Project 2523 | April 2026

SJSU SAN JOSÉ STATE UNIVERSITY

MTI MINETA TRANSPORTATION INSTITUTE

Steel Corrosion in Underground Transportation Infrastructure

Amr M. Morsy, PhD, PE Islam A. Ebo Shiv K. Janardhanan

CSUTC California State University Transportation Consortium

CALIFORNIA STATE UNIVERSITY LONG BEACH

MINETA TRANSPORTATION INSTITUTE

Founded in 1991, the Mineta Transportation Institute (MTI), an organized research and training unit in partnership with the Lucas College and Graduate School of Business at San José State University (SJSU), increases mobility for all by improving the safety, efficiency, accessibility, and convenience of our nation's transportation system. Through research, education, workforce development, and technology transfer, we help create a connected world. MTI leads the California State University Transportation Consortium (CSUTC) funded by the State of California through Senate Bill 1 and the Climate Change and Extreme Events Training and Research (CCEETR) Program funded by the Federal Railroad Administration. MTI focuses on three primary responsibilities:

Research

MTI conducts multi-disciplinary research focused on surface transportation that contributes to effective decision making. Research areas include: active transportation; planning and policy; security and counterterrorism; sustainable transportation and land use; transit and passenger rail; transportation engineering; transportation finance; transportation technology; and workforce and labor. MTI research publications undergo expert peer review to ensure the quality of the research.

Education and Workforce Development

To ensure the efficient movement of people and goods, we must prepare the next generation of skilled transportation professionals who can lead a thriving, forward-thinking transportation industry for a more connected world. To help achieve this, MTI sponsors a suite of workforce development and education opportunities. The Institute supports educational programs offered by the Lucas Graduate School of Business: a Master of Science in Transportation Management, plus graduate certificates that include High-Speed and Intercity Rail Management and Transportation Security Management. These flexible programs offer live online classes so that working transportation professionals can pursue an advanced degree regardless of their location.

Information and Technology Transfer

MTI utilizes a diverse array of dissemination methods and media to ensure research results reach those responsible for managing change. These methods include publication, seminars, workshops, websites, social media, webinars, and other technology transfer mechanisms. Additionally, MTI promotes the availability of completed research to professional organizations and works to integrate the research findings into the graduate education program. MTI's extensive collection of transportation-related publications is integrated into San José State University's world-class Martin Luther King, Jr. Library.

Report 26-10

Steel Corrosion in Underground Transportation Infrastructure

Amr M. Morsy, PhD, PE

Islam A. Ebo

Shiv K. Janardhanan

April 2026

A publication of the
Mineta Transportation Institute
Created by Congress in 1991

College of Business
San José State University
San José, CA 95192-0219

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. 26-10	**2. Government Accession No.**	**3. Recipient's Catalog No.**	
4. Title and Subtitle Steel Corrosion in Underground Transportation Infrastructure		**5. Report Date** April 2026	
		6. Performing Organization Code	
7. Authors Amr M. Morsy, PhD, PE ORCID: 0000-0002-9335-7847 Islam A. Ebo ORCID: 0009-0007-6797-8002 Shiv K. Janardhanan ORCID: 0009-0002-2213-5195		**8. Performing Organization Report** CA-MTI-2523	
9. Performing Organization Name and Address Mineta Transportation Institute College of Business San José State University San José, CA 95192-0219		**10. Work Unit No.**	
		11. Contract or Grant No. SB1-SJAUX_2023-26	
12. Sponsoring Agency Name and Address State of California SB1 2017/2018 Trustees of the California State University Sponsored Programs Administration 401 Golden Shore, 5th Floor Long Beach, CA 90802		**13. Type of Report and Period Covered**	
		14. Sponsoring Agency Code	
15. Supplemental Notes 10.31979/mti.2026.2523			

16. Abstract

Corrosion of buried steel, a critical component of American transportation infrastructure, remains one of the most insidious challenges due to the uncertainty associated with its estimates. Predicting when and how this corrosion happens is very difficult. This uncertainty grows exponentially with time, making corrosion estimation in the long term even more challenging, especially with buried steel and steel structures, which cannot even be monitored visually. While significant advancement has been made to understand the effect of the various corrosion parameters on soil corrosivity, there is a lack of a comprehensive understanding of how these factors collectively contribute to corrosion as they vary simultaneously and continuously with time. This project evaluates soil resistivity and corrosivity in controlled, constant conditions, considering the various key parameters that contribute to corrosion of buried steel. The project involved devising a new experimental protocol, developing and implementing a comprehensive experimental program by varying one testing parameter at a time. The results of the testing program showed the potential of the experimental approach to provide the necessary data to develop empirical prediction models for soil resistivity. Additionally, a new experimental method was piloted for this project to capture the variation in soil resistivity in a continuously varying environment. Finally, the researchers compiled a large digital database of real-world corrosion measurements and site information. Using advanced data analysis techniques, they created a model that can help predict corrosion in buried steel structures and estimate the level of uncertainty in those predictions. Better predictions of corrosion can help engineers and infrastructure managers identify risks earlier, plan maintenance more effectively, and extend the life of critical infrastructure such as pipelines, bridges, and transportation systems, therefore helping to reduce costly failures, improve safety, and support more reliable systems.

17. Key Words Deterioration, corrosion, wear, aging, asset management.	**18. Distribution Statement** No restrictions. This document is available to the public through The National Technical Information Service, Springfield, VA 22161.		
19. Security Classif. (of this report) Unclassified	**20. Security Classif. (of this page)** Unclassified	**21. No. of Pages115** 53	**22. Price**

Form DOT F 1700.7 (8-72)

DOI: 10.31979/mti.2026.2523

Mineta Transportation Institute
College of Business
San José State University
San José, CA 95192-0219

Tel: (408) 924-7560
Fax: (408) 924-7565
Email: mineta-institute@sjsu.edu

transweb.sjsu.edu/research/2523

ACKNOWLEDGMENTS

This study was supported by the California State University Transportation Consortium (CSUTC) through California Senate Bill 1 (SB 1): The Road Repair and Accountability Act of 2017. The opinions presented in this report are those of the author and are not necessarily those of the CSUTC. Amr Morsy recognizes the support of Dr. Hilary Nixon of the Mineta Transportation Institute (MTI). The authors thank Lisa Rose and Editing Press for editorial services, as well as MTI staff Project Assistant, Rajeshwari Rajesh, and Graphic Design Assistant, Katerina Earnest. Report cover image is by AdobeStock/evgavrilov®.

Contents

LIST OF FIGURES

LIST OF TABLES

Executive Summary

Corrosion imposes a significant threat to our aging transportation infrastructure, including bridges, highways, railroads, and a multitude of subsurface structures such as metallic culverts, drainage pipelines, reinforced soil structures, and metallic deep foundation systems that support our transportation network. The annual cost of corrosion for highway bridges in the US is estimated to be between $6.3 and $10.15 billion. Corrosion of buried steel remains one of the most insidious challenges due to the uncertainty associated with its estimates. This uncertainty grows exponentially with time, making corrosion estimates in the long-term challenging for transportation asset management, especially since buried steel structures cannot be inspected visually.

Recently, the National Academies of Sciences, Engineering, and Medicine (NASEM) published a report in 2023 titled "*Corrosion of Buried Steel at New and In-Service Infrastructure,*" documenting the research gap in modeling and estimating corrosion of buried steel. The report emphasizes the urgent need for conducting further investigations to improve the current outdated corrosion models, which were developed primarily using data from the 1950s. Improving our understanding of long-term corrosion and our ability to quantify such corrosion based on the continuously varying subsurface environment will be of great assistance to engineers in evaluating the service life of existing underground steel transportation structures, as well as in the future design of new structures.

While significant advancement has been made to understand the effect of the various corrosion parameters on soil corrosivity, there is a lack of a comprehensive understanding of how these factors collectively contribute to corrosion as they vary simultaneously and continuously with time (Morsy & Ebo, 2025). This research aimed to evaluate soil resistivity and corrosivity, considering the various key parameters that contribute to corrosion of buried steel. Three studies were completed and documented in this report:

- Element-Scale Soil Resistivity Testing in Controlled Constant Conditions. This is an experimental study that aimed to evaluate soil resistivity in controlled, constant conditions. This study involved devising a new experimental protocol that was evaluated for its ability to produce repeatable results. An experimental program was devised and implemented to evaluate the effect of testing parameters known to be the most relevant in soil corrosivity by varying one parameter at a time. The results of the testing program showed the potential of the experimental approach to provide the necessary data to develop empirical prediction models for soil resistivity. The influence of each parameter was evaluated separately. Soil resistivity was observed to decrease with increasing soil moisture, chloride content, sulfate content, acidity, alkalinity, and soil temperature.

- Element-Scale Soil Resistivity Testing in Controlled Variable Conditions. This is a pilot experimental study that aimed to evaluate soil resistivity in controlled, variable conditions. This study involved devising and piloting a new experimental method capable of capturing the variation in soil resistivity in a continuously varying environment. The results indicated great potential for the proposed experimental approach to provide data necessary to calibrate soil resistivity prediction models in a continuously varying environment representative of field conditions.

- Analysis of the NBS Romanoff Corrosion Dataset. This is a data analysis study that aimed to evaluate the corrosion of buried steel using advanced data science methods. This study presented a comprehensive framework for predicting underground corrosion of uncoated ferrous materials using a supervised machine learning approach based on the ExtraTrees algorithm. Through data curation from the NBS Romanoff (1957) corrosion program and structured model validation using site-exclusive nested cross-validation, the model achieved consistent performance transferable to unseen sites.

1. Introduction and Background

The annual costs associated with metallic corrosion in the US are estimated to range from 3 to 4% of the GDP and account for $2.5 trillion of the global GDP (Koch, 2017). The annual cost of corrosion for highway bridges in the US is estimated to be between $6.3 and $10.15 billion, as per the American Galvanizing Association (AGA). Replacing corroded steel can be quite costly; for example, the Los Angeles Department of Water and Power developed a $1.34 billion plan to replace 435 miles of deteriorating pipelines over 10 years (NASEM, 2023, citing LADWP Water Infrastructure Plan). According to Caltrans Transportation Asset Management (TAMP, 2022), one of the main reasons for drainage system replacement is deterioration, typically because of corrosion among other reasons.

1.1 Overview of the Problem

Corrosion of buried steel remains one of the most insidious challenges due to the uncertainty associated with its estimates. This uncertainty grows exponentially with time, making corrosion estimation in the long term quite challenging, especially with buried steel and steel structures which cannot even be monitored visually. Improving our understanding of long-term corrosion and improving our ability to quantify such corrosion will be of great assistance to engineers in evaluating the service life of existing underground steel structures, as well as in the future design of new structures. A vast number of transportation infrastructure features will benefit from the outcomes of this research, including reinforced soil retaining walls (e.g., mechanically stabilized earth walls, soil nail walls, anchored walls, sheet-pile walls), steel deep foundations (e.g., H-piles, pipe piles, sheet piles), steel culverts, steel pipelines, underground reservoirs, and several other underground ancillary metallic structures.

Although time is not itself a corrosive parameter, it has been widely used to quantify cumulative corrosion of metals buried in corrosive environments. Cumulative corrosion of buried steel has typically been estimated as a function of time using a power law, as follows: $X = K.t^m$, where X is the average corrosion thickness loss, K is a coefficient that depends on soil corrosivity, t is the time, and m is a time exponent (Romanoff, 1956; Bastick & Jallioux, 1992; Elias et al., 2009; Fishman &Withiam, 2011; Fishman et al., 2021; NASEM, 2023). A number of predictive models were developed to evaluate the corrosion of buried steel (Stuttgart, 1980; Darbin, 1988; Elias et al., 1990; Fishman & Withiam, 2011). These models were developed to be used in the design of new structures and are deemed overly conservative to account for the uncertainties associated with corrosion and its progression with time. However, it may be impractical to use such overly conservative models for the evaluation of reinforcement corrosion in existing, aging structures for asset management purposes. Moreover, current corrosion models were developed to provide predictions of reinforcement corrosion as a function of time for fills having characteristics that are deemed mildly corrosive and drainable.

While significant advancement has been made to understand the effect of the various corrosion parameters on soil corrosivity, there is a lack of a comprehensive understanding of how these factors collectively contribute to corrosion as they vary simultaneously and continuously with time (Morsy & Ebo, 2025). Through unique and comprehensive experimental and analytical approaches, this study aimed to assess the corrosion of steel structures buried in soil with a focus on steel structures integral to transportation infrastructure, such as reinforced soil retaining walls (*e.g.*, mechanically stabilized earth walls, soil nail walls, sheet-pile walls), steel deep foundations (*e.g.*, H-piles and pipe piles), steel culverts, steel pipelines, and underground reservoirs.

1.2 Research Objectives

This research aimed to evaluate soil resistivity and corrosivity, considering the various key parameters that contribute to corrosion of buried steel. The research objectives included the following:

- Evaluating soil resistivity in controlled, constant conditions. This study involved devising a new experimental protocol, developing and implementing a comprehensive experimental program by varying one testing parameter at a time. This approach allowed evaluating the influence of each parameter on soil resistivity separately.

- Evaluating soil resistivity in controlled, variable conditions. This study involved devising and piloting a new experimental method capable of capturing the variation in soil resistivity in a continuously varying environment.

- Evaluating corrosion of buried steel using advanced data science methods. This study involved compiling and curating a digital database of field corrosion measurements and site parameters. The database was used to devise a prediction model capable of predicting corrosion of underground steel with a prescribed uncertainty tolerance.

1.3 Report Organization

This report consists of five sections:

- Section 1. Introduction and Background. This section presents an overview of the problem and research questions, research aim and objectives, and report organization.

- Section 2. Element-Scale Soil Resistivity Testing in Controlled Constant Conditions. This section presents an experimental study aimed at evaluating the soil resistivity in controlled, constant conditions through a comprehensive experimental program by varying one testing parameter at a time.

- Section 3. Element-Scale Soil Resistivity Testing in Controlled Variable Conditions. This section presents a pilot experimental study aimed at evaluating the soil resistivity in controlled, variable conditions through a new experimental method capable of capturing the variation in soil resistivity in a continuously varying environment.

- Section 4. Analysis of the NBS Romanoff Corrosion Dataset. This section presents a data analysis study aimed at evaluating the corrosion of buried steel using advanced data science methods.

- Section 5. Summary and Conclusions. This section presents a summary of the completed studies and the conclusions drawn from them.

2. Element-Scale Soil Resistivity Testing in Controlled Constant Conditions

An experimental study was conducted to evaluate soil resistivity in controlled, constant conditions. This study involved devising a new experimental protocol, developing and implementing a comprehensive experimental program by varying one testing parameter at a time.

2.1 Testing Parameters

Corrosion rates are governed by the electrical conductivity of the soil in which the metallic structures are buried (Nicks et al., 2017; Bronson et al., 2020; Kolay et al., 2020). Electrical conductivity, which is the inverse of resistivity, primarily depends on soil moisture, concentration of soluble salts (chlorides and sulfates), and pH value (Elias et al., 2009). Soil resistivity has a direct relationship with the electrical currents generated during the corrosion reactions. Soil resistivity can be evaluated in the lab through an index test that measures the minimum soil resistivity according to an AASHTO test standard (AASHTO T 288). The minimum soil resistivity obtained experimentally is indicative of the soil corrosivity and can be used as a metric, along with other metrics, to evaluate soil corrosivity in practice. The key factors governing soil resistivity (and soil corrosivity) can be summarized as follows:

2.1.1 Soil Moisture

Soil moisture is a key element for corrosion of buried metals as it forms the electrolyte necessary for the corrosion-related chemical reactions to take place. Corrosion rate increases with increasing soil moisture content up to a critical level, beyond which corrosion rate decreases with further increasing moisture content (Gupta & Gupta, 1979; Noor & Al-Moubarki, 2014; Wasim et al., 2018; Ezuber et al., 2021; Liu et al., 2023) due to the decrease in the air content and the oxygen supply necessary for corrosion to take place. The critical moisture content has been reported in the literature to range from 10 to 30% based on several studies that involved a range of soils. Gupta and Gupta (1979) reported that the critical soil moisture content can be approximated to 65% of the soil field capacity (i.e., water holding capacity).

2.1.2 Chlorides and Sulfate Contents

Corrosion rates increase with increasing soil salts content (Romanoff, 1956; Bastick & Jailloux, 1992). Ions of chloride and sulfate salts are particularly classified as aggressively corrosive ions. Chlorides destroy the protective rust-stable film of metals, exposing the metal surface to further corrosion reactions. Sulfates are considered to be less corrosive compared to chlorides unless anaerobic sulfate-reducing bacteria exist. Bastick and Jailloux (1992) reported that chlorides and sulfates have a similar effect on corrosion rates up to concentrations of 100 ppm. For

concentrations higher than 100 ppm, the effect of the chlorides on the corrosion rate becomes much higher than that of sulfates.

2.1.3 pH Value

Corrosion rates decrease as pH increases in acidic environments and increase as pH increases in alkaline environments (Rossum, 1969; Mughabghab & Sullivan, 1989). Corrosion progression is generally high in neutral and acidic environments (Shreir et al., 1994; Roberge, 2000). Corrosion progression slows down at relatively high pH values, where metal oxides are stable because the spontaneous formation of oxides can provide protection (in the form of rust scales) to the underlying metal (NASEM, 2023).

2.1.4 Soil Temperature

Corrosion rates generally increase with increasing temperature up to a critical level, beyond which corrosion rates decrease with increasing temperature due to the decrease in oxygen solubility in water (Davalos et al., 1992; NASEM, 2023). Maximum corrosion rates have been reported to typically occur at 70°C (NASEM, 2023).

2.2 Testing Materials

The materials used in this experimental program include clean, uniform sand, deionized water, salts, and pH buffer solutions. All materials were processed and handled in a manner that ensured consistent test conditions across the experimental program.

2.2.1 Soil Environment

The soil used in the experimental program was a commercial standard sand that conforms to ASTM C33. This sand was selected for its insignificant fines content, uniform gradation, and homogeneity. Additionally, this sand is consistent with typical sands used as select fills in many infrastructure applications. The soil has a specific gravity, G_S, of 2.62 (ASTM D854), a maximum void ratio, e_{max}, of 0.728 (ASTM 4254 Method A), and a minimum void ratio, e_{min}, of 0.437 estimated based on a correlation with e_{max} for comparable sands (Cubrinovski & Ishihara, 2002). The maximum dry unit weight of the soil is 18.1 kN/m^3 per Standard Proctor test (ASTM D698), which corresponds to an optimum moisture content of 13.7% (equivalent to a degree of saturation of 85%).

Before use, the sand was processed to minimize the impact of its minerals on the results of the soil resistivity tests. Although the fines content in this sand is insignificant, the sand was washed on a No. 200 sieve to remove any fine particles. The sand was then soaked overnight and rinsed repeatedly using deionized water for several days to reduce any native minerals. The concentrations of chlorides and sulfates in the soaking water were measured every day after soaking and before

rinsing using photometers. For the soil used in this study, these concentrations of chlorides and sulfates dropped to insignificant levels after 7 days of repeated soaking and rinsing. The concentrations of chlorides and sulfates from the native soil minerals dropped from 21 mg/L to 3.5 mg/L for chlorides and from 39 mg/L to 1.0 mg/L for sulfates. Notably, the accuracy of the photometers used to measure chlorides and sulfates was ±0.5 mg/L (±6% of reading) and ±5 mg/L (±3% of reading), respectively. The cleaned sand was then oven-dried and stored in sealed containers.

2.2.2 Deionized Water

The water used to hydrate the soil to the desired degree of saturation before each test was conditioned to a target specific porewater chemistry. This was achieved by using deionized water conditioned to target pH value and concentrations of chlorides and sulfates. The deionized water was also used as the baseline porewater in the baseline control testing series to evaluate the influence of varying each corrosion parameter independently.

2.2.3 Chloride and Sulfate Salts

Sodium Chloride (NaCl) and Sodium Sulfate (Na_2SO_4) food-grade salts were used to control the concentrations of chlorides (Cl^-) and sulfates (SO_4^{2-}) in the porewater. Salt dosages were measured using a precision scale with a resolution of 0.0001 g. Porewater solutions were prepared at varied concentrations as per the designed testing matrix.

2.2.4 pH Buffers

Laboratory-grade nitric acid (HNO_3) and caustic soda (NaOH) were used as buffer solutions to offset the pH value of the porewater to varied levels of acidity and alkalinity. The pH values of the prepared solutions were verified using a soil pH tester with ±0.05-pH accuracy and pH indicator papers. Porewater solutions were prepared at varying pH levels, ranging from pH 1.15 to pH 14.00, as per the designed testing matrix.

2.3 Experimental Approach

Based on the selected parameters and ranges, an experimental program was developed using a one-factor-at-a-time (OFAT) approach to evaluate the effect of each testing parameter over its practical range with a minimum number of experiments. This approach facilitates conducting experiments necessary to derive an empirical relationship between corrosion of buried steel and corrosive environment variables (i.e., corrosion parameters). The testing process involved three main stages: sample preparation, resistivity measurement, and data curation.

2.3.1 Sample Preparation

The sample preparation for the soil resistivity testing involved mixing a conditioned porewater solution with an oven-dried soil sample to a target degree of saturation. To account for the variability in sample preparation, the soil water content was measured after mixing. This allowed record-keeping of the actual water content of each tested sample in the experimental program, which could have deviated slightly from the target water content during sample mixing.

The sample was then placed in the soil resistivity box and compacted to a target dry unit relative density. The target relative density was 40%, which corresponds to a dry unit weight of 16 kN/m^3 (equivalent to 88% of the maximum dry unit by the Standard Proctor test). The corresponding target void ratio was 0.61. To account for the variability in sample preparation, the unit weight was measured after sample compaction. This allowed record-keeping of the actual relative density of each tested sample in the experimental program, which could have slightly deviated from the target relative density during sample compaction. The void ratio of the samples prepared in this experimental program was maintained between 0.60 and 0.62 (i.e., target void ratio 0.61±0.01).

After each soil resistivity test, the soil temperature of the sample was recorded. The average soil temperature across the tests in the experimental program was 20 ˚C, except for the testing series where soil temperature was varied.

2.3.2 Soil Resistivity Measurement

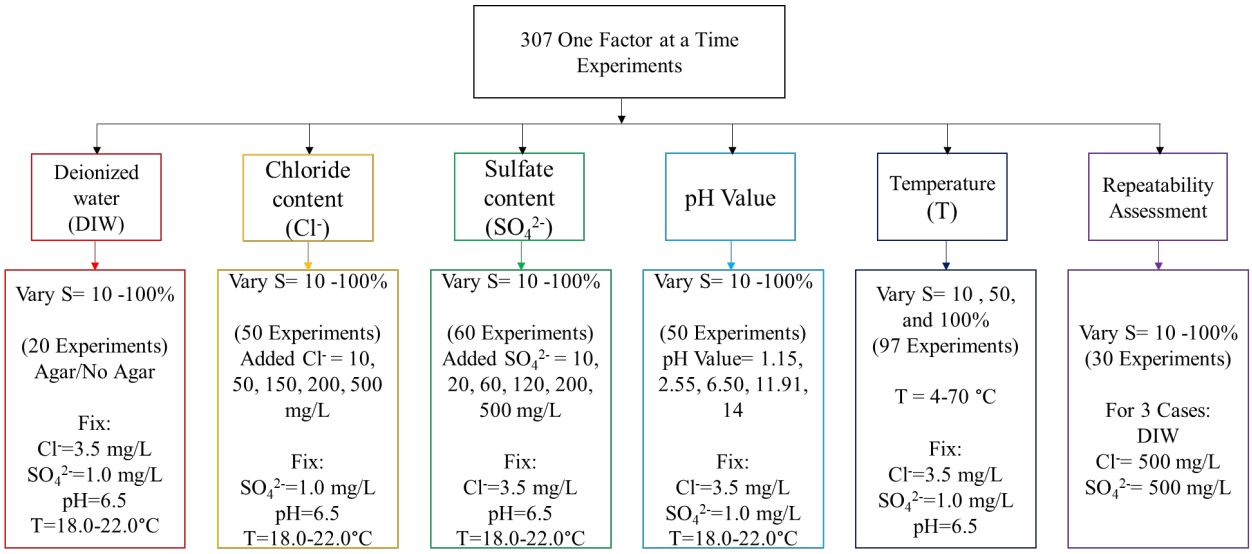
Soil resistivity was measured using a benchtop resistivity meter based on the Wenner four-electrode method (ASTM G57). The soil box that contained the soil and the four electrodes was 2.7×10^5 mm^3 in volume with a cross-sectional area of 1280 mm^2 (40 mm by 32 mm). The voltage pin separation length was 128 mm.

2.4 Experimental Program

The testing matrix was developed to ensure a systematic variation of key corrosion parameters: soil moisture content, pH level, chloride and sulfate concentrations, and soil temperature. The experiments presented in this section were conducted at controlled, constant conditions where all corrosion parameters were held constant throughout the testing duration. For each test, the soil resistivity was measured using a four-electrode measurement system. A total of 307 tests were completed to evaluate the effect of each testing parameter on soil resistivity.

The moisture content was varied from 2% to 24% (gravimetric basis), chloride concentrations ranged from 0 to 500 mg/L, and sulfate concentrations ranged from 0 to 500 mg/L. pH levels were adjusted between 1.15 and 14.00, while temperature was monitored across 4 ˚C to 70 ˚C. The full sequence of sample preparation, parameter control, and testing procedures is summarized in Figure 1, outlining the workflow followed across all experimental series.

Figure 1. Soil Resistivity Experimental Program



2.5 Results and Analysis

The results from the experimental program are presented and analyzed in six separate series. Each series isolates a single parameter while holding all other conditions constant to evaluate the direct influence of the varied parameter on the soil resistivity.

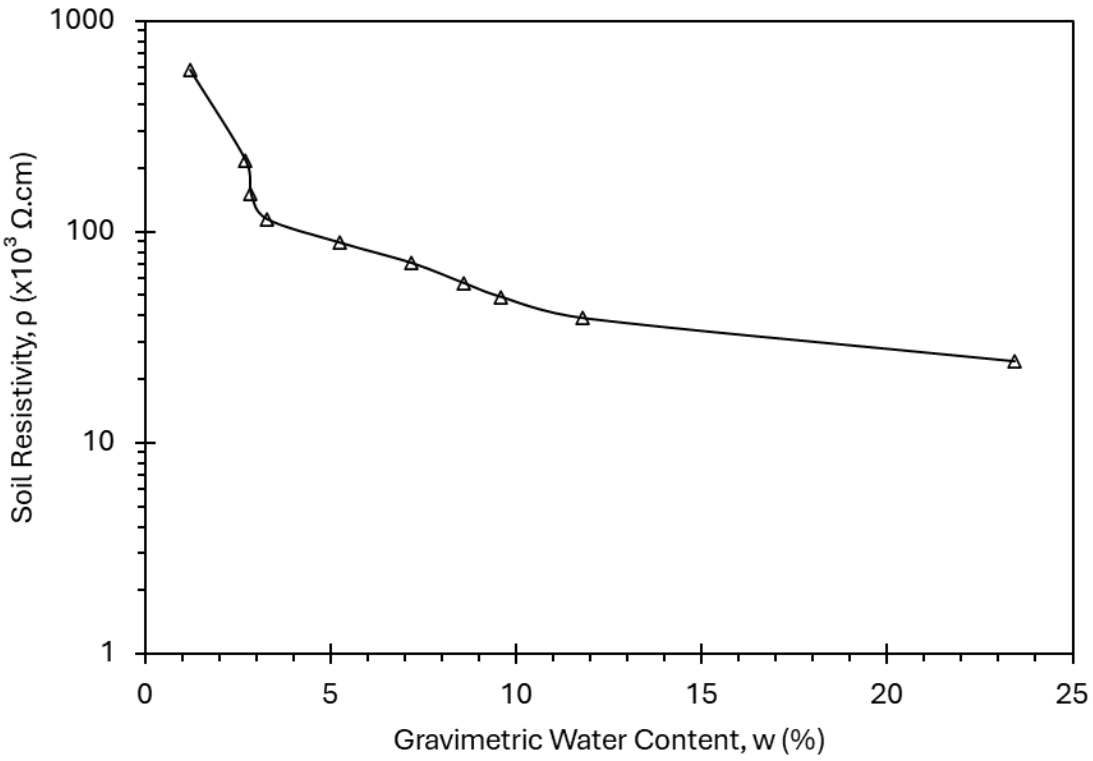
2.5.1 Repeatability Assessment Series

To evaluate the repeatability of the experimental procedure, three testing sets (a total of 30 soil resistivity tests) were repeated. The results indicated good repeatability, confirming the suitability of the devised experimental procedure for conducting the experimental program involved in this study.

2.5.2 Baseline Series

In the baseline series, deionized water was used to condition the soil at varied degrees of saturation. The results indicate an inverse relationship with gravimetric moisture content, as shown in Figure 2. The sensitivity of soil resistivity to water content is much higher in drier soils (film and capillary water) than in wetter soils.

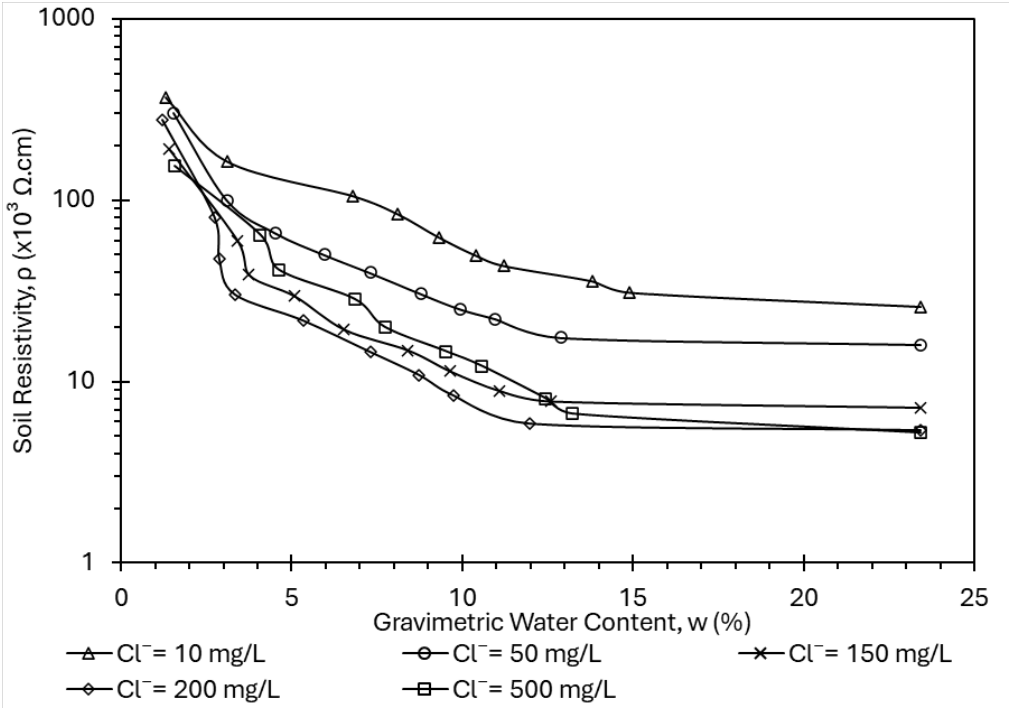
Figure 2. Baseline Soil Resistivity vs. Gravimetric Water Content



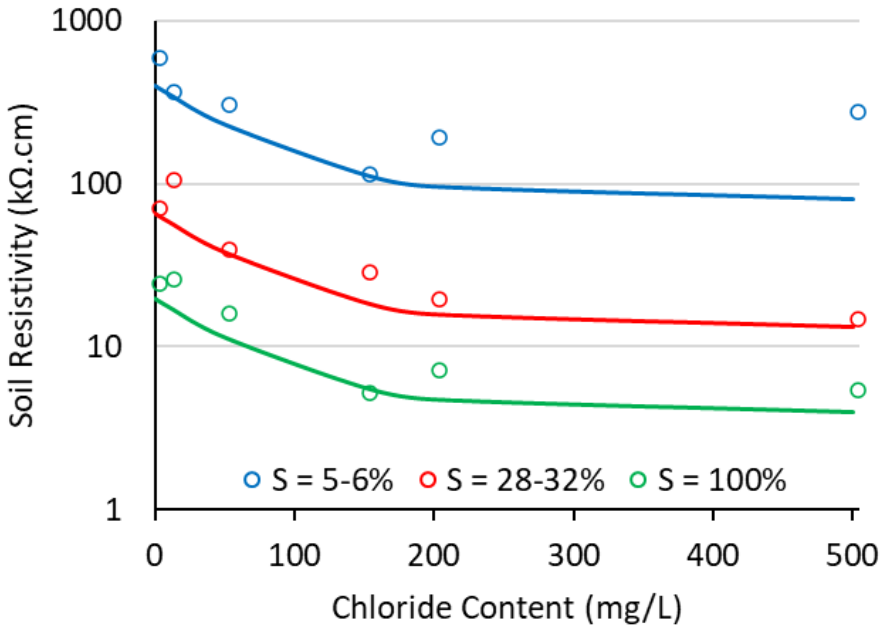
2.5.3 Chloride Content Series

As shown in Figure 3, increasing NaCl concentration resulted in a notable reduction in soil resistivity at all moisture levels. Soil resistivity exhibits a nonlinear decline with increasing chloride concentration.

Figure 3. Effect of Chloride Content on Soil Resistivity: (a) Soil Resistivity vs. Gravimetric Water Content at Varied Chloride Contents; and (b) Soil Resistivity vs. Chloride Content at Varied Saturation Degrees



(a)



(b)

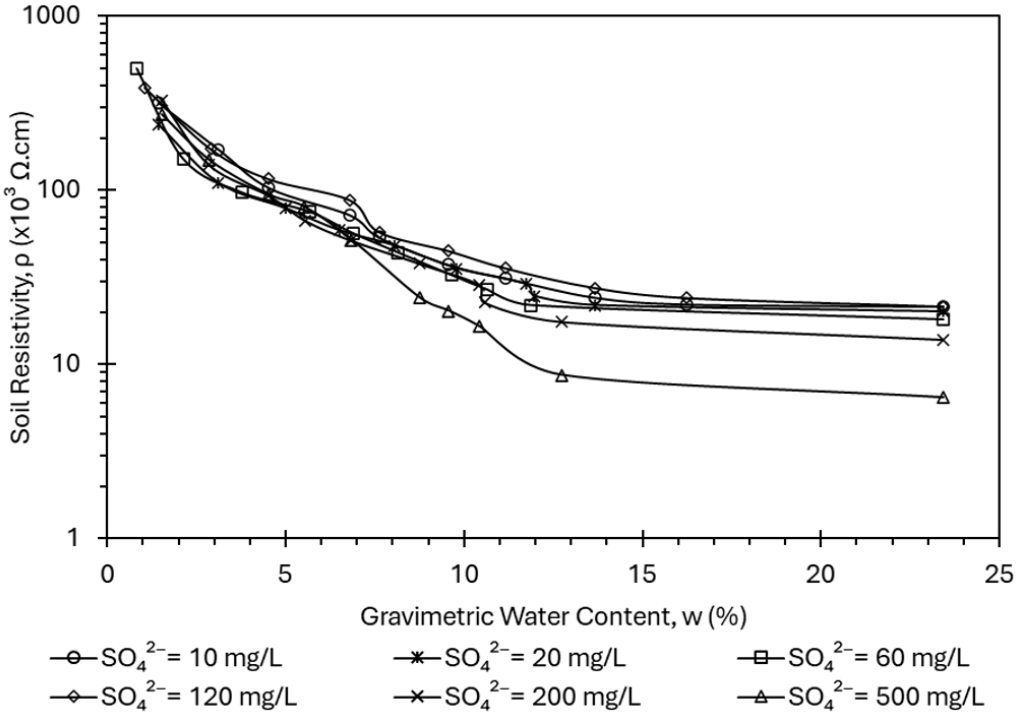
2.5.4 Sulfate Content Series

As presented in Figure 4, increasing Na_2SO_4 concentration resulted in a reduction in soil resistivity at all moisture levels. Soil resistivity exhibits a nonlinear decline with increasing sulfate concentration. The effect of the sulfate concentration on soil resistivity was comparable to that of the chloride concentration.

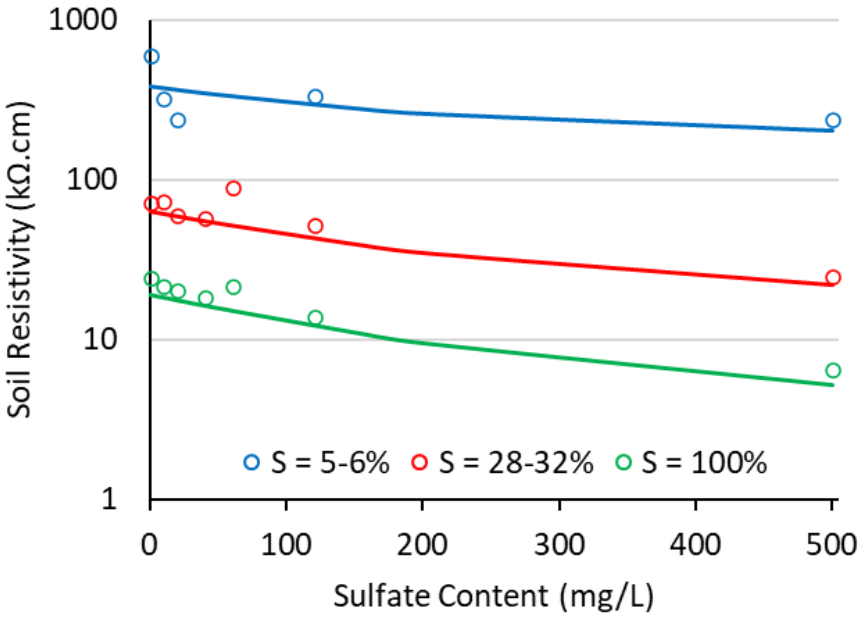
2.5.5 pH Value Series

The relationship between soil resistivity and pH levels was evaluated under controlled moisture conditions, as shown in Figure 5. The study covered a full pH range from 1.15 (highly acidic) to 14.00 (highly basic). The results indicated a drop in soil resistivity with decreasing pH value in the acidic range (below 7) or with increasing the pH value in the alkaline range (above 7). The effect of pH was higher in the acidic range than in the alkaline range. It was also observed from the data that the effect of the pH value on soil resistivity increases with increasing soil moisture.

Figure 4. Effect of Sulfate Content on Soil Resistivity: (a) Soil Resistivity vs. Gravimetric Water Content at Varied Sulfate Contents; and (b) Soil Resistivity vs. Sulfate Content at Varied Saturation Degrees

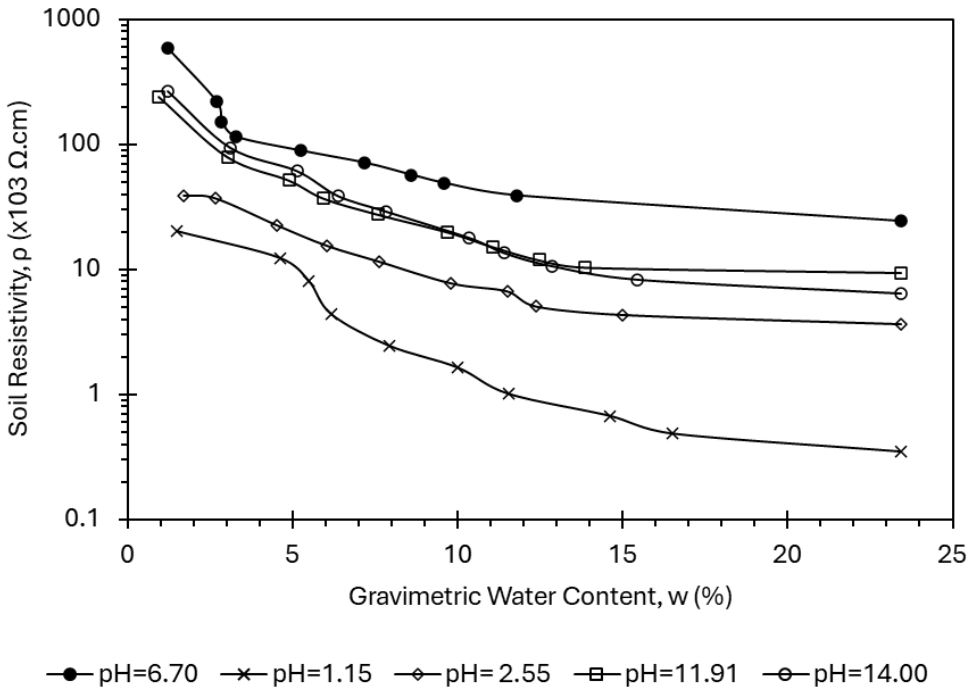


(a)

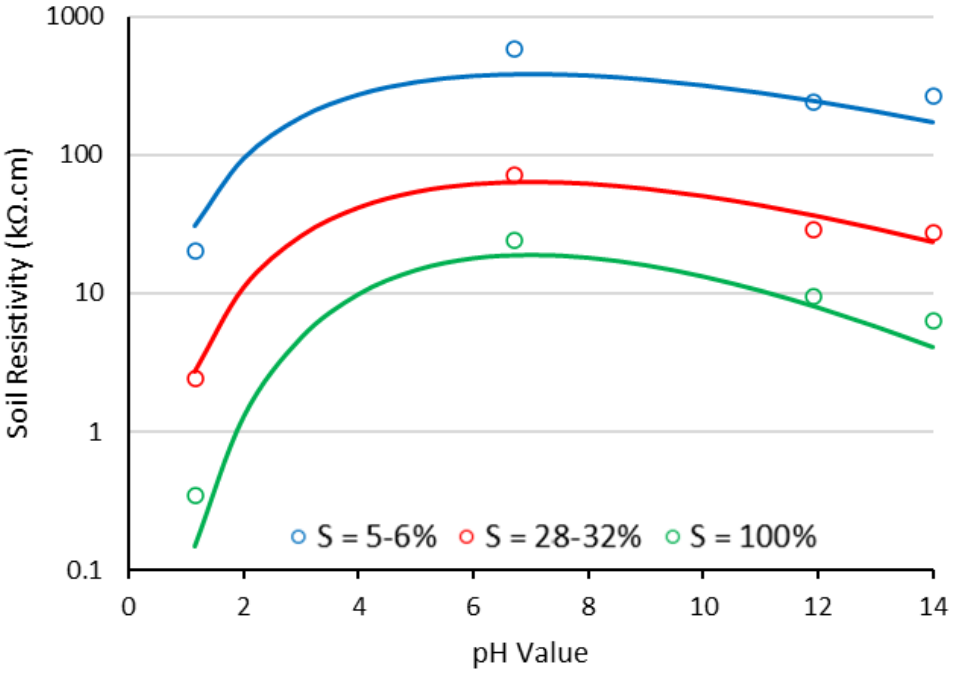


(b)

Figure 5. Effect of pH Value on Soil Resistivity: (a) Soil Resistivity vs. Gravimetric Water Content at Varied pH Values; and (b) Soil Resistivity vs. pH Value at Varied Saturation Degrees



(a)

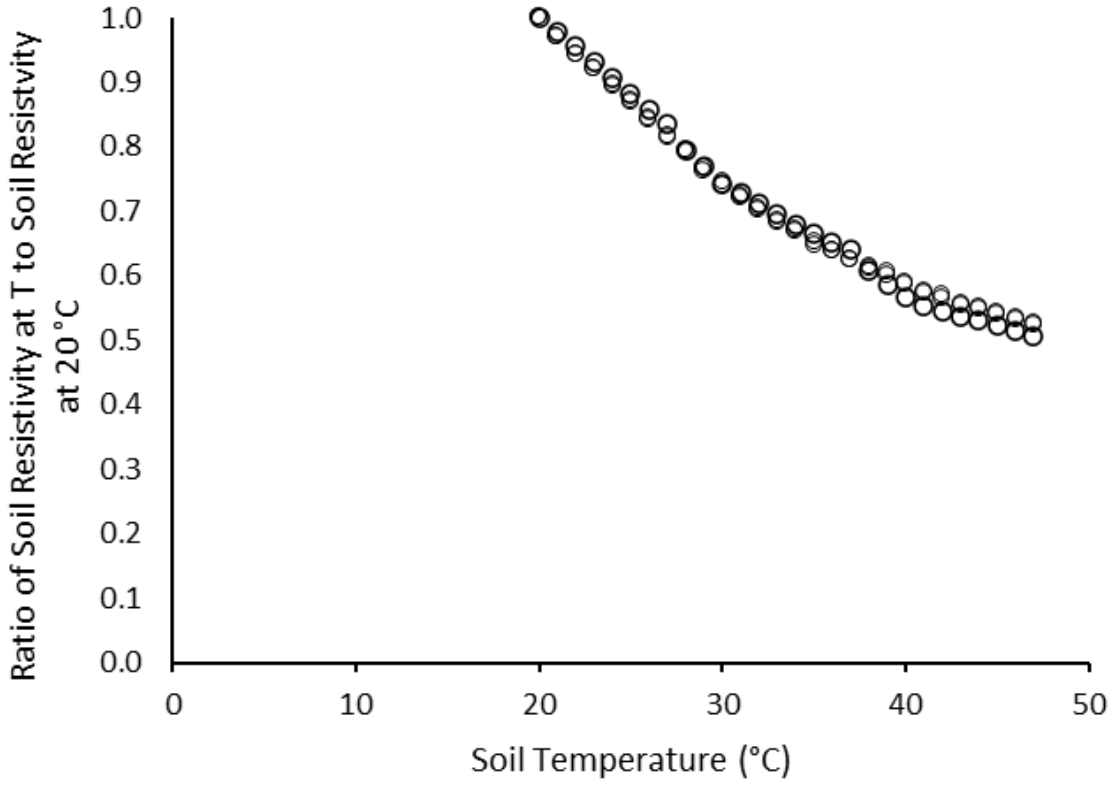


(b)

2.5.6 Soil Temperature Series

The influence of soil temperature on resistivity was examined across a range of 20–50 ˚C at three different saturation levels: 10%, 50%, and 100%, as shown in Figure 6. The results indicate that the soil resistivity decreases with increasing temperature. Unlike other parameters, no effect was observed for soil moisture on the dependency of the soil resistivity on the soil temperature. A few tests were conducted at elevated temperatures between 50–75 ˚C. Those data may not be practical for the applications discussed herein. However, the results indicated a sudden reduction in soil resistivity above 65 ˚C. Such a change in trend has been observed and reported by Wen et al. (2020) for yellow soil at about 80 ˚C.

Figure 6. Effect of Soil Temperature on Soil Resistivity



3. Element-Scale Soil Resistivity Testing in Controlled Variable Conditions

Following the tests from Section 2, a pilot testing setup was devised to monitor changes in soil moisture and resistivity under simulated rainfall events. The setup aimed to evaluate the suitability of the controlled constant tests (Section 2 of this report) in producing data that could be used to predict the soil resistivity variation in a soil environment with continuously varying moisture, representative of field conditions.

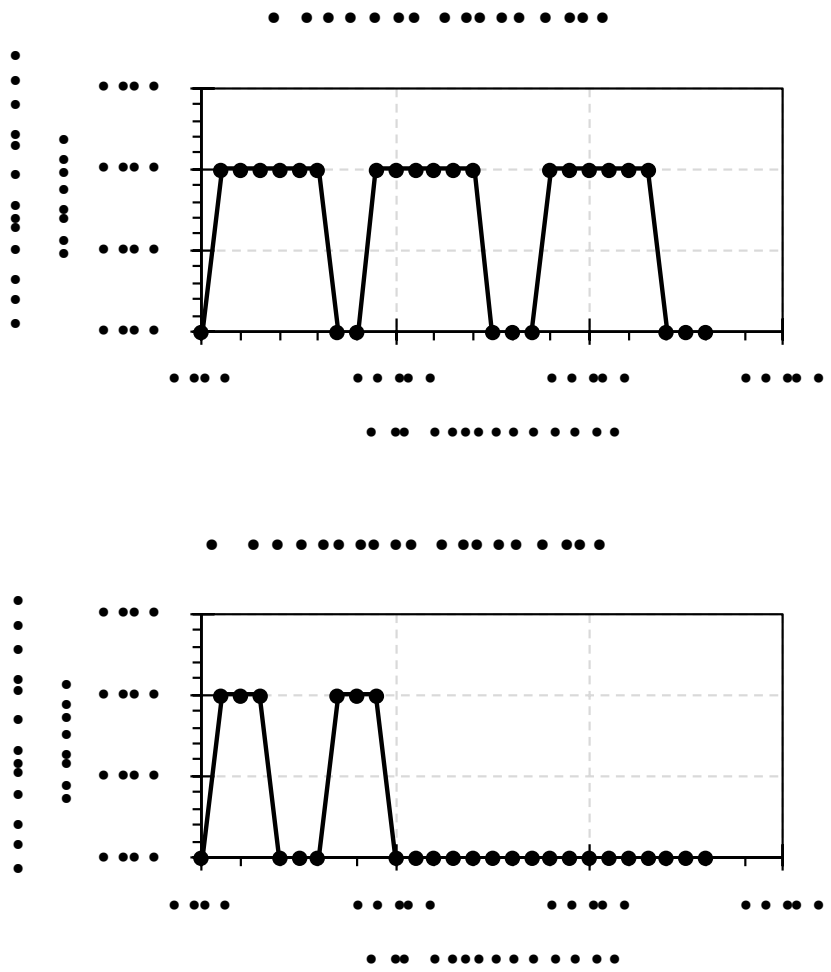
3.1 Testing Setup

A custom 3D-printed test box, of the same size as that specified by the ASTM G57, was designed to allow for controlled water infiltration and percolation. A programmable micro-pump was used to control the water infiltration rate and schedule. The bottom of the box was perforated to allow for percolation. The setup involved sensors to monitor the soil moisture, soil temperature, and ambient humidity. A capacitive soil moisture sensor was used to measure dielectric changes in the soil based on the principle of a parallel plate capacitor and Gauss's Law. The sensor was calibrated against the Miller 400D soil resistivity meter. An Arduino Nano microcontroller was used to collect and synchronize data from all sensors. A microSD card module was integrated to store time-stamped multivariable readings at 2-second intervals throughout the test duration.

3.2 Experimental Approach

A pilot test was conducted using the standard soil used in this study, compacted to the same target void ratio of 0.61 and fully saturated with deionized water. The water was applied using the micropump at a rate of 1000 mL/min. Two hypothetical rainfall scenarios (for pilot testing purposes) were applied at set intervals to simulate heavy and moderate wetting events, as shown in Figure 7. The first scenario simulated a heavy rain event, delivered in three 5-second cycles. This provided the required amount of water to saturate the soil. The second scenario simulated a light rain event, applied 10 hours after the first event, using two 3-second cycles. No changes were made to the water chemistry (i.e., the only variable in this pilot test was the soil moisture).

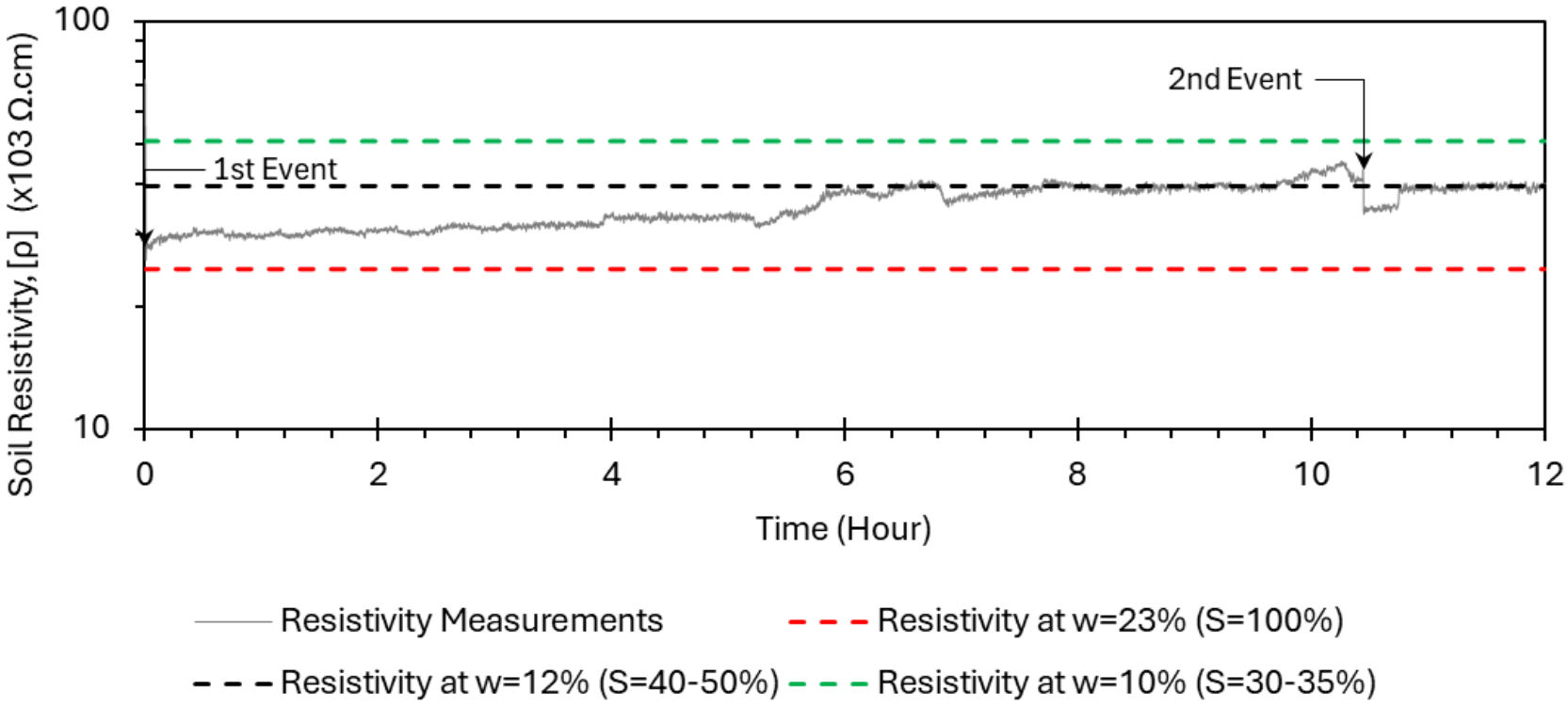
Figure 7. Rain Simulation Pulses Used in the Pilot Test.



3.3 Pilot Results and Analysis

Figure 8 presents the soil resistivity measurements with time. The soil resistivity rises from its value at saturation to lower values as the soil drains. A spike in the soil resistivity was observed during the second infiltration event. Figure 8 also shows three soil resistivity values at water contents of 10, 12, and 23% obtained from the controlled constant conditions tests (Section 2 of this report).

Figure 8. Soil Resistivity Variation with Time for a Pilot Test



4. Analysis of the NBS Romanoff Corrosion Dataset

This study re-examined the Romanoff (1957) corrosion dataset of the National Bureau of Standards (NBS) through a contemporary data science framework. A unified, digitized database was constructed and a generalizable machine learning model was developed for site-level prediction of corrosion mass loss.

4.1 Curation of the Romanoff (1957) Corrosion Database

This research focused only on the corrosion of uncoated ferrous materials; therefore, only relevant data were compiled into the database, which are included in Tables 6, 10, 13, 14, 15, 21, 27, and 29 of the NBS report (Romanoff, 1957). Specifically, site-level soil chemistry and environmental attributes were obtained primarily from Table 6. This was supplemented with soil texture, organic/inorganic classification, redox condition, and aeration from Tables 7, 9, and 19. Additionally, the Thornthwaite Moisture Index (TMI) was derived using site information to capture climatic variability, which correlates with near-surface soil moisture patterns. Material-specific metallurgical and geometric properties were obtained from Table 10, which required normalization and the creation of a unique Material ID to resolve non-unique specimen symbols and ensure consistent relational joins within the database. Corrosion outcomes were sourced from Tables 13, 14, 15, 19, 21, and 27, documenting weight loss (oz/ft^2) as a function of site, material, and exposure duration, with an explicit "specimen status" field added to identify destroyed specimens.

Data from all tables were integrated using Site Number and Material ID to form specimen–site–year records, including soil, environmental, material, and performance variables. Feature selection emphasized variables with clear physical relevance to corrosion processes, adequate completeness across sites, and compatibility with site-exclusive model validation, while excluding identifiers and artefacts not expected to generalize. A summary of the database parameters is presented in Table 1 of this report.

Table 1. Parameters Curated in the Corrosion of Buried Steel Database.

NBS Report Table No.	Data Curated
Table 6	Site Number, Soil, County, State, Internal Drainage, Resistivity at 60 °F (Ohm-cm), pH Value, Total Acidity, {Na+K as Na, Ca, Mg, CO_3, HCO_3, Cl, SO_4} (mg-eq/100g of soil), Temperature (°F), Annual Precipitation (inches), Moisture Equivalent (%), Air-pore space (%), Apparent Specific Gravity, Volume Shrinkage (%)
Table 7	Site Number, Depth (inches), Sand (%), Clay (%), Silt (%)
Table 9	Inorganic/Organic Soils, Oxidizing/Reducing Soils
Table 10	Material, Symbol, Year, Number of specimens, Form, Width or Diameter, Length, Thickness, C (%), Si (%), Mn (%), S (%), P (%), Cr (%), Ni (%), Cu (%), Mo (%), Other Elements, C Free (%), C Combined (%)
Tables 13, 14, 15, 19, 21, 27	Site Number, Soil Type, Year, Duration, Material Name, Type, Material ID, Material Code, Weight Loss in oz/ft^2

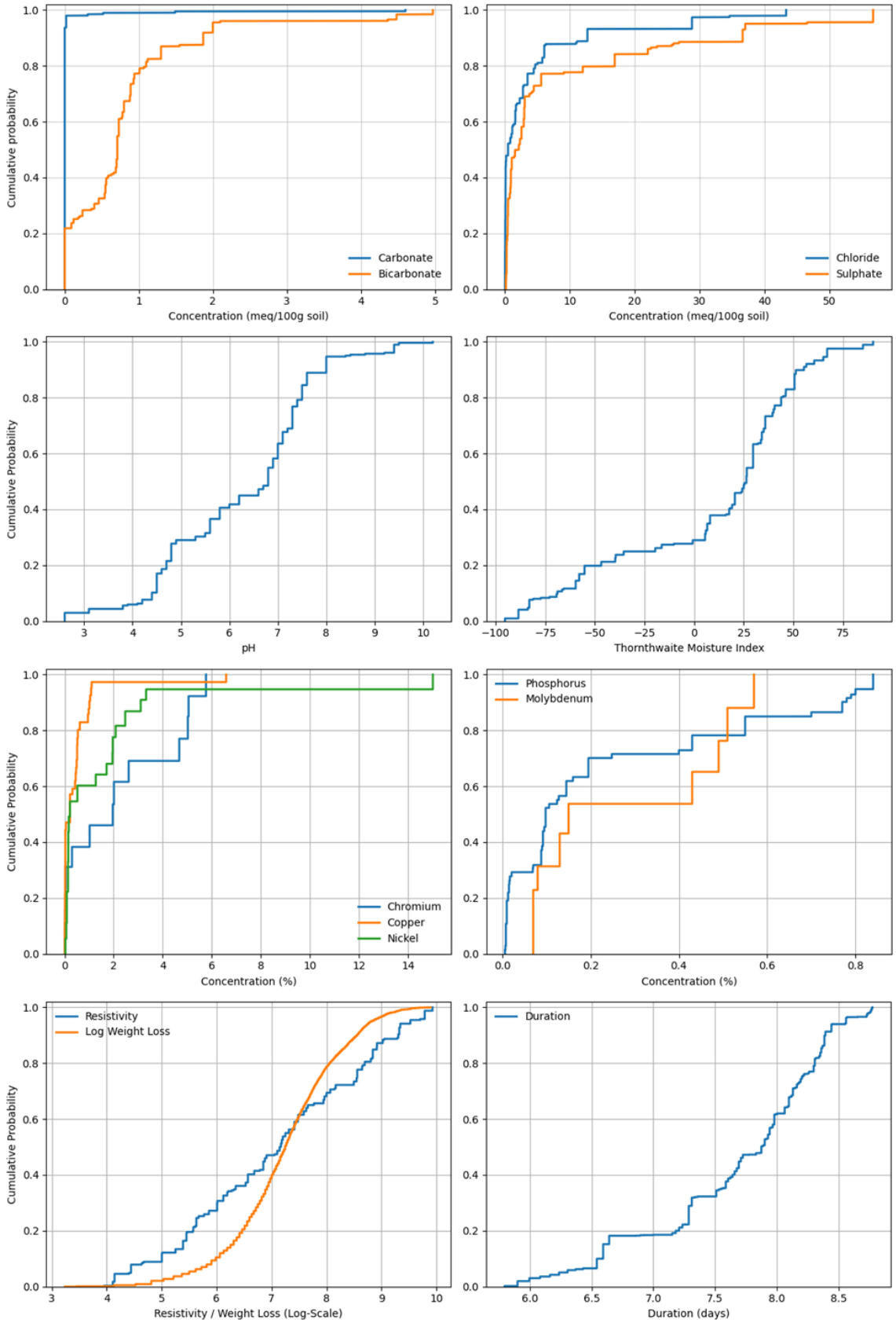
4.2 Feature Selection for Model and Data Preprocessing

Feature selection was guided by three primary considerations: (i) physical and chemical relevance to underground corrosion processes, (ii) data availability and completeness across sites, and (iii) suitability for robust generalization under a site-exclusive validation framework. Emphasis was placed on variables with plausible causal influence on corrosion behavior, while avoiding features that primarily served as identifiers, proxies, or artifacts of the original experimental design. Table 2 provides a summary of the selected features, their sources, assumptions made, and any preprocessing they underwent during curation. Figure 9 presents the statistical distributions of the model features to show the range and distribution uniformity of each feature in the database.

Table 2. Selected Features Used in Developing the Machine Learning Model for Corrosion of Buried Steel Prediction

Feature	Source	Physical / Modelling Assumption	Pre-processing
Corrosion weight loss (target)	NBS Tables 13, 14, 15, 19, 21, 27	Intact-specimen mass loss represents cumulative underground corrosion damage	Converted from oz/ft^2 to g/m^2; log-transformed
Exposure duration	NBS Tables 13, 14, 15, 19, 21, 27	Corrosion accumulates over time; duration captures temporal effects more directly than calendar year	Converted from years to days
Soil resistivity	NBS Table 6	Primary electrochemical control on corrosion rate; highly nonlinear influence	Log-transformed
Soil pH	NBS Table 6	Governs soil acidity/alkalinity and corrosion thermodynamics	None
Internal drainage	NBS Table 6	Proxy for soil aeration, moisture movement, and oxygen availability	Ordinal encoding (VP = 0, P = 1, F = 2, G = 3)
Chloride (Cl^-)	NBS Table 6	Increases soil conductivity and promotes aggressive corrosion	None
Sulfate (SO_4^{2-})	NBS Table 6	Influences conductivity, buffering, and corrosion kinetics	None
Carbonate (CO_3^{2-})	NBS Table 6	Contributes to alkalinity and buffering capacity	None
Bicarbonate (HCO_3^-)	NBS Table 6	Regulates soil buffering and corrosion chemistry	None
Phosphorus (P)	NBS Table 10	Alloying element affecting corrosion resistance	None
Chromium (Cr)	NBS Table 10	Promotes passivation and corrosion resistance	None
Nickel (Ni)	NBS Table 10	Enhances corrosion resistance and alloy stability	None
Copper (Cu)	NBS Table 10	Improves corrosion resistance in ferrous alloys	None
Molybdenum (Mo)	NBS Table 10	Increases resistance to localized corrosion	None
Thornthwaite Moisture Index (TMI)	ArcGIS TMI-1974 Map by Olaiz et al. (2017)	Integrated indicator of long-term climatic moisture controlling soil wetness	Nearest-neighbor spatial assignment
Site number (grouping only)	NBS Table 6	Enforces site-exclusive validation and prevents data leakage	Used only for GroupKFold CV

Figure 9. Statistical Distributions of the Model Features Based on Data from Romanoff (1957)



4.3 Development of Prediction Model

Modeling decisions were shaped by three primary constraints: (1) substantial and uneven missingness across features and sites; (2) strongly nonlinear and multicollinear relationships among soil chemistry, environmental conditions, and alloying elements; and (3) the requirement for generalization to unseen sites. Given the substantial proportion of missing data within the dataset, tree-based ensemble methods were identified as the most suitable modeling approach as they are intrinsically resilient to missing values and multicollinearity and do not necessitate extensive feature scaling or transformation. In particular, scikit-learn's ExtraTreesRegressor incorporates missing values directly within the tree-construction process rather than requiring prior imputation. This integrated handling of incomplete observations eliminates the need for external imputation procedures and enables the model to learn from the observed data distribution without artificially modifying feature values.

Several supervised learning algorithms were evaluated, including Random Forest, XGBoost, and LightGBM. The ExtraTrees regressor (Breiman, 2001; Geurts et al., 2006; Chen & Guestrin, 2016) was eventually selected for its ensemble-based architecture, resilience to multicollinearity, and capability to capture complex nonlinear interactions without requiring extensive feature scaling or transformation. To approximate real-world use across new geographies, site exclusivity was enforced using a nested GroupKFold design (Pedregosa et al., 2011; Roberts et al., 2017), with inner folds performing a randomized hyperparameter search and outer folds providing unbiased estimates on entirely unseen sites. All models were developed with a fixed random seed (SEED = 42) to ensure full reproducibility of results.

4.3.1 Training and Validation

Model development and validation were conducted using a nested 5-Fold cross-validation (CV) framework to ensure rigorous and unbiased estimation of model generalization across distinct sampling sites. The grouping variable "site no." was used to prevent data leakage between the training and testing folds, thereby ensuring that all samples originating from a given site were exclusively included in either the training or testing subset. This approach was designed to simulate real-world prediction scenarios where the model encounters data from entirely unseen sites.

The outer loop of the nested CV (five folds) was used for the final model evaluation, whereas the inner loop (four folds) was used for hyperparameter optimization via randomized search. Sixty candidate parameter combinations were sampled from predefined search spaces, covering key model parameters such as tree depth, number of estimators, and feature sampling strategies.

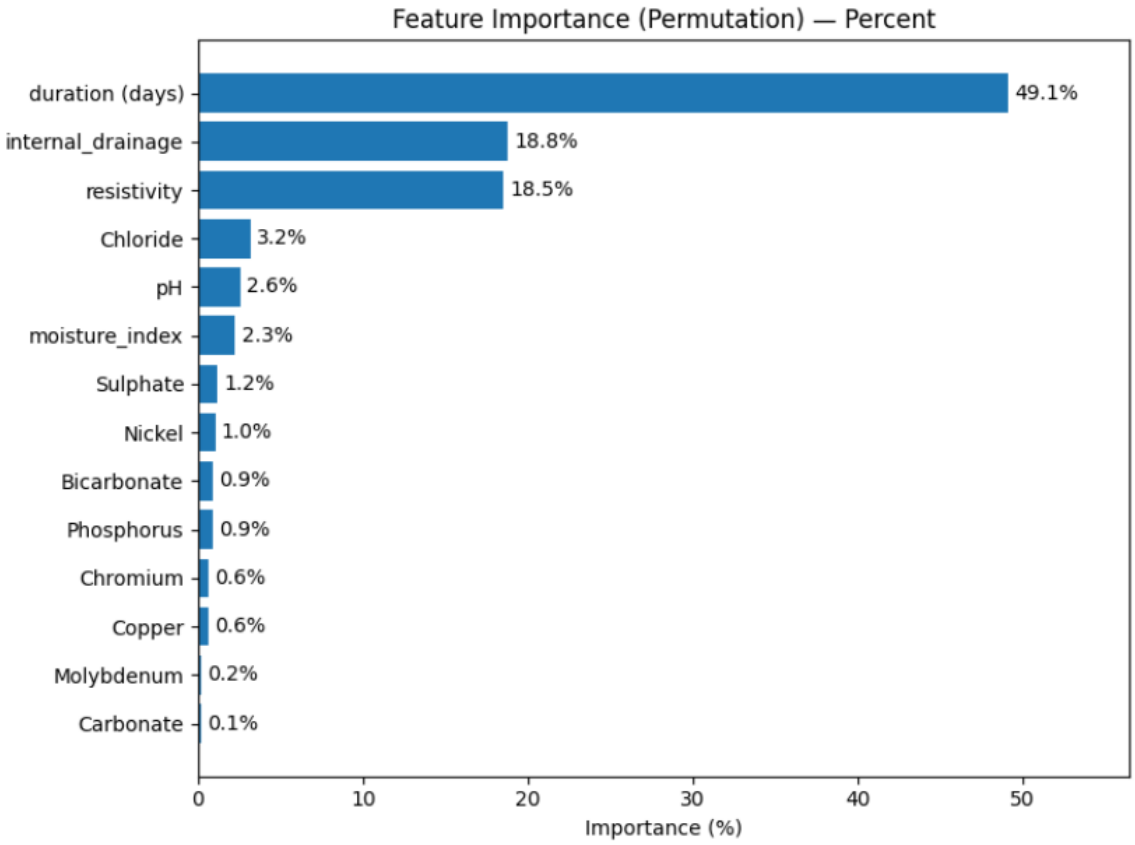
4.3.2 Feature Importance

Feature importance was quantified using permutation importance (Breiman, 2001; Fisher et al., 2019) based on the ExtraTrees regressor, with results expressed as percentages for interpretability,

as shown in Figure 10. For each outer test fold within the nested GroupKFold framework (grouped by site), each feature was permuted independently on the hold-out dataset, and the corresponding decrease in model performance (ΔR^2) was recorded. The ΔR^2 values were then averaged across all folds to produce a global importance score for each feature. These aggregated values were subsequently normalized to percentages summing to 100% to compare the relative contributions of model features.

In this formulation, a higher percentage denotes a greater reduction in predictive performance when the corresponding feature is randomized, indicating stronger model dependence on that variable. Conversely, lower percentages reflect features with minimal or redundant predictive influence. These values represent relative importance with respect to model generalization across unseen sites.

Figure 10. Prediction Model Feature Importance



4.3.3 Model Evaluation

Model performance was assessed using the outer test folds of a nested GroupKFold cross-validation framework, in which all observations from a given site were held out together. This site-exclusive design provides an unbiased estimate of generalization performance for previously unseen sites, consistent with the intended deployment scenario.

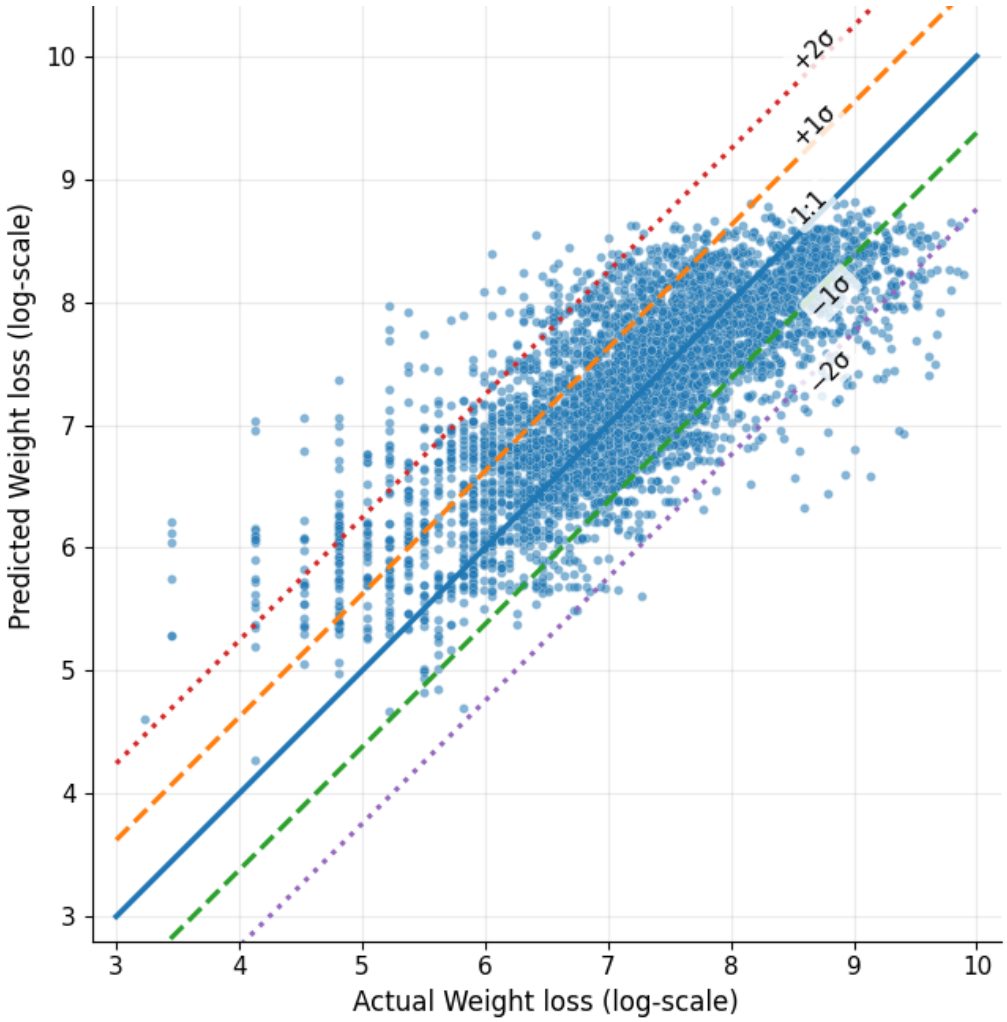
Across five outer folds, the ExtraTrees regressor achieved a mean coefficient of determination of approximately 0.60 in log-space, indicating that about 60% of the variance in log-transformed corrosion weight loss is explained for unseen sites. Performance showed moderate fold-to-fold variability, reflecting genuine heterogeneity among sites rather than overfitting. The log-scale RMSE corresponds to a typical multiplicative prediction error of approximately 1.9 (i.e., predictions are generally within a factor of two of observed values). Table 3 summarizes the aggregated performance metrics across outer folds.

Table 3. Summary of the Aggregated Performance Metrics Across Outer Folds.

Metric	Mean ± Standard Deviation
R^2 (log-space)	0.605 ± 0.100
Adjusted R^2 (log-space)	0.600 ± 0.101
RMSE (log-space)	0.616 ± 0.106
RMSE (grams)	1735.74 ± 438.87

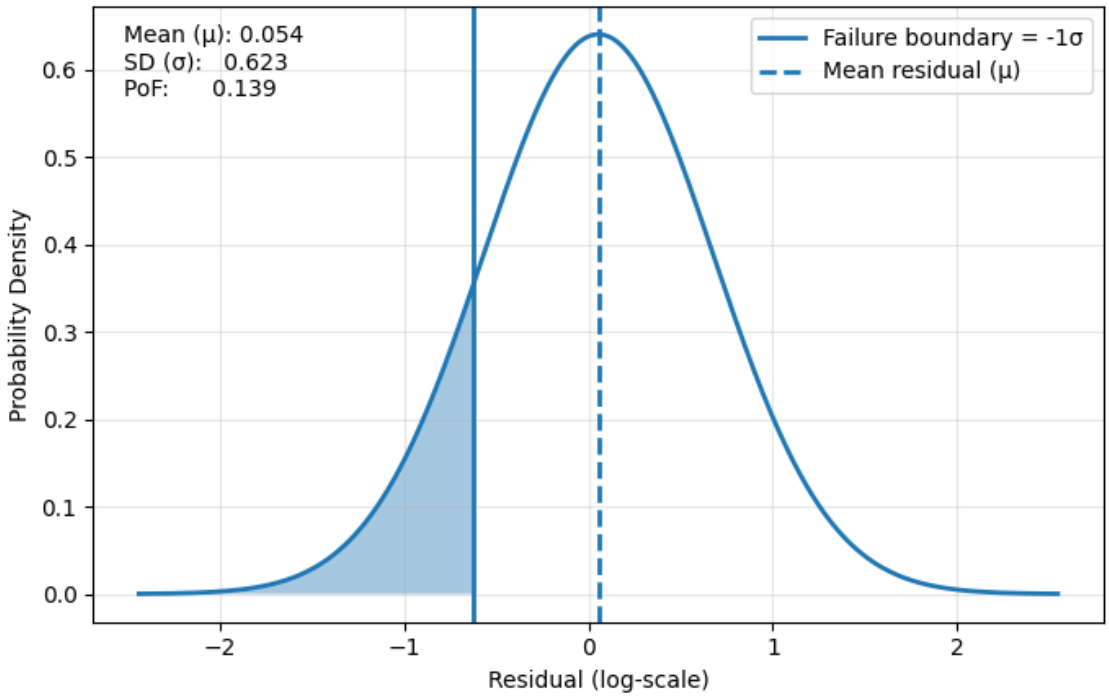
The predicted-versus-measured plot (Figure 11) shows a strong positive correlation along the 1:1 line, demonstrating that the model captures the dominant structure of the data and generalizes across geographically distinct sites. However, the scatter plot shows systematic regression-to-the-mean behavior, where corrosion losses at the lower end of the range are overpredicted, while extreme losses are underpredicted. This pattern indicates increasing uncertainty at the tails of the response distribution and that the model is most reliable within the central range of observed corrosion severity.

Figure 11. Predicted vs. Measured Corrosion Weight Loss



The residual distribution plot (Figure 12) further characterizes model uncertainty. Residuals in log-space are approximately symmetric and centered near zero, with a slightly positive mean bias. A fitted normal distribution provides an approximation for uncertainty quantification, allowing estimation of exceedance probabilities relative to specified tolerance or failure thresholds. The shaded area to the left of this boundary represents the Probability of Failure (PoF = 0.14), meaning that about 14% of model predictions on unseen sites are expected to underpredict the true corrosion by more than one standard deviation. Overall, the model is nearly unbiased but exhibits enough variability that a non-negligible fraction of predictions fall into the defined failure region.

Figure 12. Corrosion Model Prediction Residual Distribution



Overall, the nested, site-exclusive evaluation demonstrates that the ExtraTrees model could provide stable and generalizable predictions under realistic conditions. While conservative at the extremes, its consistent performance across unseen sites supports its applicability for site-level inference and large-scale predictive mapping of underground corrosion risk.

4.4 Impact of Individual Features

To understand the individual impact of the features, ceteris paribus analysis was used, where all variables except the one being examined are held fixed at benchmark values (Table 4). The benchmark values were chosen as the median of each feature across the dataset. The median is preferred because many environmental and soil composition variables are skewed and contain outliers; using the median provides more realistic estimates that are less sensitive to extreme values. This allows the model to respond to the variable of interest to be interpreted as a local effect under typical conditions, rather than under an unrealistic or outlier scenario.

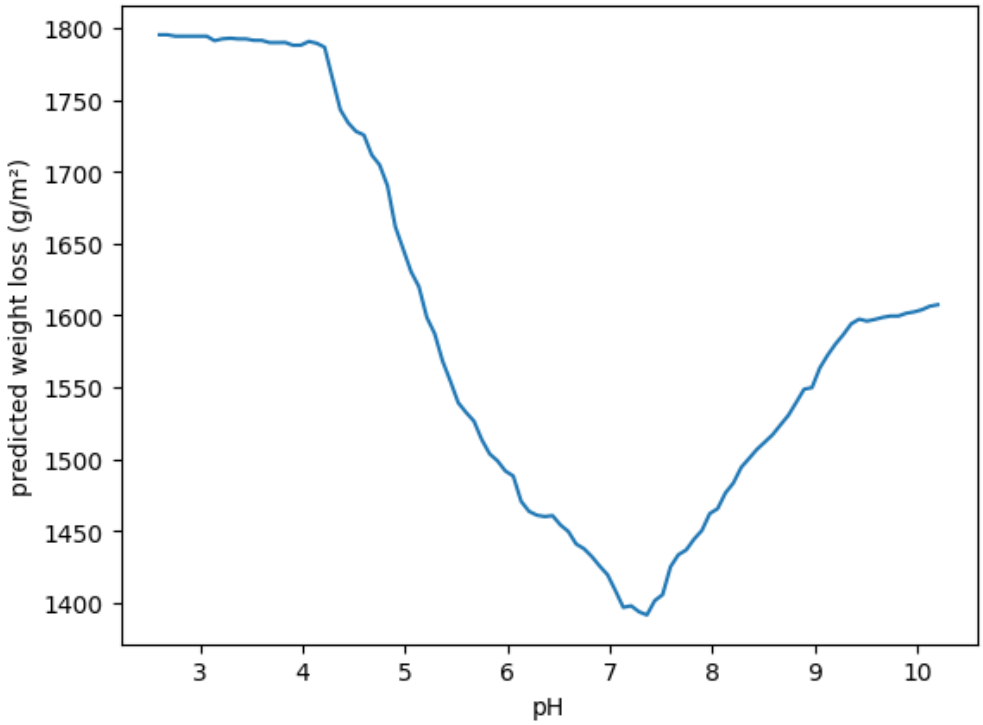
Table 4. Feature Benchmark Values Used in the Evaluation of Individual Feature Impacts

Feature	Benchmark
Bicarbonate (mg-eq/100 g of soil)	0.71
Carbonate (mg-eq/100 g of soil)	0.0
Chloride (mg-eq/100 g of soil)	0.4
Chromium (%)	1.96
Copper (%)	0.22
Molybdenum (%)	0.15
Nickel (%)	0.22
Phosphorus (%)	0.098
Sulfate (mg-eq/100 g of soil)	2.13
Internal Drainage	Fair

4.4.1 Impact of pH Value

Figure 13 shows the model prediction of the pH value effect on weight loss. Predicted corrosion decreases with increasing pH in acidic conditions (pH < 7), reaches a minimum near neutral conditions (pH = 7), and increases with increasing pH in alkaline conditions (pH > 7). This behavior suggests that the model captures the expected effect of pH on corrosion. Acidic soils promote corrosion through increased proton availability, whereas more alkaline soils often contain higher concentrations of dissolved salts and mobile ions which improve electrical conductivity and allow corrosion currents to flow more readily (Romanoff, 1957).

Figure 13. Model Prediction of the pH Value Effect on Weight Loss



Duration = 7.2 years; TMI = 26.39; ρ = 1271 Ohm-cm;
Sulfates = 2.13 mg-eq/100 g; Chlorides = 0.4 mg-eq/100 g.

4.4.2 Impact of Moisture Index

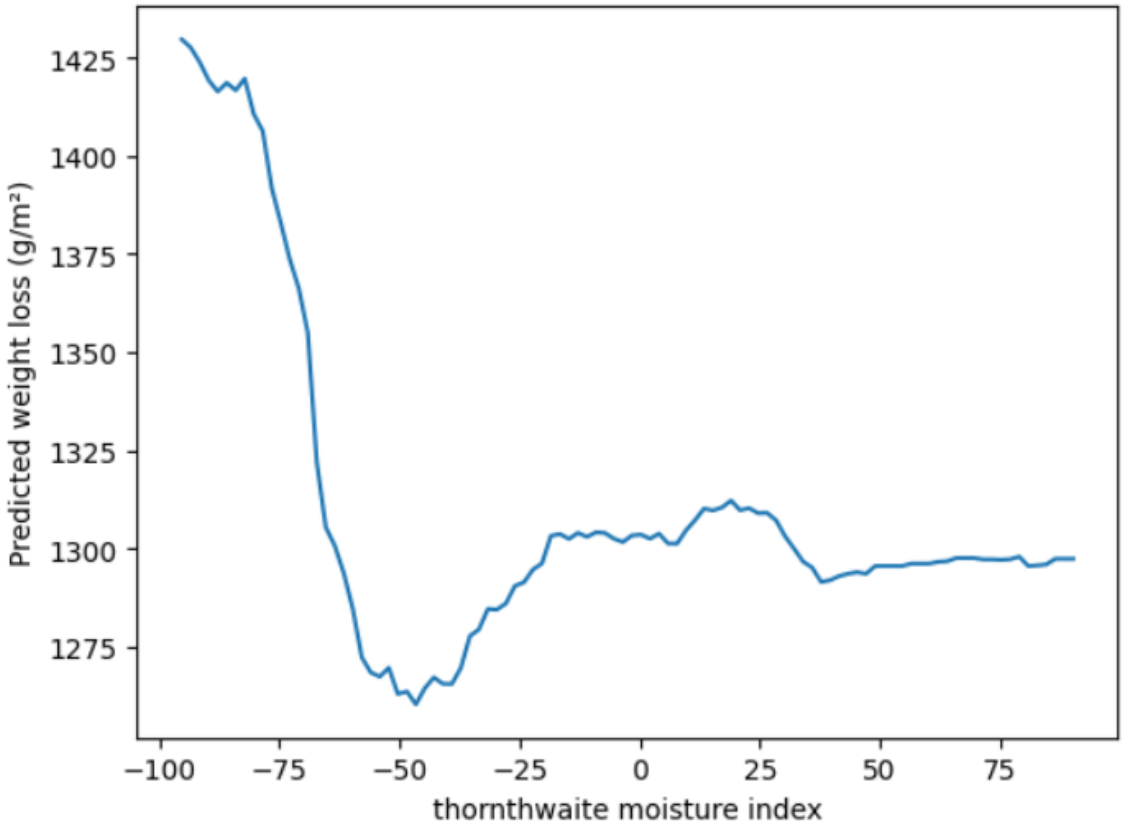
Figure 14 shows the model prediction of the moisture effect on weight loss. The relationship between the Thornthwaite Moisture Index (TMI) and the predicted corrosion initially appears to contradict established corrosion theory, which generally associates increasing moisture availability with increased corrosion activity. However, this apparent inconsistency arises from the interpretation of the TMI and from the interaction of moisture with other controlling soil properties. TMI represents a long-term climatic moisture balance derived from precipitation and potential evapotranspiration and does not directly quantify soil moisture conditions at the specimen burial depth. Consequently, low TMI values should not be interpreted as uniformly dry environments at the specimen-soil interface.

Analysis of the dataset indicates that locations with very low TMI values (−100 to −60) are associated with substantially lower soil resistivity compared to locations with moderate to high TMI values (−60 to +90). Low soil resistivity is indicative of high ionic conductivity, which is commonly caused by elevated concentrations of dissolved salts. These dissolved ions enhance electrochemical reactions by facilitating charge transport within the soil electrolyte, thereby increasing corrosion rates even under climatically dry conditions. In contrast, locations with higher

TMI values generally exhibit higher soil moisture. In wetter soils, corrosion may also be partially limited by restricted oxygen diffusion due to water-filled pore spaces, which can suppress cathodic reaction rates. As a result, increased moisture does not necessarily translate to increased corrosion, particularly when oxygen transport becomes a limiting factor. The observed non-monotonic behavior of the TMI–corrosion curve is therefore attributed to the combined influence of electrolyte conductivity, dissolved salt content, and oxygen availability rather than moisture alone.

The model response reflects correlations present within the dataset, where arid climatic regions are frequently associated with salt-affected, highly conductive soils that promote corrosion. Finally, it is noted that the interpretation of model behavior at TMI values below −60 should be undertaken with caution. In the original Thornthwaite formulation, −60 represents the lower bound of the moisture index, and values below this threshold arise from revised or extended classifications (Thornthwaite, 1948; Feddema, 2005). As a result, predictions in this range may be influenced by reduced data density and stronger feature interactions, which can lead to irregular or non-intuitive trends in model-derived response curves.

Figure 14. Model Prediction of the Moisture Effect on Weight Loss

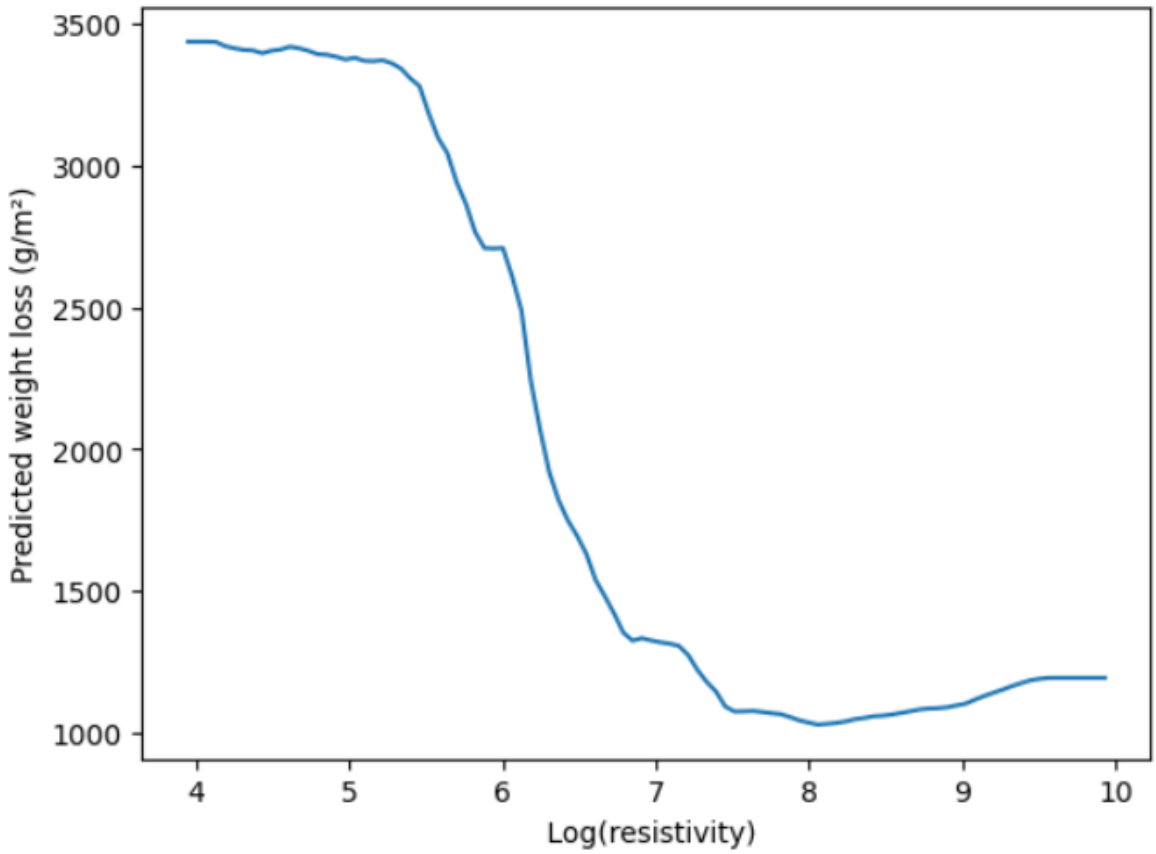


Duration = 7.2 years; ρ = 1271 Ohm-cm; pH = 6.8;
Sulfates = 2.13 mg-eq/100 g; Chlorides = 0.4 mg-eq/100 g.

4.4.3 Impact of Soil Resistivity

Figure 15 shows the model prediction of the soil resistivity effect on weight loss. Corrosion decreases with increasing soil resistivity. A steep corrosion decline rate was predicted between log(ρ) of 5.5 and 7.5. The slight increase beyond this range is attributed to regime change (i.e., a shift in the dominant controlling mechanisms) and feature interactions rather than a causal effect of resistivity. Examination of the data distribution indicates that the high-resistivity subset (log(ρ) > 8) is characterized by a distinctly different data structure compared to the lower-resistivity range. While resistivity values extend to very high levels in this subset, several accompanying soil chemistry variables exhibit limited availability or negligible variability, with a large proportion of records lacking values for features such as carbonate, bicarbonate, chloride, and sulfate. Consequently, predictions in the high resistivity region are more influenced by other available variables, most notably pH and moisture index, as well as by the learned decision structure of the model.

Figure 15. Model Prediction of the Soil Resistivity Effect on Weight Loss

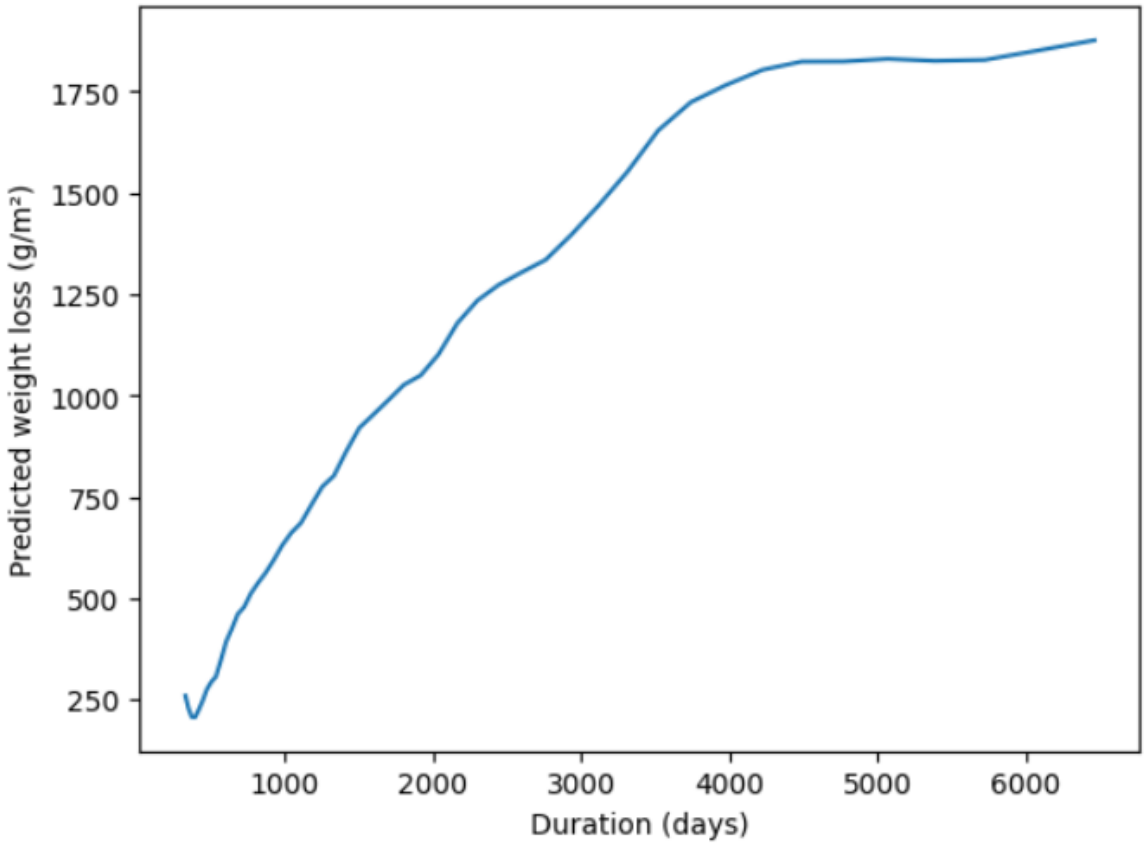


Duration = 7.2 years; TMI = 26.39; pH = 6.8;
Sulfates = 2.13 mg-eq/100 g; Chlorides = 0.4 mg-eq/100 g.

4.4.4 Impact of Exposure Duration

Figure 16 shows the model prediction of the exposure duration effect on weight loss. The model captured the corrosion increase with increasing exposure duration, at a decelerating corrosion rate over time. This suggests that the model could learn long-term corrosion rate deceleration typically caused by the buildup of rust scales, which serves as a protection against further corrosion.

Figure 16. Model Prediction of the Exposure Duration Effect on Weight Loss



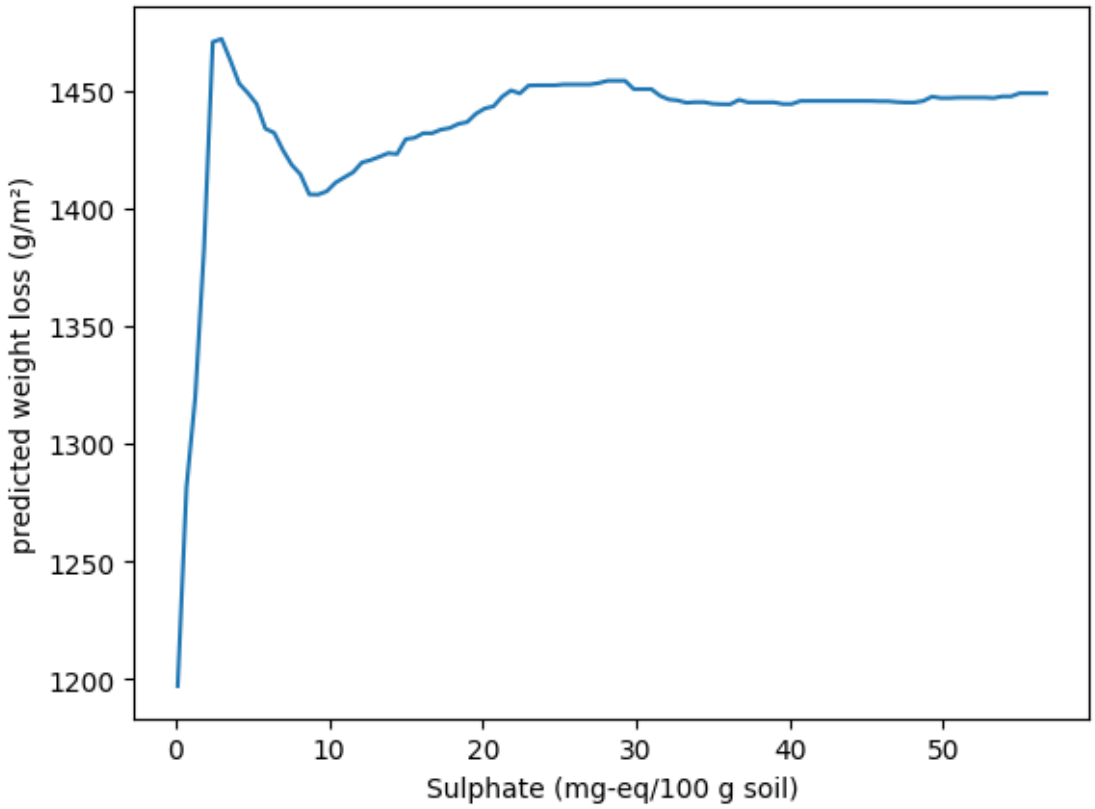
Duration = 7.2 years, ρ = 1271 Ohm-cm; TMI = 26.39; pH = 6.8;
Sulfates = 2.13 mg-eq/100 g; Chlorides = 0.4 mg-eq/100 g.

4.4.5 Impact of Sulfate Concentration

Figure 17 shows the model prediction of the sulfate concentration effect on weight loss. The predicted influence of sulfate concentration on corrosion must be interpreted with consideration to its highly skewed data distribution, in which most observations are concentrated at low sulfate levels, with relatively few samples at higher concentrations. Predicted corrosion loss increases with increasing sulfate concentration at low sulfate concentration ranges (ranges most densely represented in the dataset). At higher sulfate concentration ranges (ranges of a limited number of

observations in the dataset), the predicted response gradually levels off. This behavior suggests that the model captures the expected effect of sulfate concentration on corrosion.

Figure 17. Model Prediction of the Sulfate Concentration Effect on Weight Loss

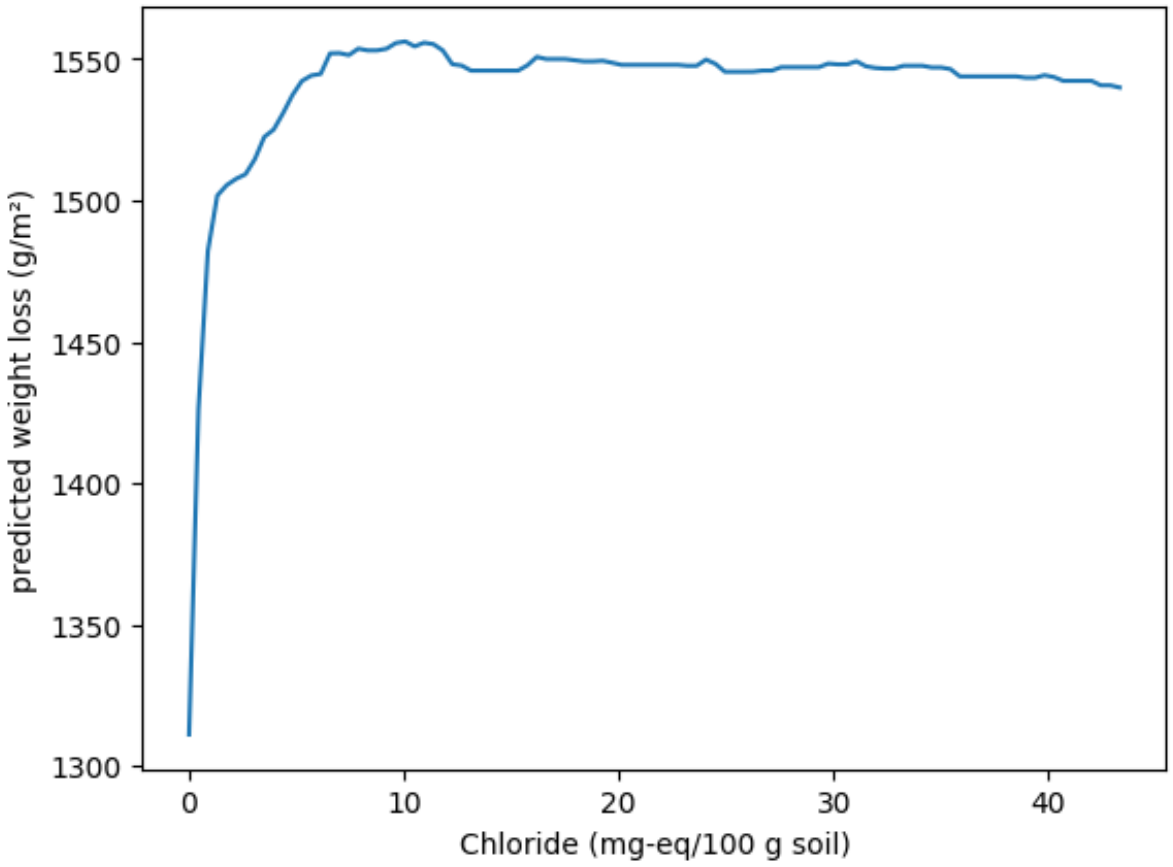


Duration = 7.2 years, ρ = 1271 Ohm-cm; TMI = 26.39; pH = 6.8;
Chlorides = 0.4 mg-eq/100 g.

4.4.6 Impact of Chloride Concentration

Figure 18 shows the model prediction of the chloride concentration effect on weight loss. Similar to sulfate concentration, corrosion increases with increasing chloride concentration at low chloride concentration ranges (ranges most densely represented in the dataset). At higher chloride concentration ranges (ranges of a limited number of observations in the dataset), the predicted response gradually levels off. This behavior suggests that the model captures the expected effect of chloride concentration on corrosion.

Figure 18. Model Prediction of the Chloride Concentration Effect on Weight Loss.



Duration = 7.2 years, ρ = 1271 Ohm-cm; TMI = 26.39; pH = 6.8;
Sulfates = 2.13 mg-eq/100 g.

5. Summary and Conclusions

This research aimed to evaluate soil resistivity and corrosivity, considering the various key parameters that contribute to corrosion of buried steel. Three studies were completed and documented in this report:

- Element-Scale Soil Resistivity Testing in Controlled Constant Conditions. This is an experimental study that aimed to evaluate the soil resistivity in controlled, constant conditions. This study involved devising a new experimental protocol that was evaluated for its ability to produce repeatable results. An experimental program was devised and implemented to evaluate the effect of testing parameters known to be the most relevant in soil corrosivity by varying one parameter at a time. The results of the testing program showed the potential of the experimental approach to provide the necessary data to develop empirical prediction models for soil resistivity. The influence of each parameter was evaluated separately. Soil resistivity was observed to decrease with increasing soil moisture, chloride content, sulfate content, acidity, alkalinity, and soil temperature.

- Element-Scale Soil Resistivity Testing in Controlled Variable Conditions. This is a pilot experimental study that aimed to evaluate the soil resistivity in controlled, variable conditions. This study involved devising and piloting a new experimental method capable of capturing the variation in soil resistivity in a continuously varying environment. The results indicated a great potential for the proposed experimental approach in providing data necessary to calibrate soil resistivity prediction models in a continuously varying environment representative of field conditions.

- Analysis of the NBS Romanoff Corrosion Dataset. This is a data analysis study that aimed to evaluate the corrosion of buried steel using advanced data science methods. This study presented a comprehensive framework for predicting underground corrosion of uncoated ferrous materials using a supervised machine learning approach based on the ExtraTrees algorithm. Through data curation from the NBS Romanoff (1957) corrosion program and structured model validation using site-exclusive nested cross-validation, the model achieved consistent performance transferable to unseen sites.

Bibliography

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32.

Bronson, A., Castillo, C., Hinojos, J., Nazarian, S., & Borrok, D. (2020). Relating corrosion of mechanically stabilized earth reinforcements with fluid conductivity of backfill soils. *Journal of Materials in Civil Engineering*, *32*(11), 04020346.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* Association for Computing Machinery, New York, NY, United States, 785–794.

Darbin, M., Jailloux, J. M., & Montuelle, J. (1988). Durability of reinforced earth structures: The results of a long-term study conducted on galvanized steel. *Proceedings of the Institution of Civil Engineers*, *84*(5), 1029–1057.

Davalos, J., Gracia, M., Marco, J. F., & Gancedo, J. R. (1992). Corrosion of weathering steel and iron under wet-dry cycling conditions: Influence of the rise of temperature during the dry period. *Hyperfine Interactions*, *69*, 871–874.

Elias, V. (1990). *Durability/corrosion of soil reinforced structures* (Report No. FHWA-RD-89-186). Federal Highway Administration (FHWA).

Elias, V., Fishman, K., Christopher, B. R., Berg, R. R., & Berg, R. R. (2009). *Corrosion/degradation of soil reinforcements for mechanically stabilized earth walls and reinforced soil slopes* (No. FHWA-NHI-09-087). National Highway Institute (US).

Feddema, J. J. (2005). A revised Thornthwaite-type global climate classification. *Physical Geography, 26*(6), 442–466.

Fisher, A., Rudin, C., & Dominici, F. (2019). All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously. *Journal of Machine Learning Research*, *20*(177), 1–81.

Fishman, K. L., & Withiam, J. L. (2011). *NCHRP Report 675: LRFD metal loss and service-life strength reduction factors for metal-reinforced systems*. Transportation Research Board.

Fishman, K. L., Nazarian, S., Walker, S., & Bronson, A. (2021). NCHRP Research Report 958: Electrochemical test methods to evaluate the corrosion potential of earthen materials. Transportation Research Board.

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*, 3–42.

Gupta, S. K., & Gupta, B. K. (1979). The critical soil moisture content in the underground corrosion of mild steel. *Corrosion Science*, *19*(3), 171–178.

Kolay, P. K., Tajhya, D., & Mondal, K. (2020). Corrosion of steel in MSE walls due to deicers and backfill aggregates. *Geotechnical and Geological Engineering*, *38*(3), 2493–2507.

Liu, L., Li, W., Deng, Z., Xu, S., Xu, Y., Zeng, L., ... & Zhong, Z. (2023). Effect of moisture on corrosion behavior of Q235 steel in bentonite clay. *International Journal of Electrochemical Science*, *18*(6), 100164.

Morsy, A. M., & Ebo, I. A. (2025). *Development of physics-based deterioration models for reinforced soil retaining structures* (Report No. 24-43). Mineta Transportation Institute (MTI).

Mughabghab, S. F., & Sullivan, T. M. (1989). Evaluation of the pitting corrosion of carbon steels and other ferrous metals in soil systems. *Waste Management*, *9*(4), 239–251.

NASEM. (2023). *Development of guidance for the use of non-destructive testing in transportation asset management* (Project 26686). National Academies of Sciences, Engineering, and Medicine.

Noor, E. A., & Al-Moubaraki, A. H. (2014). Influence of soil moisture content on the corrosion behavior of X60 steel in different soils. *Arabian Journal for Science and Engineering*, *39*, 5421–5435.

Olaiz, A. H., Singhar, S. H., Vann, J. D., & Houston, S. L. (2017). *Interactive Thornthwaite Moisture Index (TMI) GIS application* [Web application]. ArcGIS Online.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Roberge, P. (2000). *Handbook of corrosion engineering*. New York: McGraw-Hill Education.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., ... & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, *40*(8), 913–929.

Romanoff, M. (1956). *Underground Corrosion.* National Bureau of Standards Circular 579, U.S. Department of Commerce.

Rossum, J. R. (1969). Prediction of pitting rates in ferrous metals from soil parameters. *Journal-American Water Works Association*, *61*(6), 305–310.

Shreir, L. L., Jarman, R. A., & Burstein, G. T. (1994). *Corrosion* (Vol. 1, 3rd edition). Butterworth Heinemann Ltd.

TAMP. (2022). *California Transportation Asset Management Plan*. California Department of Transportation (Caltrans).

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical Review, 38*(1), 55–94.

Wasim, M., Shoaib, S., Mubarak, N. M., Inamuddin, & Asiri, A. M. (2018). Factors influencing corrosion of metal pipes in soils. *Environmental Chemistry Letters*, *16*, 861–879.

About the Authors

Amr M. Morsy, PhD, PE

Dr. Amr Morsy is a professional civil engineer with experience in both academia and industry, focusing on geotechnical engineering, transportation geotechnics, environmental geotechnics, and climate adaptation. He obtained his B.Eng. and M.Sc. degrees in civil engineering from Cairo University in 2011 and 2013, respectively, and obtained his PhD degree in civil engineering from The University of Texas at Austin in 2017. He worked as a postdoctoral fellow at The University of Texas at Austin in 2018 and as a practicing geotechnical engineer from 2018 to 2020. He later worked as a research associate at Loughborough University on the ACHILLES program grant from 2020 to 2022. He has been working as an Assistant Professor at California State University, Long Beach since 2022.

As part of his academic experience, Dr. Morsy conducts research on geotechnical infrastructure deterioration and asset management, climate change impacts on geotechnical infrastructure, and geotechnical solutions for sustainable built environments. He has excelled in physical and numerical modeling of geotechnical and geoenvironmental engineering systems, infrastructure instrumentation, and laboratory experimentation. He participated in research projects sponsored by the Transportation Research Board of the National Academies of Sciences, Engineering, and Medicine, the Engineering and Physical Sciences Research Council of the UK Research and Innovation, the US Federal Highway Administration, the Geosynthetic Institute, the Departments of Transportation of Texas and Indiana, and geosynthetic manufacturers.

As part of his professional consulting experience, Dr. Morsy conducts rigorous analyses, designs, and forensic evaluations for a range of slopes, retaining walls, reinforced soil structures, deep excavations, bridge foundations, waste containment facilities, tailings dams, and embankment dams. He assisted expert witnesses in cases involving collapse and poor performance of earth retaining structures. He provided solutions to geotechnical problems in a number of environmental remediation projects involving cleanup of superfund sites. He conducted multi-phase flow analyses for several infrastructure features, including earthworks, embankment dams, and cover systems. Some of the consulting projects he participated in served the US Environmental Protection Agency, New York State Department of Environmental Conservation, New York State and Indiana Departments of Transportation, Tennessee Valley Authority, New Jersey Transit, and several multinational private and public corporations.

Islam A. Ebo

Islam Ebo is a marine and geotechnical engineer with experience in both industry and academic research. He obtained his bachelor's degree in civil engineering from Alexandria University in 2018 and his master's degree in civil engineering from California State University, Long Beach.

Islam worked on a variety of large-scale civil and geotechnical engineering projects with a focus on infrastructure development, particularly in marine and coastal environments. His work involves using advanced engineering tools to ensure the structural stability and environmental sustainability of these projects.

In his research role, Islam is focused on studying soil-structure interactions, coastal protection strategies, and the effects of environmental factors on marine infrastructure. He has developed strong skills in geotechnical analysis, project management, and field assessments, working closely with multidisciplinary teams to address engineering challenges.

Shiv K. Janardhanan

Shiv Karthee Janardhanan is a second-year graduate student in Information Systems at California State University, Long Beach, with a strong interest in data analytics and predictive modeling. His academic focus is on leveraging data and information systems to support informed decision-making and solve complex technical and organizational problems.

Shiv earned his bachelor's degree in computer science and engineering from India, where he built a solid foundation in software development, data processing, and analytical problem-solving. His graduate studies expand on this background through an emphasis on data lifecycle management, modeling, and applied analytics.

Through this research experience, Shiv has worked hands-on with complex datasets across the entire data lifecycle, from data collection and integration to cleaning, preprocessing, modeling, and evaluation, to develop predictive models for corrosion behavior. He is motivated by his goal of using data to create meaningful, positive change and to contribute solutions that improve systems, enhance sustainability, and support better engineering and organizational outcomes.

MTI FOUNDER

Hon. Norman Y. Mineta

MTI BOARD OF TRUSTEES

Directors