higher hydrogen sulfide generation due to conversion of sulfate to sulfide as microorganisms consume chemically-bound oxygen, and thus faster rates of corrosion. Increased liquid-phase DO, although decreasing liquid-phase sulfide formation, may lead to increased oxidation of gas-phase hydrogen sulfide to sulfuric acid, and thus increased corrosion. Also, a rise in liquid temperature results in increased volatilization of the hydrogen sulfide gas into the manhole headspace, thus increasing the rate of corrosion.

The model also considers the effect of interaction between the gas phase temperature and age of the manhole. This is very interesting, as literature has suggested that with age the concrete properties change, and uptake rate of the hydrogen sulfide by the concrete has been found to change with temperature depending on the age of concrete due to increased diffusion of H_2S in air and increased chemical and biological sulfide oxidation rates[(43)]: new concrete showed a 17% increase in hydrogen sulfide uptake rate compared to 26% for old corroded concrete with increase in temperature[(43)]. Thus, the model identifies the interaction, although the sign would be expected to be positive, because increasing temperature and manhole age should increase corrosion rates.

The presence of right-angle bend or multiple inlets, the size of the inlet pipe being larger than the outlet pipe, and increased drop height decrease the rate of corrosion. Although this is surprising, it is in agreement with our Chapter 4 assumption that manholes which do not suffer from drastic changes in concrete environments may show higher rates of corrosion, since the microbial community have more a consistent environment to thrive and carry out corrosion. The presence of the drop may result in washing off the microbial community present on the concrete

surface, thus reducing the rate of corrosion. This is also an interesting result that showcases how manhole design plays an important role in determining corrosion rate, in contrast to sewer pipes, where such a scenario would be unlikely, because in sewer pipes the physical design is horizontal; thus, presence of a high drop is unlikely. In addition, since the manhole is a vertical structure and near to ground surface, the microbial community could vary depending on the temperature (temperature increases with depth of the manhole, as seen in Chapter 4), oxygen, and pH variability along the depth of the manhole, especially for really deep manholes, whereas in sewer pipe the conditions may not depend on these factors. Also, we assumed that lower sections of the manhole are more corroded for corrosion determination; this might imply that due to the heavy flow from the drop, those areas may not support the microbial community, thus lowering corrosion in lower sections, but higher corrosion may be noticeable in other areas. More studies in controlled environments should be utilized to further investigate this assumption. In addition, corrosion is a slow process; thus, a controlled environment study with varied designs spanning over 10 years may address such assumptions.

The primary importance of the study is to identify the key variables that contribute towards corrosion in manholes. Corrosion rate is a function of various microbial community, sewer system design characteristics and wastewater characteristics, and thus this research identifies major variables that can be controlled or regulated to reduce corrosion in manhole or measured regularly to understand the condition and life expectancy of the manholes. The corrosion rate equations available are specific only in certain conditions such as 1977 Pomeroy-Parkhurst equation which will not work for dissolved oxygen concentrations greater than 0.5 mg/l[33]. Also, corrosion rate equations already available are specific for sewer pipes; however. the gas phase conditions in the manhole are very different from sewer pipes. A manhole has a vertical headspace which is near to the ground surface, thus there is an influence of ambient temperature and precipitation on corrosion rates. Also, oxygen can easily diffuse into the manhole's headspace and increase rates of corrosion. In addition, previous research such as Wells et al.[43] have focused more on laboratory simulated environment, whereas this study captures the real-time changes.

In conclusion, we can say that though the model has accuracy of only 35.08%, it does a good job at identifying the key variables of corrosion in a manhole. Also, the corrosion model suggests the effect of design on the corrosion, indicating that variables that control corrosion rate in sewer pipes are likely different for manholes.

This first of its kind study helps to identify the important parameters that future research may focus on for further improving the hydrogen sulfide generation and volatilization model, and rate of corrosion model. Also, the study may help cities and government authorities to avoid vulnerable designs in manhole in their future work.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

The primary goal of the research was to build predictive models to determine the hydrogen sulfide generation/volatilization and corrosion rates occurring in manholes as a function of wastewater characteristics, manhole designs, and ambient weather conditions. Since no similar research had previously been conducted, the project focused on understanding the relationship between the various predictor variables and the response variables of hydrogen sulfide and corrosion rates. To achieve this goal, 366 manholes in the city of Arlington were sampled and 48 hours of liquid and gas phase parameters were collected. Manhole diameters were measured, and depth of corrosion was estimated. Preliminary analyses were conducted to verify our assumptions.

6.1 FINDINGS AND CONCLUSIONS

6.1.1 DATA CHARACTERIZATION

Findings and conclusions from characterizing the data are:

a. Gas phase hydrogen sulfide concentration was maximum 1 to 5 feet above the inlet, and the gas phase temperature increased with vertical depth along the manhole shaft.
b. In terms of manhole design categories, hydraulic jump design had the highest hydrogen sulfide concentration in the manhole's headspace, followed by right angle bends. Thus, design of the manhole did play a role in determining the concentration of hydrogen sulfide in the manhole's headspace to some extent.
c. Location of manholes also determined the amount of hydrogen sulfide in the manhole's headspace, with manhole downstream of the lift station recording the highest hydrogen sulfide in comparison to others.
d. Strong correlation was identified between the ambient temperature, gas phase temperature of the manhole, and liquid phase temperature of the wastewater. The average variation between the gas and ambient temperature was found to be 6.22 °F, and average variation between the gas phase and liquid phase temperature was found to be 3.48 °F. This resulted in diurnal patterns in hydrogen sulfide concentrations in the manhole irrespective of the design. Thus, higher hydrogen sulfide concentrations were recorded in the

afternoon in comparison to the early mornings. Also, hydrogen sulfide concentrations in the manhole vary seasonally, with summer recording highest hydrogen sulfide concentrations.

e. An inverse relationship between the wastewater pH and hydrogen sulfide was noticed, with highest pH recorded in the early mornings in the manholes.

f. Manhole designs and location also regulate the dissolved oxygen concentrations in the wastewater. Acute bend showed highest dissolved oxygen concentrations of 4.4 mg/l; also, supercritical flow, smaller to larger pipe size change, and larger to smaller pipe size change show values higher than the average DO, thus implying addition of dissolved oxygen via agitation and mixing. Lowest dissolved oxygen concentrations were recorded for the manhole downstream of lift station, due to the detention of wastewater and solids settling out in lift station, resulting in anaerobic conditions with relatively low oxygen.

g. Lower corrosion rates recorded in manhole designs expected to have turbulent flow (e.g. drops) might have been a result of washing away of the microbial community present on the manhole surface; however, in calmer flows, such as subcritical flow or obtuse bends, the microbial community has a chance to thrive, thus increasing corrosion. This result could have stemmed from our assumption that the lower sections of the manhole were more corroded for manhole corrosion determination. Maximum corrosion might occur in other areas.

6.1.2 MODEL FOR HYDROGEN SULFIDE GENERATION AND VOLATILIZATION

Multiple linear regression was conducted to develop a comprehensive hydrogen sulfide generation and volatilization equation in a manhole, Eq. 5.2, repeated here for convenience:

$$Log_{10}(0.01 + Y) = 1.514 + 0.3099V - 0.444Pr - 0.3251pH - 0.2374PS_{sl} - 0.376St + 0.497US_{H2S} - 0.1613(Super * DO) - 0.311Flow_{hydrualic} + 0.1072(Sulfate * DO) \ldots\ldots \text{Eqn } 5.2$$

Where:

Y = Gas phase hydrogen sulfide concentration in the manhole, ppm

V = Velocity of wastewater, ft/s

Pr = Precipitation, inches

pH = pH of the wastewater

PS_{sl} = Change in pipe size diameter (inlet pipe smaller than outlet pipe) (binary)

St = Absence of bends and multiple inlets (binary)

US_{H2S} = Upstream of a high hydrogen sulfide manhole (binary)

Supercritical*DO = Interaction between supercritical flow and dissolved oxygen (mg/l)

$Flow_{hydraulic}$ = Presence of hydraulic jump (binary)

Sulfate*DO = Interaction between the liquid sulfate concentration and dissolved oxygen (mg/l)

The model has a R^2 value of 35.97%, and adjusted R^2 value of 30.38%. Model validation showed that the modeled hydrogen sulfide concentration was an average of 42.98% above or below the measured hydrogen sulfide values in the manhole's headspace.

Model relationships are reasonable as follows:

a. Increasing velocity results in higher hydrogen sulfide concentrations, which is expected due to greater turbulence causing volatilization.
b. Precipitation decreases the hydrogen sulfide concentrations since precipitation would absorb the gas phase hydrogen sulfide.
c. Increasing pH decreases hydrogen sulfide concentrations, since at high pH the hydrogen sulfide in the wastewater will dissociate to ionic forms and thus is less available for volatilization to the manhole's headspace.
d. Absence of bends and multiple inlets decreases hydrogen sulfide, as there would be limited turbulence.
e. Being present upstream of a manhole with high hydrogen sulfide levels increases hydrogen sulfide concentrations.
f. The interaction of supercritical flow with dissolved oxygen decreases hydrogen sulfide, implying that if supercritical flow incorporates oxygen into the wastewater due to agitation, then hydrogen sulfide concentrations will lessen. Supercritical flow, however, could also strip more hydrogen sulfide into the gas phase due to turbulence. According to the model, the former effect predominates.

However, the reason that a smaller-to-larger pipe size change and hydraulic jump decreases hydrogen sulfide concentrations is unclear. Also, the interaction of sulfate and dissolved oxygen increases hydrogen sulfide according to the model. Increased sulfate would be expected to increase formation of sulfides; however, dissolved oxygen should reduce their formation.

6.1.3 MODEL FOR CORROSION RATE

Multiple linear regression was conducted to develop Equation 5.15 (repeated here for convenience) for rate of MICC in a manhole:

$$Log_{10}(0.01 + Y) = -2.683 + 0.1611V - 0.53Bend_{right} - 0.582PS_{ls} + 0.000521BOD - 0.052Drop_{max} + 0.362\left(Log_{10}(0.1 + DO)\right) + 0.03105T_{liq} - 0.23Inlets_{multiple} - 0.2169(T_{gas} * Age) \ldots\ldots \text{ Eqn. 5.15}$$

Where:

Y = Rate of corrosion, mm/yr

V = Velocity of wastewater, ft/s

$Bend_{right}$ = Presence of a right-angle bend

PS_{ls} = Change in pipe size diameter (inlet pipe larger than outlet pipe) (binary)

BOD = Wastewater BOD (mg/l)

$Drop_{max}$ = Maximum drop height (ft)

DO = Wastewater dissolved oxygen (mg/l)

T_{liq} = Wastewater temperature (F)

$Inlets_{multiple}$ = Presence of multiple inlets

T_{gas}*Age = Interaction between the manhole gas phase temperature and age (F*yrs)

The model has a R^2 value of 35.1%, and adjusted R^2 value of 28.7%. Model validation shoed that the modeled corrosion rate was an average of 30.42% above or below the measured hydrogen sulfide values in the manhole's headspace.

Corrosion model relationships can be explained as follows:

a. Corrosion rate increases with velocity which is in accordance with our assumption that increased velocity will result in increased wastewater turbulence, thus higher hydrogen sulfide volatilization, and thus more hydrogen sulfide been available for the microorganisms to cause corrosion, thus faster rate of corrosion.
b. An increase in BOD of the wastewater increased the corrosion rate, which is reasonable as higher BOD will result in faster DO consumption, thus more anaerobic conditions, and excessive hydrogen sulfide production and volatilization, thus higher rates of corrosion.
c. An increase in liquid temperature results in higher hydrogen sulfide volatilization and thus higher rates of corrosion, as anticipated.
d. The model also confirmed that concrete age does play an important role in determining rates of corrosion. A relationship between the temperature and age was identified as an important factor by the MLR model in the corrosion rate equation.
e. Three designs expected to have higher turbulence - drops, right angle bends, and multiple inlets – actually had decreased rates of corrosion. This may be due to washout of the microbial community on the concrete surface. A larger pipe leading into a smaller pipe also decreased rate of corrosion.

6.1.4 OVERALL IMPORTANCE OF STUDY

The primary importance of the study is it helps to identify the key parameters involved in hydrogen sulfide generation and volatilization and corrosion rate in manholes. The study, being the first of its kind, provides an opportunity to cities and government authorities to focus their repair and rehabilitation efforts on the manhole designs identified by the MLR model. As discussed in Ch. 5, the MLR equations developed to estimate hydrogen sulfide concentrations in this research are very different from existing equations for estimating hydrogen sulfide concentrations in sewers, indicating that the sewer equations do not apply to manholes, and the new equations are needed.

This study indicates that to minimize hydrogen sulfide generation and release, wastewater system/manhole designers should minimize wastewater flow velocity, avoid supercritical flow, and avoid multiple manhole inlets and bends as

much as possible. It also indicates that reducing wastewater flow velocities will reduce corrosion. Future research should investigate whether increasing right-angle bends, drops, and multiple inlets reduces corrosion, as this study seems to suggest, presumably by reducing the thickness of the microbial biofilm in which hydrogen sulfide is converted to sulfuric acid.

6.2 RECOMMENDATIONS FOR FUTURE RESEARCH

6.2.1 CONSIDER ADDITIONAL FACTORS TO EXPLAIN VARIABILITY IN HYDROGEN SULFIDE CONCENTRATIONS AND CORROSION RATES

The 36% and 35% R^2 values of the hydrogen sulfide and corrosion models, respectively, indicate that the models are explaining 36% and 35% of the variability in the data. Factors influencing microbial concrete corrosion fall into 4 categories: those related to the microbial community, wastewater flow and composition, manhole design (which impacts wastewater flow and composition), and ambient environment. In this study, we collected data concerning wastewater composition and the ambient environment, and used modeled data provided by the city concerning wastewater flow, as well as information provided by the city and our own observations concerning manhole design. Measurements of the first factor, however – microbial community – were not conducted for reasons of safety (manhole entry would have been required to sample microorganisms along the surface of the manhole shaft) as well as expense (PCR identification of microorganisms is expensive). The microbial community is the factor responsible for hydrogen sulfide generation and its conversion to sulfuric acid. Thus, future research to identify the microorganism population in the concrete, and its diversity based on wastewater characteristics, could reduce model uncertainty. To avoid the issues associated with field safety, growth of microbial communities on manhole shafts could be replicated in the laboratory to assess the change in microbial diversity with time and changing concrete characteristics as the corrosion progresses.

Since manholes are vertical structures, assessing the change in microbial community along the depth of manhole, especially in very deep manholes, might suggest which sections of the manhole might be more prone to corrosion.

Moreover, one of the conclusions we arrived in our research was that in less turbulent conditions, microbial community has a steady environment to grow and thrive, and thus a reason we see higher corrosion rates in those manhole designs, whereas in turbulent designs such as hydraulic jump or drop, the wastewater strips away the microbial community with constant water splashing. Thus, an in-house study confirming this assumption by the comparing less turbulent design microbial community with that of turbulent design will validate the conclusion.

In addition to microbial community, concrete pH was not measured because it would have required manhole entry. Concrete pH might control the rate of microbial growth and could be an important factor for corrosion. Future lab studies could examine this factor.

6.2.2 IMPROVE THE ACCURACY OF CORROSION RATE MEASUREMENTS

Since our manholes were old, it became difficult to evaluate the rate of corrosion occurring in the manhole, as documentation regarding the original manhole diameter was unavailable. As an alternative to field data collection, cities could select designs from the corrosion rate model recommend in this research, and develop in-house constructed designs under a controlled environment, where the corrosion rate can be frequently measured. Since corrosion is a slow process, long term corrosion measurements over a long period, such as 10 years, would be helpful. As another idea the city could carry out regular manhole diameter measurements in the field using the CleverScan for a long period of time such as 10 years.

Another possibility to improve the accuracy of corrosion measurements would be to assess corrosion near the drops. In our study we noticed that hydrogen sulfide concentration was highest 1 to 5 feet above the inlet; thus, assessing corrosion in the drop in addition to the lower 20% might incorporate any localized corrosion that might be taking place near the drops. Also, measurements of hydrogen sulfide near the drops and near the water surface, in addition to corrosion measurement, may help to develop a relation between the concentration of hydrogen sulfide and rate of corrosion in these localized regions.

6.2.3 ADDITIONAL FIELD AND LAB DATA COLLECTION

6.2.3.1 MANHOLE DESIGN WITH LIMITED NUMBER OF MANHOLES

The city has an unequal distribution of manhole designs: radical designs such as hydraulic jump, or downstream of a lift station manhole, were limited. Sampling from more of these manholes will add more comprehensive design consideration to the model. We were limited to manholes with modeled velocity data. Expanding the modeling could provide more manholes with radical designs that could be field sampled. If data can be collected for these designs with different wastewater characteristics and ambient conditions, it may train the model to anticipate more diverse scenarios, thus improving its prediction. In our research we collected data only for 48 hours; however, collection of more than 48 hours of data for few manholes may provide a clearer view of the trends occurring in the wastewater variables, thus helping to draw better conclusions on the relation of hydrogen sulfide generation and volatilization based on design and wastewater characteristics.

6.2.3.2. VELOCITY DATA

Velocity was a key factor in our models, however the velocity used in our study was modelled. In future, measuring real-time velocities might be show how change in velocity is affecting hydrogen sulfide volatilization in different manhole designs.

6.3.2.3 PRECIPITATION DATA

Our research could not focus extensively on high precipitation days due to technical and safety reasons; however, Eqn. 5.2 showed precipitation as an important criterion in reducing the hydrogen sulfide concentrations in the manhole. Use of water sprayers in in-house lab experiments to replicate the effect of rain could help clarify the effect of precipitation on hydrogen sulfide concentrations. Also, effect of precipitation on concrete over long periods of rain can be evaluated through this method. Precipitation may also change microbial community functions and there might be interactions which can be added to the model to improve its accuracy.

6.2.3.3 DEEPER MANHOLES

Manholes more than 30 feet in depth could not be considered due to safety concerns; however, in future research can focus on deeper manholes. Replicating deeper manhole designs in the laboratory might be helpful in this regard. In addition, the effect of precipitation on deeper manholes can also be studied evaluated, as with depth the effect of precipitation might reduce or change. Also, we were able to study the variation in manhole's gas phase temperature and hydrogen sulfide concentrations with depth. However, measuring oxygen and relative humidity variation with depth might be useful to understand the reactions and interactions between the various wastewater constituents and gas phase variables.

REFERENCES

1. National Oceanic and Atmospheric Administration; Record of Climatological Observations – Station Arlington municipal Airport; August 2017 – May 2020. Infor from URL: https://www.ncdc.noaa.gov/cdo-web/datasets/GHCND/stations/GHCND:USW00053907/detail
2. Gas Logger: OdaLog Gas Logger User Manual. App-Tek International Pty Ltd. (2010). P/N: 10-1000
3. Kestrel Drop: Wireless Environmental Data Loggers. Getting Started. Global Test Supply. Info from URL: http://www.globaltestsupply.com/pdfs/cache/www.globaltestsupply.com/kestrel/temperature_humidity_recorder/0720blu/datasheet/kestrel_0720blu_temperature_humidity_recorder_datasheet.pdf
4. ToxiRAE Pro: User's Guide. Rae Systems by Honeywell. (July 2015). P/N G02-4009-000 Rev. E.
5. Aqua TROLL 600 Multiparameter Sonde: Operator's Manual. Part Number 0096400. In-Situ. (August 2016).
6. HI98194, HI98195, HI 98196 Multiparameter Meters: Instruction Manual. Hanna Instruments. (May 2015). Romania. Info from URL: https://www.fieldenvironmental.com/assets/files/Manuals/Hanna-Manual-98194-95-96_29_05_15.pdf
7. 6712 Portable Samplers: Installation and Operation Guide. Teledyne ISCO. (May 2015).
8. Power Products Guide: Installation and Operation Guide. Teledyne ISCO. (April 2011).
9. Methods for Chemical Analysis of water and wastes. Environmental Monitoring and Support Laboratory, Office of Research and Development, United States of Environmental Protection Agency. Cincinnati, Ohio. (March 1983).
10. Rodger B. Baird, Andrew D. Eaton, Eugene W. Rice. Standard Methods for the Examination of Water and Wastewater: 23rd Edition. American Public Health Association, American Water Works Association, Water Environment Federation. Washington, DC. (2017). Electronic.
11. Method 9030B - Acid-soluble and Acid-Insoluble Sulfides: Distillation. United States of Environmental Protection Agency. (December 1996).
12. Richard Pomeroy. Auxiliary Pretreatment of Zinc Acetate in Sulfide Analyses. *Analytical Chemistry* 1954 *26* (3), 571-572. DOI: 10.1021/ac60087a047

13. Method 9034 – Titrimetric Procedure for Acid-soluble and Acid-Insoluble Sulfides. United States of Environmental Protection Agency. (December 1996).
14. Code of Federal Regulations, Title 40, Chapter 1, Subchapter C, Part 60, Appendix A-5, Method 11. 36 FR 24877, Dec. 23, 1971. Info from URL: https://www.ecfr.gov/cgi-bin/text-idx?SID=48672866df47b98d250106a369473cf7&mc=true&node=pt40.9.60&rgn=div5#ap40.9.60.a_65
15. Method 9038 - Acid-soluble and Acid-Insoluble Sulfides: Distillation. United States of Environmental Protection Agency. (December 1996).
16. Spectrophotometer DR 2800: User Manual Edition 4. Hach Company. (August 2013).
17. Sulfate, Turbidimetric TNTplus Method(900 mg/l), Edition 9. Hach. (2019).
18. Sulfate, Turbidimetric TNTplus Method(150 mg/l), Edition 8. Hach. (2019).
19. Clifford C. Hach, Robert L. Klein, Jr., Charles R. Gibbs. Introduction to Biochemical Oxygen Demand, Technical Information Series – Booklet No. 7. Hach Company. (1997).
20. Oxygen Demand, Biochemical, Edition 9. Hach Company. (August 2015).
21. BOD LDO Probe – Model LBOD10101, Edition 3. Hach Company. (September 2014).
22. HQd Portable Meter: User Manual, Edition 4. Hach Company. (June 2013).
23. CleverScan. Environsight. Info from URL: https://www.envirosight.com/dwnld/CleverScan.pdf
24. Michael H. Kutner, Christopher J. Nachtsheim, John Neter, William Li. Applied Linear Statistical Models, Fifth edition. (2004).
25. What is Monte Carlo Simulation; Palisade. Info from URL: https://www.palisade.com/risk/monte_carlo_simulation.asp
26. Developing a Manhole or Catch Basin Numbering System; A Systematic Approach for Electronic Data Storage and Linkage. United States of Environmental Protection Agency. Info from URL: https://www3.epa.gov/region1/sso/manhole-id.html
27. Manholes. Designing Buildings Wiki. Info from URL: https://www.designingbuildings.co.uk/wiki/Manhole
28. Manhole – Purpose, Types, and Construction. The Constructor. Info from URL: https://theconstructor.org/practical-guide/manhole-purpose-types-features/36941/

29. Sewer construction. Designing Buildings Wiki. Info from URL: https://www.designingbuildings.co.uk/wiki/Sewer_construction#Manhole_chambers
30. Collection systems technology fact sheet; Sewers, conventional gravity. United States of Environmental Protection Agency. Info from URL: https://nepis.epa.gov/Exe/ZyPDF.cgi/P10053D9.PDF?Dockey=P10053D9.PDF
31. Subchapter c: conventional collection systems §§217.51 - 217.71; Chapter 217 - Design Criteria for Domestic Wastewater Systems. Texas Commission on Environmental Quality. December 4, 2015. Info from URL: https://www.tceq.texas.gov/assets/public/legal/rules/rules/pdflib/217c.pdf
32. Rule §317.2, Sewer collection system; Chapter 317 – Design criteria prior to 2008; Part I - Texas Commission on Environmental Quality; Title 30 – Environmental Quality. Texas Administrative Code. Info from URL: https://texreg.sos.state.tx.us/public/readtac$ext.TacPage?sl=T&app=9&p_dir=F&p_rloc=174662&p_tloc=14842&p_ploc=1&pg=2&p_tac=&ti=30&pt=1&ch=317&rl=2
33. Sulfide Task Group of the Water Pollution Management Committee of the Environmental Engineering Division of the American Society of Civil Engineers. Sulfide in wastewater and collection systems. ASCE Manual and Reports on Engineering Practices 69. (1989). ISBN 0-87262-681-4.
34. Water Environment Federation, American Society of Civil Engineers, Water Environment Federation. Odor control in wastewater treatment plants. WEF Manual of Practice 22. (1995).
35. Joanne B Hughes. Manhole Inspection and Rehabilitation. ASCE Manual and Reports on Engineering Practices 92. (2009).
36. Manholes: The Ultimate Resource. Fibertech inc. Info from URL: https://fibertechinc.net/rhino-manholes/manholes-ultimate-resource/
37. Hydrogen sulfide corrosion in wastewater collection and treatment systems; Technical Report. United States of Environmental Protection Agency. September 1991.
38. State of technology for rehabilitation of wastewater collection systems. United States of Environmental Protection Agency. Office of research and development, Washington. July 2010.

39. Detection, Control, and Correction of hydrogen sulfide corrosion in existing wastewater systems. United States of Environmental Protection Agency. Office of wastewater enforcement and compliance. September 1992.
40. Hydrogen sulfide corrosion in wastewater collection and treatment systems. United States of Environmental Protection Agency. Office of water, Washington. May 1991.
41. Tony Wells; R.E. Melchers: Phillip Bond. Factors involved in the long term corrosion of concrete sewers. January 2009. Info from URL: https://www.researchgate.net/publication/43527972_Factors_involved_in_the_long_term_corrosion_of_concrete_sewers
42. T. Wells, R.E. Melchers (2014). An observation-based model for corrosion of concrete sewers under aggressive conditions. Cement and Concrete Research, 61-62, 1-10.
43. Xiaoyan Sun, Guangming Jiang, Philip L. Bond, Tony Wells, Jurg Keller. A rapid, non-destructive methodology to monitor activity of sulfide-induced corrosion of concrete based on H_2S uptake rate. Water research 59 (2014) 229-238.
44. Tony Wells , Robert Melchers , Antony Joseph , Phil Bond , Dammika Vitanage , Heriberto Bustaante , John De Grazia , Thomas Kuen , John Nazimek , Ted Evans. A collaborative investigation of microbial corrosion of concrete sewer pipe in australia. Australian Water association. Conference paper. 2012. ISBN:9781921335198
45. Romanova, A., Mahmoodian, M., & Alani, M. A. (2014). Influence and interaction of temperature, H_2S and pH on concrete sewer pipe corrosion. International Journal of Civil, Architectural, Structural, Urban Science and Engineering, 8(6), 592-595.
46. Gutiérrez-Padilla, M. G. D., Bielefeldt, A., Ovtchinnikov, S., Hernandez, M., & Silverstein, J. (2010). Biogenic sulfuric acid attack on different types of commercially produced concrete sewer pipes. Cement and Concrete Research, 40(2), 293-301.
47. Jiang, G., Keller, J., & Bond, P. L. (2014). Determining the long-term effects of H_2S concentration, relative humidity and air temperature on concrete sewer corrosion. Water research, (65) 157-169.
48. Amir M. Alani., Asaad Faramarzi., Mojtaba Mahmoodian., Kong Fah Tee (2014). Prediction of sulphide build-up in filled sewer pipes. Environmental Technology, (35), 1721-1728.

49. P.A. (Tony) Wells & R.E. Melchers (2014). Findings of a 4 year study of concrete sewer pipe corrosion. Corrosion & Prevention 2014 Paper 22. Info from URL: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&uact=8&ved=2ahUKEwjaptWG5LHpAhUqhHIEHfwlC6wQFjAEegQIAhAB&url=https%3A%2F%2Fnova.newcastle.edu.au%2Fvital%2Faccess%2Fservices%2FDownload%2Fuon%3A9061%2FATTACHMENT01&usg=AOvVaw3OcPjovy6C5mSsp5aaY7gA

50. Melchers. 18th International Corrosion Congress 2011. Conference paper. Info from URL: https://nova.newcastle.edu.au/vital/access/manager/Repository/uon:14382

51. Tony Wells, Robert Melchers. Concrete sewer pipe corrosion – findings from an Australia field study. Info from URL: https://pdfs.semanticscholar.org/b1ff/dd03e19c0da5e5617a0e51ced9d2454ac8b6.pdf

52. Xuan Li, Faezehossadat Khademi, Yiqi Liu, Mahmoud Akbari, Chengduan Wang, Philip L. Bond, Jurg Keller, Guangming Jiang (2019). Evaluation of data-driven models for predicting the service life of concrete sewer pipes subjected to corrosion. Journal of environmental management, (234) 431-439.

53. Kaempfer W. Estimation of service life of concrete pipe in sewer network. 8th Conference on Durability of Building Materials and Components. National Research Council of Canada, Ottawa, 1999, Vol. 1, pp. 36–45.

54. Odor and corrosion control in sanitary sewerage systems and treatment plants. Design manual. United States of Environmental Protection Agency. Office of research and development, Washington. October 1985. Info from URL: https://play.google.com/books/reader?id=4jGq_R-dRK4C&hl=en_GB&pg=GBS.PA110

APPENDIX A

REAGENT PREPARATION METHODOLOGY

A.1 PRESERVATIVE PREPARATION FOR SULFIDE ANALYSIS

We will be adding 60 grams of sodium hydroxide to 250 ml of deionized water to make 6 mol/liter (6N) of the NaOH solution. To this, we will be adding approximately 25 ml of 2N (1.15M) of zinc acetate to maintain the pH level of the NaOH solution at 9.

A.2 CALCULATION FOR PRESERVATIVE PREPARATION

The following calculation meets the EPA guidelines of preservative addition.

Recommendation: 0.20 ml of 2N (220g in 870 ml) of zinc acetate + 0.20 ml of 6N (240 g in 1000 ml) to 100 ml of the sample solution. Thus, 0.6072 g (0.00276 M) of zinc acetate is added to 100 ml of sample solution, and 0.48 g (0.012 M) of sodium hydroxide is added to 100 ml of sample solution. According to standard methods, zinc acetate concentration can be changed based on our sample's conditions[(10)].

For our research: 5 ml of preservative solution which is made up of 25 ml of 2N (220g in 870 ml) of zinc acetate + 250 ml of 6N (240 g in 1000 ml) is added to 500 ml of the sample solution. Thus, 0.1595 g (0.00073 M) of zinc acetate is added to 500 ml of sample solution, and 2.4 g (0.006 M) of sodium hydroxide is added to 500 ml of sample solution.

A.3 TITRANT - SODIUM THIOSULFATE PREPARATION FOR SULFIDE TITRIMETRIC METHOD

250 ml of 0.1N Sodium thiosulfate was added to 750 ml of deionized water to prepare a 0.025N sodium thiosulfate solution to be used a titrant.

A.4 IODINE PREPARATION OFR SULFIDE TITRIMETRIC METHOD

24 g of potassium iodide (KI) was added in 30 ml of deionized water to prepare the potassium iodide (KI) solution. 12.7 g of resublimed iodine was added to the KI solution, and shaken to dissolve the iodine. The 0.1N iodine solution was made up to 1 liter and stored in a dark place. 250 ml of the 0.1N iodine solution was added to 750 ml of deionized water to prepare 0.025N of iodine solution to be used in the sulfide titration[(14)].

A.5 DILUTION WATER PREPARATION FOR BOD ANALYSIS

For dilution water preparation we will be using the HACH BOD Nutrient Buffer Pillow. The method followed for dilution water preparation is HACH Method 8043[20]. Distilled water will be used for the preparation of the dilution water. Distilled water will be aerated using an air compressor for 24 hours before use[19,20]. 1 HACH BOD Nutrient Buffer Pillow is shaken and emptied into 6 liters of aerated distilled water to prepare the dilution water[20].

APPENDIX B

EFFECT OF DIFFERENT SEASONS ON HYDROGEN SULFIDE RESULTS

B.1 EFFECT OF DIFFERENT SEASONS ON GAS PHASE TEMPERATURE

To better understand the relationship between seasons and its effect on hydrogen sulfide concentration, 24 manholes were repeated in different seasons. The Fig B.1, and Fig B.2 shows the relationship between ambient temperature with gas phase temperature inside the manhole and average hydrogen sulfide concentration for 6 manholes. In Fig B.1, the bar chart represents the average ambient temperature in Fahrenheit on the days the particular manhole was done, and the markers represent the average gas phase temperature in Fahrenheit inside the manhole. Similarly, in Fig B.3 the bar chart represents the average ambient temperature in Fahrenheit on the days the particular manhole was done, and the markers represent the average gas phase hydrogen sulfide concentration in ppm inside the manhole. The colors represent the measurements done in different seasons with green colors for spring, yellow for summer, red for fall and blue for winter.

In the Fig B.1 we can clearly see that ambient gas phase temperatures were highest during the summer and lowest in winter with fall and spring temperature values falling in the middle. Due to this variation in the ambient temperatures, we can see a clear trend in the gas phase temperatures inside the manhole with the gas phase temperatures showing very high values in comparison to winter blue square markers while the fall and spring markers falling in the middle. In addition, it seems there is an average of 21 F difference between the summer and winter gas phase temperature, with fall and spring temperatures having an average difference of 1 F with each other and 10 F with winter and summer, as shown in Table B.1. Also, the difference between the gas phase temperatures between different seasons seem to increase from Summer to Spring, which implies that highest change in gas phase temperature in the manhole is noticed between winter and spring, and lowest between summer and fall.

Table B.1 Average gas phase temperature variability in different seasons

Seasons	Average difference between gas phase temperatures
Summer to Fall	9.41
Fall to Winter	10.84
Winter to Spring	12.29
Spring to Summer	11.34
Summer and Winter	21.23
Fall and Spring	1.40

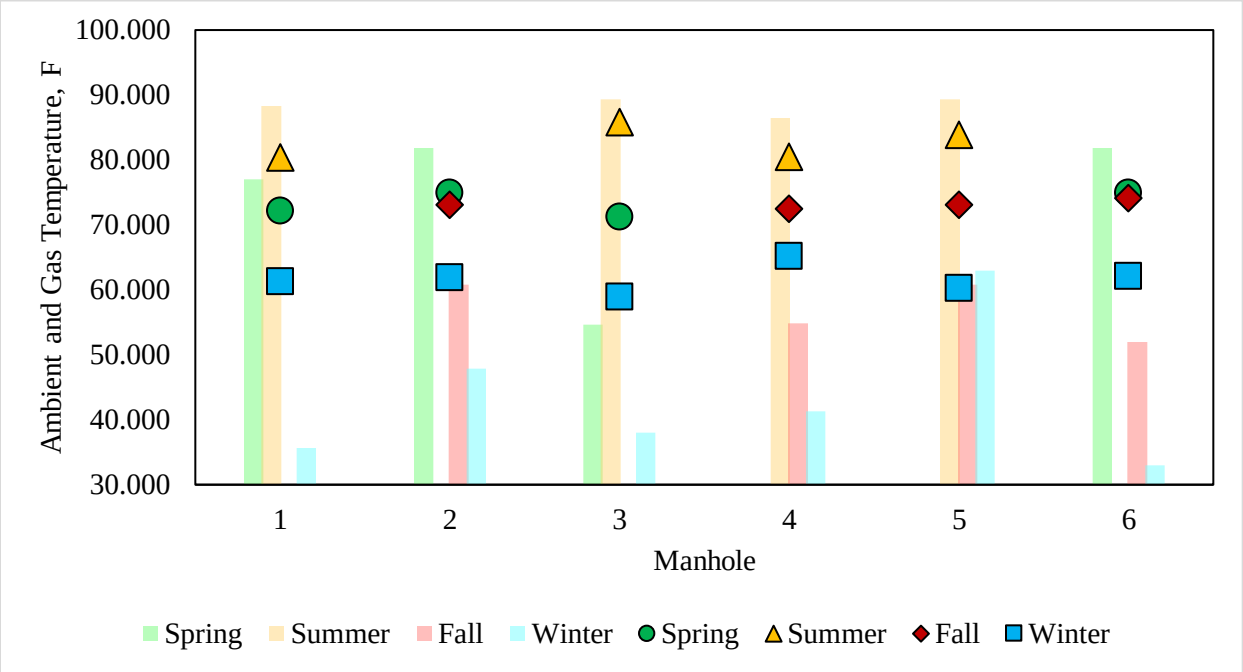


Fig B.1 Average ambient temperature vs average gas phase temperature

From the above discussion is seems there is a pattern to gas phase of temperatures inside the manhole to that of the ambient temperatures. Thus, the rate of change of gas phase temperature with respect to ambient temperatures was evaluated for 24 manholes repeated in different seasons, and it was found that for every 1 F change in ambient

temperature the gas phase temperature changes by 0.38 F, with a good R^2 value of 59.76%. The Fig 4.25 shows that the rate of change in the gas phase temperatures with respect to the ambient temperatures.

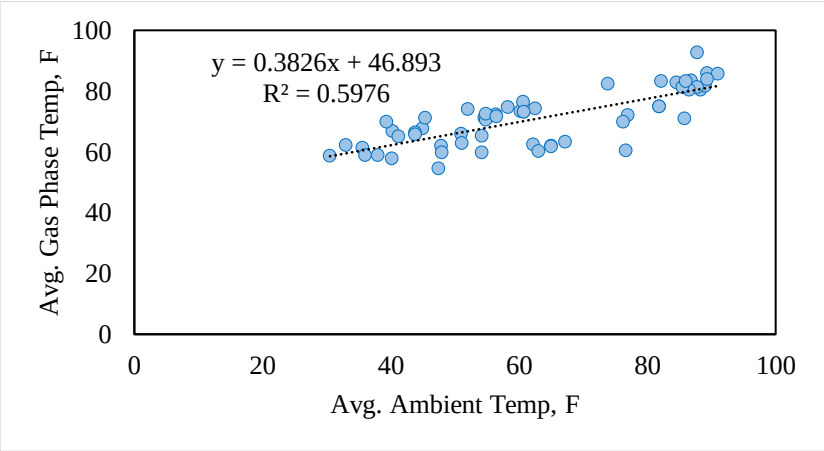


Fig B.2 Average ambient temperature vs average gas phase temperature

B.2 EFFECT OF DIFFERENT SEASONS ON HYDROGEN SULFIDE:

In the Fig B.3 we can clearly see that ambient gas phase temperatures were highest during the summer and lowest in winter, (except for manhole 5 where the winter ambient temperature recorded 3 F higher than the fall temperature), with fall and spring temperature values falling in the middle. Due to this variation in the ambient temperatures, we can see a clear trend in the hydrogen sulfide concentration inside the manhole with summer recording high values, followed by spring, with fall and winter recording the lowest values as seen in Fig. B.3 and Fig B.4. The hydrogen sulfide concentration in fall and winter does not seem to show a clear trend. This could be because at higher ambient temperature might contribute more towards hydrogen sulfide volatilization, but at lower or colder temperatures other factors might play a more important role in determining the volatilization of hydrogen sulfide. As higher temperatures increase microbial activity which then results in increased sulfide generation in the manhole thus more hydrogen sulfide available for volatilization[(37)], and thus we observing higher hydrogen sulfide concentration in the manhole. In addition, hydrogen sulfide solubility decreases with increase in temperature thus, at higher temperatures, we can see more effect of it than in lower temperatures. Also, looking at the Fig B.4 and Fig B.5, it is clear that the effect of temperature on hydrogen sulfide concentration was more prominent in summer and spring in comparison to winter and fall, with manholes showing rise of over 20 ppm in hydrogen sulfide concentration,

similar to results obtained by T.Wells who reported sewer systems having >100 ppm of hydrogen sulfide concentration in summer/autumnal months in comparison to winter which had ~40 ppm[42]. In addition, it seems the average highest change in hydrogen sulfide concentration is observed between spring and winter change of 5.98 ppm and lowest between fall and winter with an average change of 0.54 ppm as shown in Table B.2.

Table B.2 Average hydrogen sulfide variability in different seasons

Seasons	Average difference between hydrogen sulfide, ppm
Winter to Spring	5.98
Summer to Spring	3.83
Summer to Fall	1.71
Fall to Winter	0.54
Summer and Winter	0.90
Fall and Spring	3.79

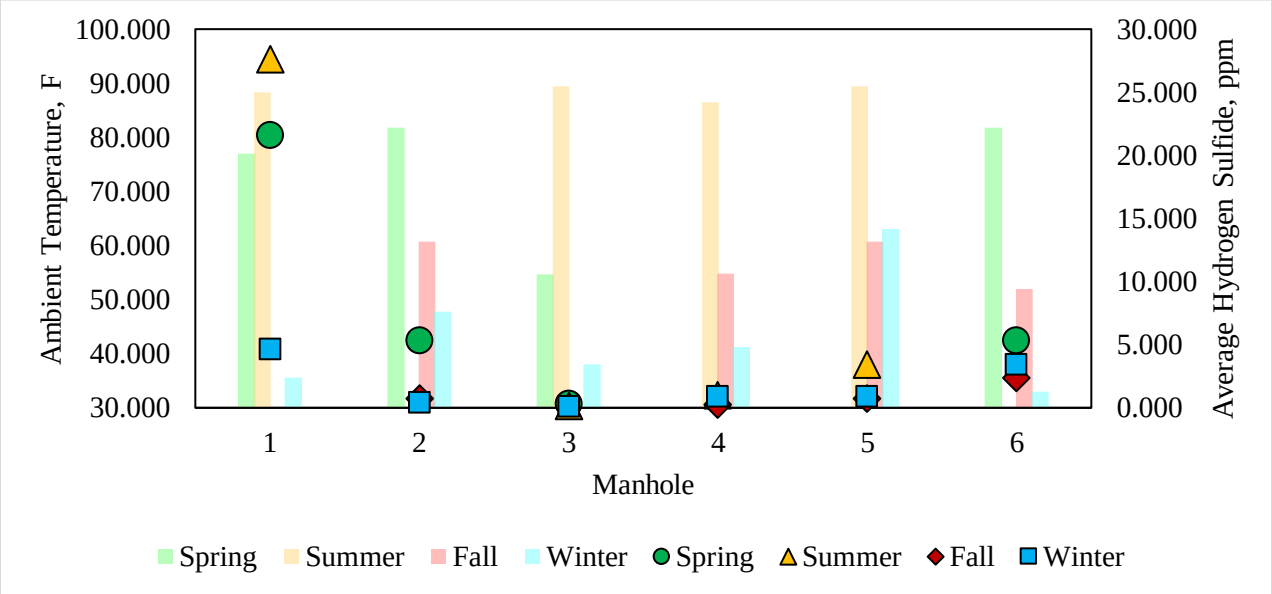


Fig B.3 Average ambient temperature vs average hydrogen sulfide

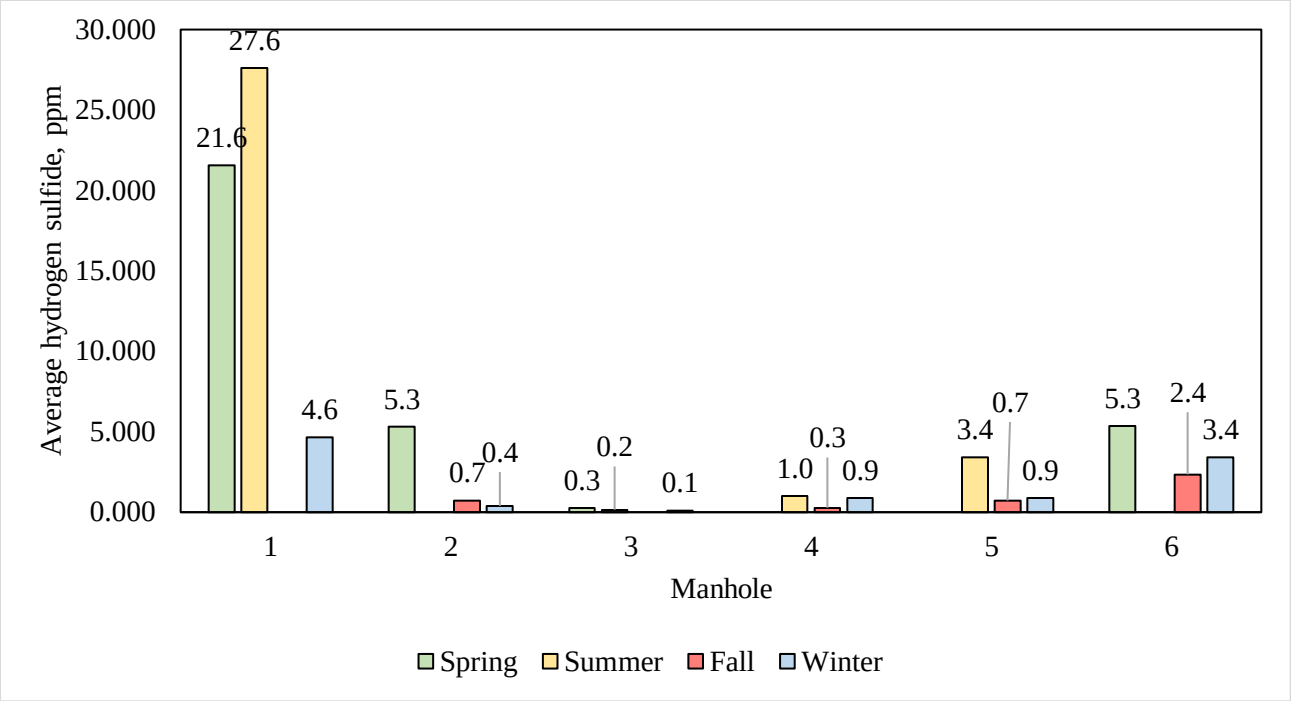


Fig B.4 Average hydrogen sulfide concentrations in different seasons

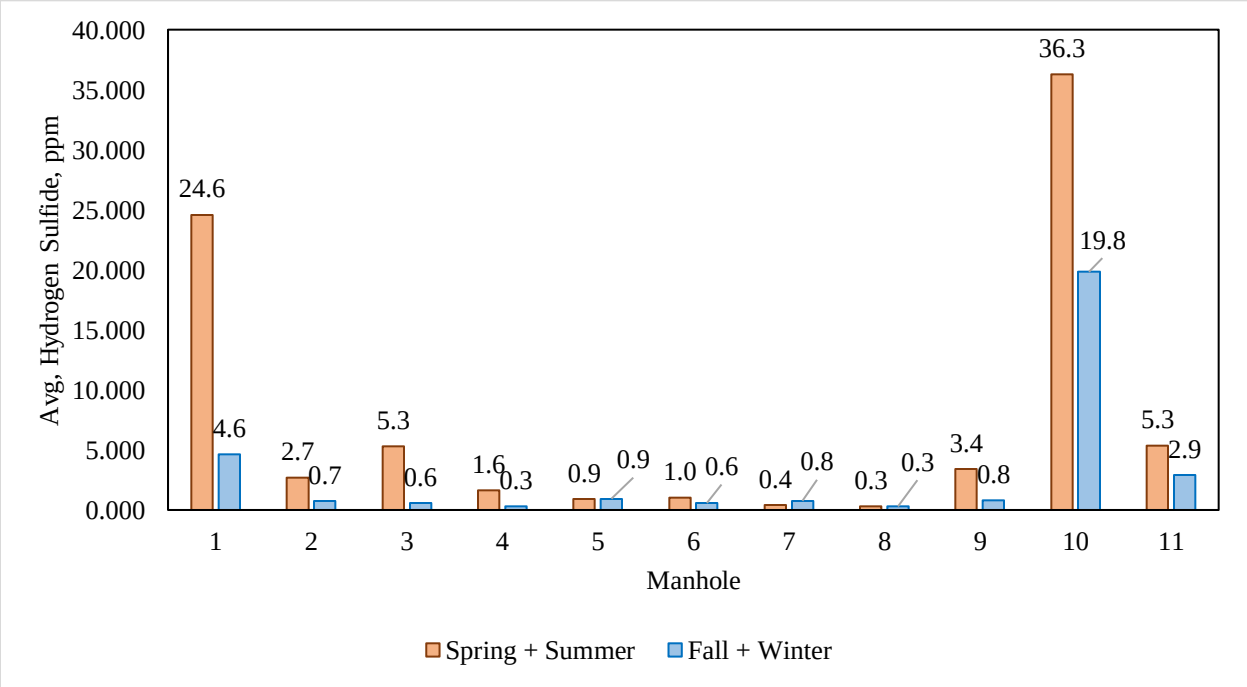


Fig B.5 Average Hydrogen Sulfide in warmer and colder seasons

From the above discussion it seems there is a pattern to change of hydrogen sulfide concentrations in the manhole to that of the ambient temperatures. Thus, the rate of change of gas phase temperature with respect to ambient temperatures was evaluated for 24 manholes, and it was found that for every 1 F change in ambient temperature the hydrogen sulfide concentration changes by 0.045 ppm, which is similar to the value reported by EPA, which stated sulfide production increases 7% for every Celsius degree increase up to 40°C[39].However the R^2 value is only 1.4%, thus implying though there is a relationship as clear from above graphs, it is not directly related and other factors might be contributing as well. The Fig B.6 shows that the rate of change in the gas phase temperatures with respect to the ambient temperatures.

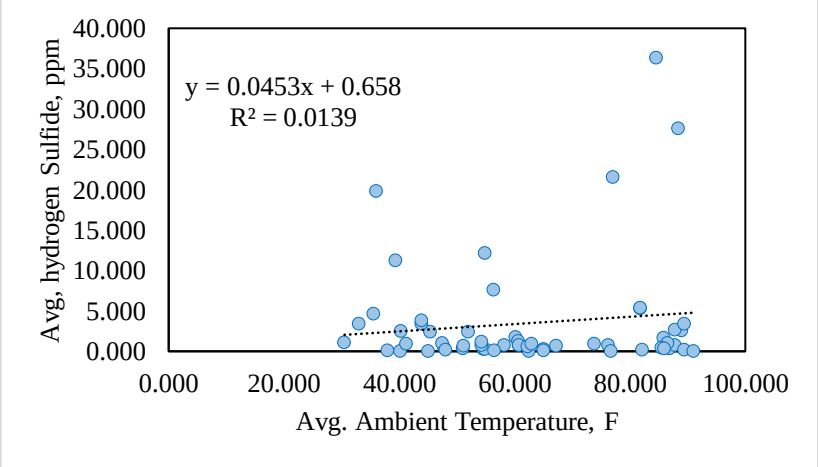


Fig B.6 Average ambient temperature vs average hydrogen sulfide

B.3 MANHOLES REPEATED IN SAME SEASONS:

To confirm if seasons did result in higher hydrogen sulfide concentration, and if similar results will be obtained next year in the same season, we collected liquid and gas temperatures values again for 20 manholes in the same seasons. Below Fig B.7 and Fig B.8, shows 7 different manholes, which were repeated in 2 different years in the same seasons. Milder shade bar on the left represents ambient temperatures from season 1 and brighter share bar on the right represent same manhole repeated in season 2. Markers in circle represent average gas phase temperatures and average hydrogen sulfide concentration in season 1 in Fig B.7 and Fig B.8 respectively. Markers in triangle represent average gas phase temperatures and average hydrogen sulfide concentration in season 2 in Fig B.7 and Fig B.8 respectively.

In the Fig B.7 we can clearly see that the average ambient temperatures are highest in summer1 and 2 in comparison to another season. The high ambient temperatures thus result in average higher gas phase hydrogen sulfide concentration in summer 1 and summer 2 for manhole 3 and 4 in the graph. Also, winter 1 has low ambient temperatures resulting in gas phase temperature for winter 1 being the lowest. However, the winter 2 ambient temperature is almost comparable to spring and fall temperature thus gas phase temperature in winter 2 is like spring values. This clearly demonstrates that gas phase temperatures are a seasonal phenomenon with gas phase temperature following an annual and diurnal pattern.

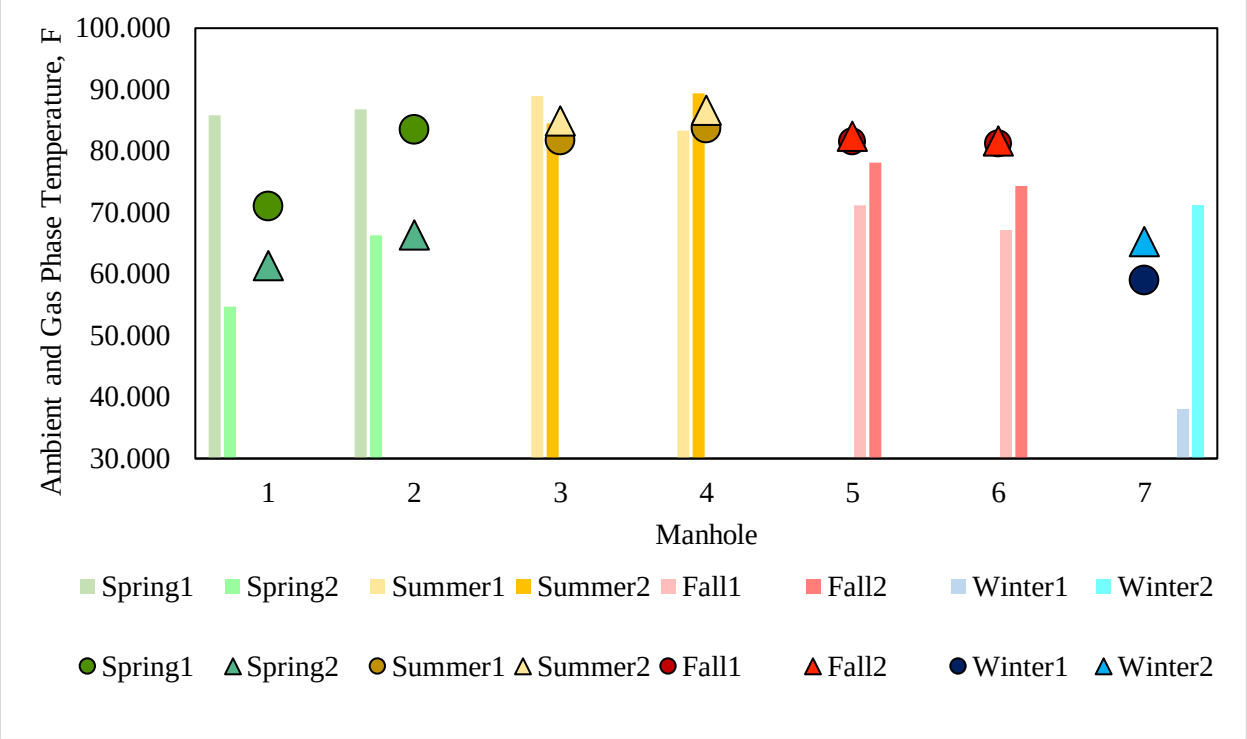


Fig B.7 Average ambient temperature vs average gas phase temperature for manholes in same seasons

In the Fig B.8, the trend is not clear which could be due to other factors such as microbial activity or manhole designs contributing to hydrogen sulfide concentrations as well. However, we can clearly see that when the ambient temperatures are higher than 70F all the manholes do show a rise in hydrogen sulfide levels, as seen in Spring 1 for manhole 1 and 2, for manholes 3 and 4 in summer 1 and 2, manhole 5 in fall 2, and winter 2 for manhole 7. Only Fall 2 for manhole 6 showed a decline in hydrogen sulfide levels even when the ambient temperatures were higher

than previous year fall. Also, when the temperatures were less than 65 F, the manhole recorded 0 ppm of hydrogen sulfide, even when the same manhole recorded higher values in the same season in another year, as seen in Spring 2 for manhole 1 and 2, and winter 1 for manhole 7. Only Fall 1 for manhole 5 did not follow this trend.

This proves that higher temperatures do effect hydrogen sulfide concentration, but lower temperatures may not have a direct or a prominent impact on the hydrogen sulfide concentration. This result is like what was observed for manholes done in different season where fall and winter manholes did not show good association with the ambient temperature however summer and spring did. This also suggest that at lower temperatures other factors such as design, or wastewater characteristics may play more important role in determining hydrogen sulfide concentration and thus controlling the corrosion rates. In addition, hydrogen sulfide solubility decreases with increase in temperature thus, at higher temperatures, thus we can see more effect of higher temperature than lower temperatures.

In conclusion, temperature does play a seasonal role in controlling hydrogen sulfide concentration in manholes, by causing more volatilization in summer.

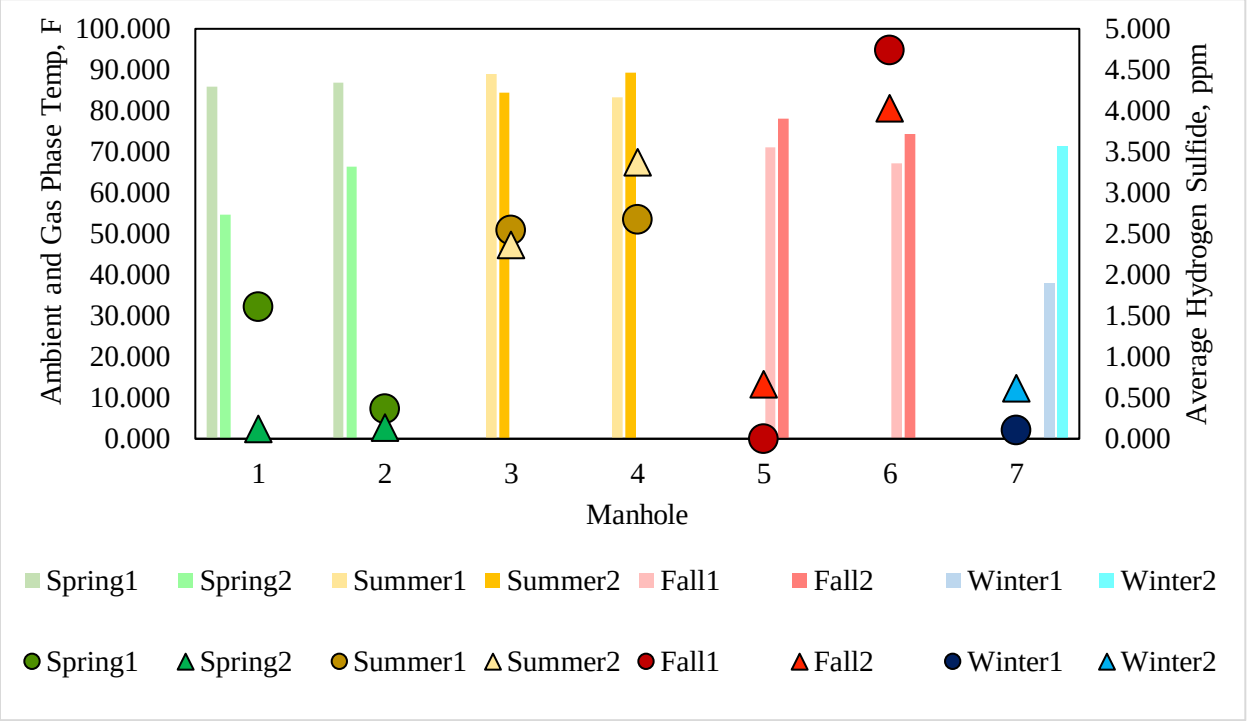


Fig B.8 Average ambient temperature vs average gas phase temperature for manholes in same seasons

B.4 EFFECT OF SEASONS ON GAS AND LIQUID PARAMETERS:

From above discussion seasonal effect on gas phase temperatures and hydrogen sulfide concentrations is clear, and seasonal pattern do effect sewer characteristics. In this section, we are going to analyze how seasons could impact gas and liquid parameters in the manhole. Fig B.9a to Fig B.9h shows the effect of seasons on various gas and liquid parameters for 7 different manholes done in different seasons.

In Fig B.9a we clearly see the effect of warmer temperatures on hydrogen sulfide concentrations, however colder temperatures of Fall and Winter do not show much association with ambient temperature, these results are similar to results discussed in previous section. Also, Table 4.7 shows the average hydrogen sulfide concentrations for 146 manholes with summer concentration of 3.53 ppm being highest followed by winter, with least recorded for Spring. The reason for winter average values being high could be because of design factors been a major factor such as presence of hydraulic jump, high drop.

In Fig B.9b and c we clearly see the effect of ambient temperatures on gas phase and liquid phase temperatures with higher temperatures reported in summer and lower gas and liquid phase temperatures reported in winter. We see higher gas and liquid phase temperatures in summer in comparison to winter. The spring and fall gas and liquid phase temperature values were similar to each other. Also, the liquid phase temperatures were higher than gas phase temperatures in all the seasons, because wastewater temperatures are usually warmer due to increased use of hot water by residential areas[41]. Table B.3 shows the seasonal trends for 146 manholes, and it can been seen that highest difference between the liquid and gas phase temperature was in winter and least in summer, a result of excessive hot water usage in winter in comparison to summer.

Table B.3 Liquid and gas phase temperatures variability in different seasons

Gas Phase Temperature				Liquid Temperature				Liquid - Gas Temp			
Spring	Summer	Fall	Winter	Spring	Summer	Fall	Winter	Spring	Summer	Fall	Winter
69.79	81.34	76.37	59.60	73.14	82.25	78.64	66.49	3.35	0.91	2.26	6.89
Average								3.35			

In Fig B.9d we see the effect of ambient temperatures on dissolved oxygen. The trend observed in dissolved oxygen is reverse of that observed for gas phase and liquid phase temperatures. For dissolved oxygen the higher concentrations were observed during spring followed by fall followed by winter and with least values during the summer. This trend could be because during summer there is an increase in microbial growth[37], thus faster oxygen consumption, however faster microbial growth also would result in higher sulfide generation which could explain the higher hydrogen sulfide values in summer (Fig B.9a), in comparison to fall and spring. Also, it has been reported that sulfide production increases 7% for every Celsius degree increase up to 40°C[39]. However in winter though we see low dissolved oxygen which could result in anerobic conditions, the reason for low hydrogen sulfide generation could be low gas and liquid phase temperatures (Fig B.9b and c) which increase hydrogen sulfide solubility[33, 34] and also lower temperatures slow microbial activity. Similar, trend is seen in Table 4.7, which shows average dissolved oxygen concentrations for 146 manholes with spring showing the highest dissolved oxygen average of 1.80 mg/l and lowest average of 0.66 mg/l for summer.

In Fig B.9e, we see the effect of ambient temperature on pH. The pH values mostly stay constant ranging between 6.2-8.2, which is neutral pH for wastewater[33,45, 56]. The high pH for manhole 7 could be due basic constituents in the wastewater.

From Fig B.9f to Fig B.9h, sulfide, sulfate, BOD data was not available, due to concerns while sample collection. So, the manhole 1 position is left blank in the 3 figures. In Fig B.9f we see the effect of ambient temperature on sulfide. The sulfide values are higher in fall, and lowest in summer. The reason for low sulfide concentration in liquid phase could be volatilization of sulfide from liquid to gas phase as hydrogen sulfide because as discussed above high temperature decrease sulfide solubility. Similarly, for spring manhole 2 and 3 showed good hydrogen sulfide concentrations in the manhole thus we see less sulfide in manhole 2 and 3, also dissolved oxygen was highest during the spring which could mean that the sulfide generated may have also been converted to sulfate. Also, winter has shown fairly good sulfide values, which seems reasonable considering during winter the dissolved oxygen concentration is less thus a good environment for anerobic condition, however due to low temperatures the build sulfide cannot escape to the manhole headspace thus higher concentration recorded in the liquid phase. The winter sulfide values are less than fall values may be because lower temperatures do slow down microbial activity, thus

there might be reduced microbial activity in winter in comparison to fall. Similar, trend is seen in Table 4.9, which shows average dissolved sulfide concentrations for 146 manholes with fall showing the highest dissolved sulfide average of 55.60 mg/l and lowest average of 39.01 mg/l for summer.

In Fig B.9g we see the effect of ambient temperature on sulfate. The sulfate values were highest during spring followed by fall, and least for summer. The higher sulfate values could be because spring manholes also recorded highest dissolved oxygen levels thus the produced sulfide is quickly converted to sulfate, thus a higher sulfate concentration is recorded for these manholes. Similarly, fall also recorded second highest dissolve oxygen levels, and second highest sulfate levels thus indicating a there might be an association between these parameters. Also, sulfate was lowest for the summer, which also correlates with lower dissolved oxygen levels observed in summer, which may be due to increased microbial activity due to high temperatures. Also, other factor for less sulfate in summer could be higher volatilization of hydrogen sulfide to the gas phase at higher temperatures. Similar, trend is seen in Table 4.7, which shows average dissolved sulfate concentrations for 146 manholes showing the highest dissolved sulfate average of 81.02 mg/l and lowest average of 58.12 mg/l for summer.

In Fig B.9h we see the effect of ambient temperature on BOD. Except for manhole 2, the BOD trends show an inverse relationship to the dissolved oxygen concentrations. For seasons in which dissolved oxygen was higher we saw a decreased BOD concentration. However, this trend was not seen for 146 manholes in Table 4.7. In Table 4.7, BOD trends followed the same trend as the liquid dissolved oxygen values with spring BOD been the highest and summer the lowest. This could imply that at higher dissolved oxygen, the sulfide if converted to sulfate thus consuming the liquid dissolved oxygen and showing high BOD. Also, it could mean that other factors such as microbial activity consuming dissolved oxygen for sulfide and sulfate production, or higher sulfur compounds in wastewater which might consume dissolved oxygen thus creating a biological oxygen demand and temperature variation in gas and liquid phase may be controlling the production and consumption of oxygen in manhole. Also, turbulence in manhole adds oxygen to wastewater which may result in unclear relationship between dissolved oxygen and biochemical oxygen demand.

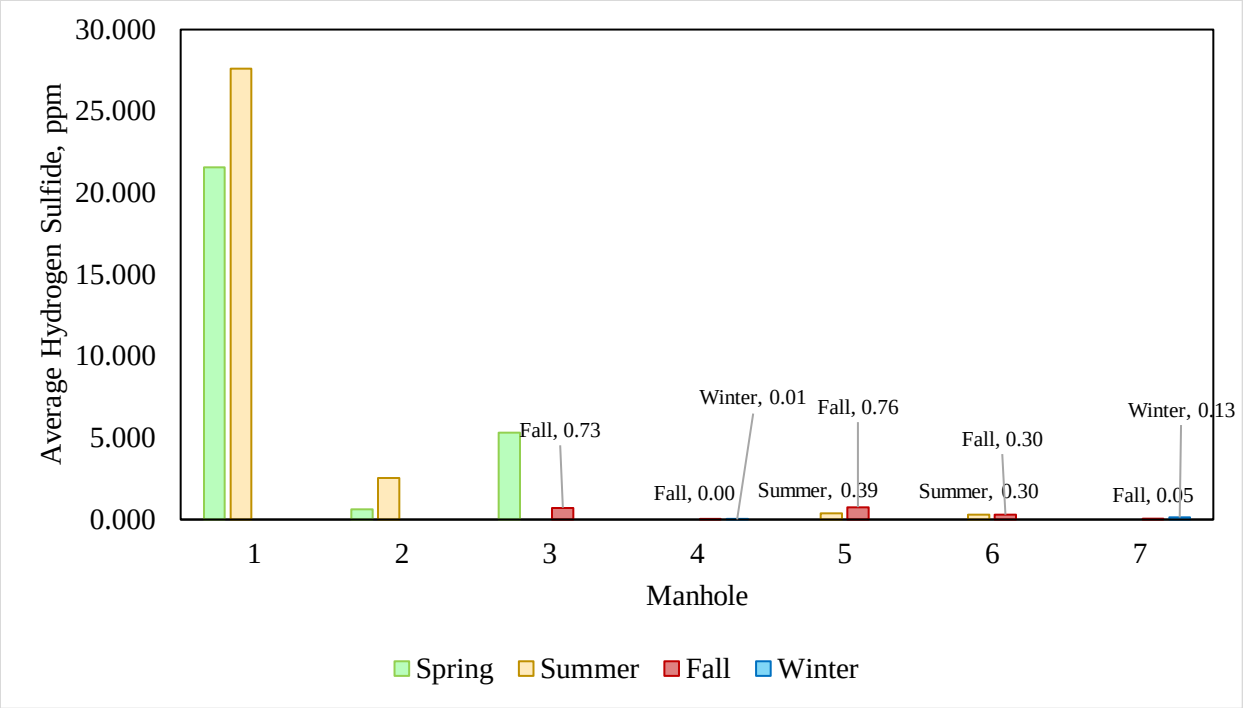
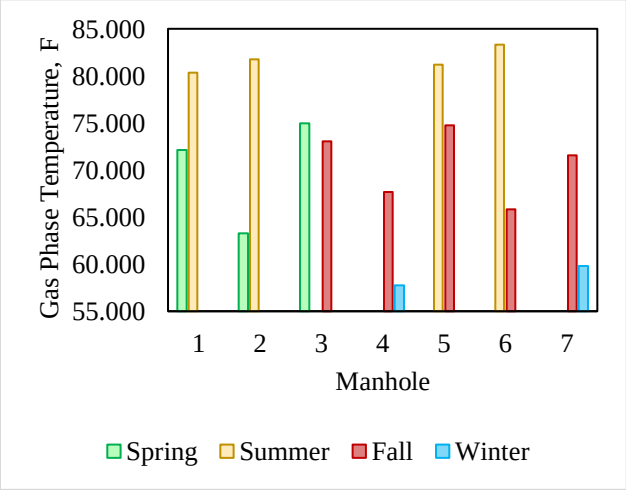


Fig B.9a Average hydrogen sulfide in different seasons



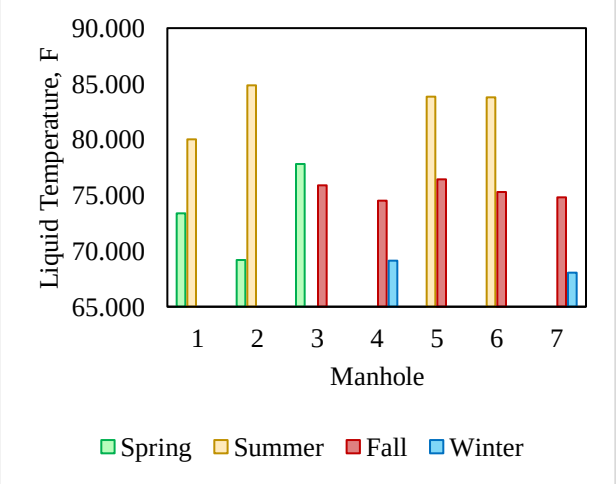


Fig B.9b Avg. gas phase temperature in different seasons

Fig B.9c Avg. liquid temperature in different seasons

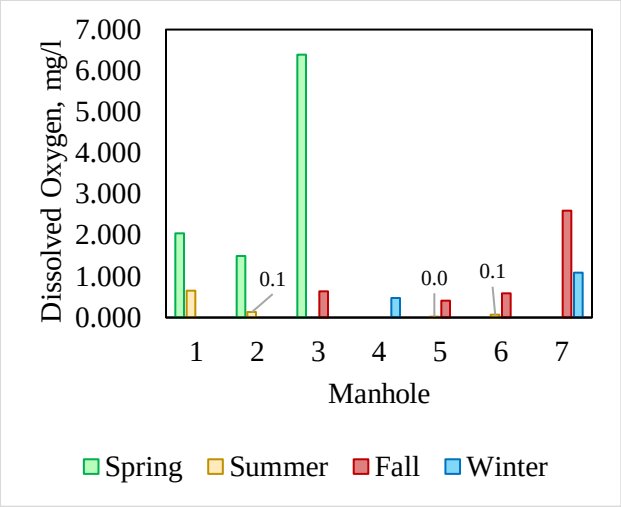


Fig B.9d Avg. DO in different seasons

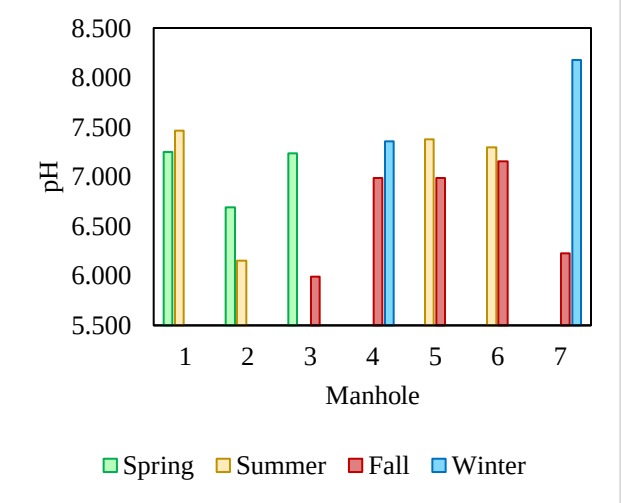


Fig B.9e Avg. pH in different seasons

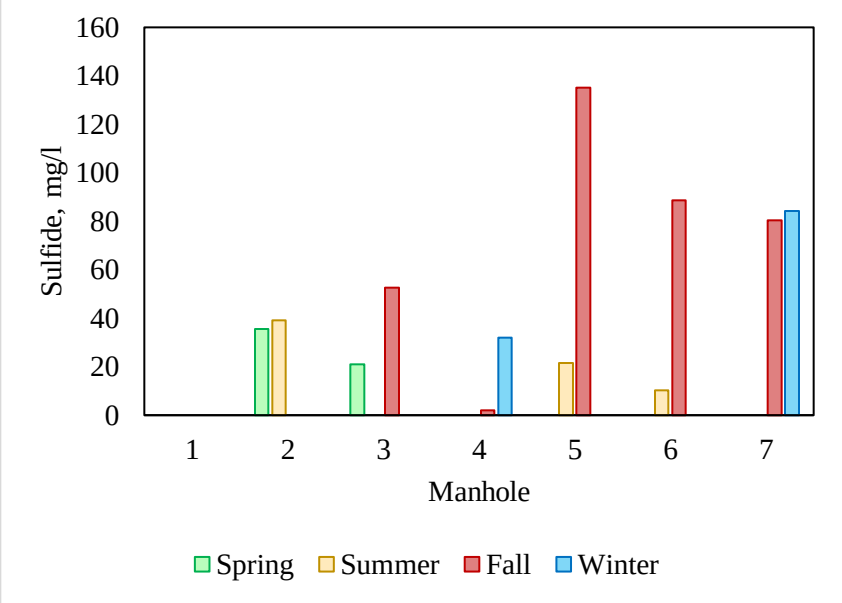


Fig B.9f Avg. sulfide in different seasons

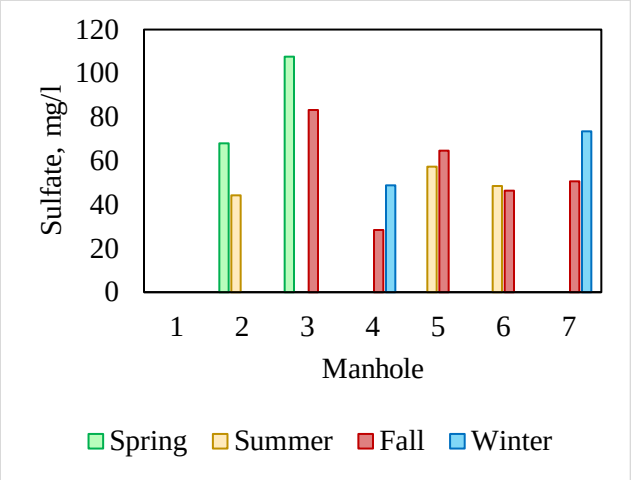


Fig B.9g Avg. sulfate in different seasons

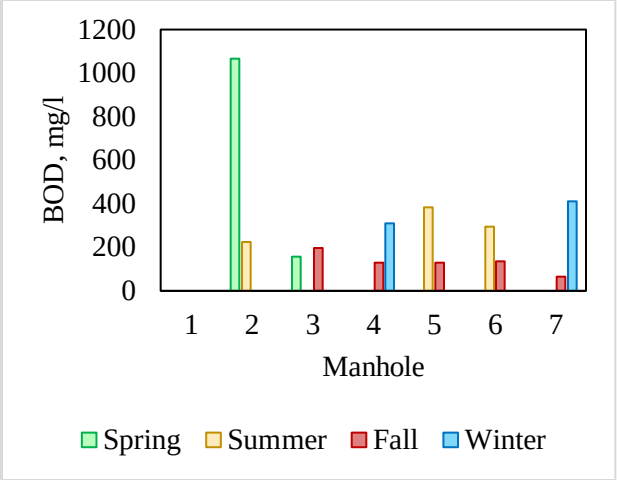


Fig B.9h Avg. BOD in different seasons

In summary we can say that,

a. There is a direct relationship between ambient temperatures and gas phase and liquid phase temperatures in the manhole.
b. The manhole does exhibit both diurnal and seasonal patterns.
c. Relationship between hydrogen sulfide and ambient temperature is more prominent at higher temperatures greater than 70 F rather than at lower temperatures lower than 65 F.
d. There is an average of 3.35 F difference between the gas phase and liquid phase temperatures.
e. For a 1 F change in ambient temperature the gas phase temperature changes by 0.38 F.
f. During higher temperatures (summer), becomes a major determining factor for hydrogen sulfide concentration in manhole, whereas at lower temperatures such as design, dissolved oxygen could play a major role in hydrogen sulfide generation.
g. Since summer results in increased gas and liquid phase temperatures thus reducing hydrogen sulfide solubility in the manhole, and it also increases microbial activity, this could mean that manholes are more vulnerable to corrosion in summer as the volatized hydrogen sulfide can be converted to sulfuric acid to degrade concrete surface.
h. During winter though dissolved oxygen is low, the low winter temperatures increase hydrogen sulfide solubility thus less sulfide available for volatilization thus less available for corrosion. Also, during winter

the microbial activity is usually slow, thus further slowing the production and volatilization of hydrogen sulfide.

BIOGRAPHICAL INFORMATION

Sunakshi Hada received her Ph.D. in Environmental Engineering from the University of Texas at Arlington. Her research focused on modeling hydrogen sulfide generation and volatilization, and rates of corrosion in manhole shafts, with an aim at reducing infrastructure damage. She earned her Master of Science degree in Environmental and Earth Science from University of Texas at Arlington, and Bachelor of Technology degree in Biotechnology from Anna University, India. For her master's project, she worked on predicting methane emission rates occurring in developing countries landfill using the CLEEN model. For her bachelor's final year project, she won an Innovative Project iQuest award from the Central Leather Research Institute in India. The Bachelor's final year project dealt with phytoremediation of tannery effluent discharging from leather industries to improve the local water quality.

Dr. Hada previously worked for the Natural Resources Management (NRM) Division at the Texas Department of Transportation (TxDOT), in Austin, Texas, and TATA Consultancy Services, India. Her research has won her many honors from Central Leather Research Institute in India, City of Arlington - Texas, Air and Waste Management Association (AWMA) Stan Curry Memorial Scholarship. She is a member of AWMA and the Solid Waste Association of North America (SWANA). Her research interests include wastewater collection systems and effluent clean up, landfill emissions and leachate treatment, toxic emissions in wastewater collection systems, and bioremediation of contaminated systems.