Quarterly Report – Public Page

Date of Report: 4th Quarterly Report-September 30th, 2020
Contract Number: 693JK31910018POTA
Prepared for: DOT PHMSA
Project Title: Mapping Indication Severity Using Bayesian Machine Learning from Indirect Inspection Data into Corrosion Severity for Decision-Making in Pipeline Maintenance
Prepared by: TEES (Texas A&M Engineering Experiment Station) and University of Dayton
Contact Information: Homero Castaneda, hcastaneda@tamu.edu, 979 458 9844.
For quarterly period ending: September 30th, 2020

1: Items Completed During this Quarterly Period:

Per the contract, Task 1 is associated with the fourth quarterly report. Task 2 include the experimental set up and activities which was concluded as planned. Due to patent and current results the laboratory activities will continue and will be included in the final report. The following activities have been completed:

Item	*Task*	*Activity/Deliverable*	*Title*	*Federal Cost*	*Cost Share*
9	1,2,3	4th Quarterly Report	4th Quarterly Report	4,000.00	0.00
4	1	Mapping available data via GIS tools and geographically co-register all datasets.	Establishing a database	20,000.00	10,000.00
6	2	Laboratory set up and electrochemistry mechanisms and corrosion assessment	Experiments and analyses to bridge gaps in prior knowledge	55,000.00	0.00

The title of the table is based on the file Technical and Deliverable Payable Milestone

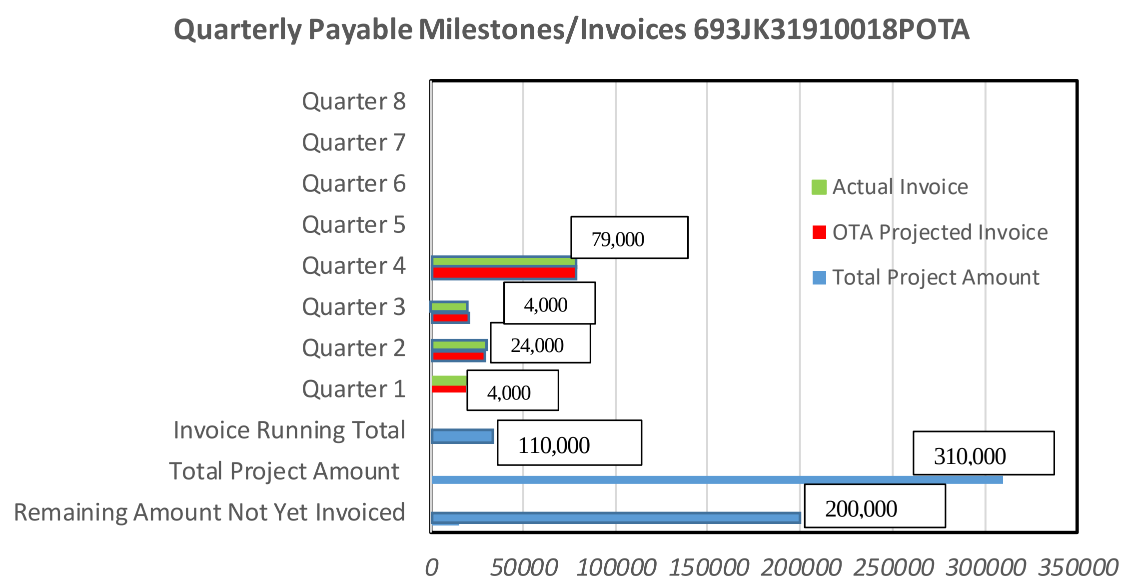
2: Items Not-Completed During this Quarterly Period:

Task number 2, extract basic corrosion model parameters started during the previous quarter report. Part of Task 2 will be cover in the following partial report. The following activities will be ready in latter reports based on the Technical and Deliverable Payable Milestone

Item #	*Task #*	*Activity/Deliverable*	*Title*	*Federal Cost*	*Cost Share*
8	2	Extract basic corrosion model and embed into the previously developed stochastic corrosion rate model framework	Experiments and analyses to bridge gaps in prior knowledge	24,000.00	0.00
10	3	Three unsupervised learning strategies (k-means, Gaussian Mixture Model, and Hidden Markov Random Field) for soil corrosivity clustering	Unsupervised learning strategies	20,000.00	3,000.00
11	3	Two supervised learning strategies (Support Vector Machine, Relevance Vector Machine) for defect type classification	Two supervised learning strategies	20,000.00	3,000.00

3: Project Financial Tracking during this Quarterly Period:

The table has been updated based on the deliverables and corrected attachment No5 *Technical and Deliverable Payable Milestone*.



4: Project Technical Status –

The following tasks are included in the project:

- *Task 1:* **Establishing a database**
- *Task 2:* **Experiments and analyses to bridge gaps in prior knowledge**
- *Task 3:* **Bayesian machine learning to bridge gaps in uncertainty quantification.**
- *Task 4* **Finalize and evaluate/validate the model.**

During the fourth quarter, the team members from Texas A&M University (TAMU) and the University of Dayton (UD) had weekly meetings and a Workshop with the sponsors. The aim of the workshop was to link, analyze and understand the field/ laboratory results and how we will use the data to analyze by using data driven techniques. In addition, the sponsor explain an update how they can use the algorithm to validate with their data for this project.

The team organized the mentioned Workshop entitled *"Indirect tools outcome as a parameter for severity and how we link the parameters with laboratory results"* with the *sponsor company*.

The outcomes of the workshop will help the PhD students and the engineers of the company to understand the corrosion mechanism and which parameters are used as primary precursors in the field following the understanding in the well control environment of the laboratory and how we can map the severity from several indirect and/or direct methods.

Description of the activities for this quarter.

Task 1 – Establishing a database.
Data handling for simulation purposes – In-line inspection data, soil survey, DCVG and CIS data for both pipelines (60 km and 110 km) have been finalized and analyzed and check for inconsistencies and consistency. Correction have being taken to resolve any issues with the data. A final Master file has been made with all this information combined to make the input for the Bayesian analysis as consolidated as possible. We started to analyze the field, environment (soil) parameters by using fundamental and semi empirical expressions to perform sensitivity analysis for parameters influencing the corrosion process and could be linked with the indirect inspection outcome. We added the parameters we found in laboratory conditions to align the outcome and distinguish the severity based on the surface conditions.

Task 2: **Experiments and analyses to bridge gaps in prior knowledge**

During this quarter, we performed the proposed experimental matrix to identify critical gaps in prior knowledge (i.e., current indirect survey (cathodic protection level), environmental conditions (soil characteristics) and other databases) and coded (or related) to deterministic and probabilistic modeling by following the corrosion mechanism. The correlation of different parameters influencing corrosion in field conditions has been considered in the laboratory set up and by the response in the laboratory experiments with mechanistic and electrochemical analysis. The results of the experimental testing could revealed the potential difference between intact (or no defect), active surface due to a coating defect and passive surface with corrosion products formed due to a coating defect.

Experiments and analyses to bridge gaps in prior knowledge and current response.
Different relationships and correlation could be found with the fundamentals of electrochemistry and electrical signals by following the transfer function techniques available in the laboratory that can be use in the field. The relationships are developed from different techniques applied to the system simulated system.

The performed set of laboratory experiments include the effects of soil resistivity (or conductivity), pH and the metallic surface condition in the presence of holidays (specifically intact, active and passive state) under cathodic protection. The experimental design performed is presented in Table1. Buffer solution (with defined conductivity and TDS Standards) is applied to adjust solution resistivity. The passive holiday can be realized by external anodic current via potentiostat (Gamry, The Interface 600plus™). NS4 solution with composition (unit: g/L) of KCl: 0.122, $NaHCO_3$: 0.483, $CaCl_2$: 0.093 and $MgSO_4$: 0.131 is used to simulate soil conditions.

Sequence of non-destructive techniques for parameters that are equivalent either with DCVG or CIPS measurements; the methods include: the evolution of open circuit potential

(OCP) or cell potential, bias or polarized potential, bias Electrochemical impedance spectroscopy (EIS), and surface analysis will be carried out in sequence to measure the response of experimental set-ups for the analysis of CIS and DCVG database. For the electrochemical measurement system, the three-electrode method will be used.

Samples	Coatings Thickness	Soil Composition	Distribution of soil (resistivity)	Severity based on active-passive concept	pH
AISI 1008/API X52	10-20 mils	NS4	1036 ohm cm	Active Holiday	4
AISI 1008/API X52	10-20 mils	NS4	1036 ohm cm	Active Holiday	7
AISI 1008/API X52	10-20 mils	NS4	1036 ohm cm	Active Holiday	10
AISI 1008/API X52	10-20 mils	NS4	1036 ohm cm	Passive Holiday	4
AISI 1008/API X52	10-20 mils	NS4	1036 ohm cm	Passive Holiday	7
AISI 1008/API X52	10-20 mils	NS4	1036 ohm cm	Passive Holiday	10
AISI 1008/API X52	10-20 mils	NS4	714 ohm cm	Active Holiday	4
AISI 1008/API X52	10-20 mils	NS4	714 ohm cm	Active Holiday	7
AISI 1008/API X52	10-20 mils	NS4	714 ohm cm	Active Holiday	10
AISI 1008/API X52	10-20 mils	NS4	714 ohm cm	Passive Holiday	4
AISI 1008/API X52	10-20 mils	NS4	714 ohm cm	Passive Holiday	7
AISI 1008/API X52	10-20 mils	NS4	714 ohm cm	Passive Holiday	10
AISI 1008/API X52	10-20 mils	NS4	870 ohm cm	Active Holiday	4
AISI 1008/API X52	10-20 mils	NS4	870 ohm cm	Active Holiday	7
AISI 1008/API X52	10-20 mils	NS4	870 ohm cm	Active Holiday	10
AISI 1008/API X52	10-20 mils	NS4	870 ohm cm	Passive Holiday	4
AISI 1008/API X52	10-20 mils	NS4	870 ohm cm	Passive Holiday	7
AISI 1008/API X52	10-20 mils	NS4	870 ohm cm	Passive Holiday	10

Table 1 Experimental design matrix for electrochemical measurements

Table 1 includes the simulation conditions of the field and the parameters that we consider first level or critical for the detection and characterization parameters needed and the severity mapping. The active included a metallic surface with a defined defect area with all parameters in the table simulating the most sensitive. The severity simulated marked three conditions, the intact coating condition, the active and the passive condition. The sequence of methods performed for each condition resulted in different outcomes and concepts for the severity mapping.

Table 2 includes the experimental techniques used for the quantitative analysis of the severity of the metallic surface under soil simulation conditions. The parameters that can connected the field and laboratory conditions cover the chemical, physical and electrical properties. The outcome is set in tables for the analysis of different precursors for corrosion.

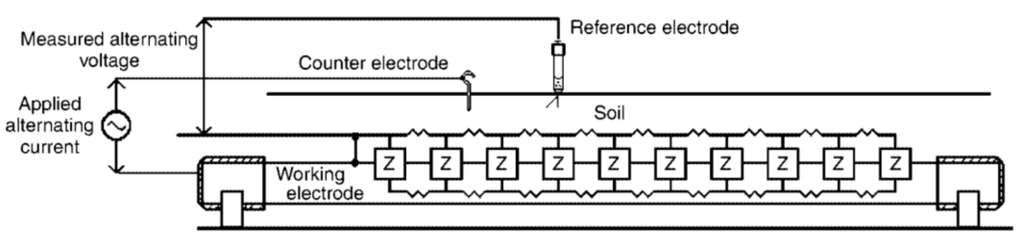
Experimental technique	***Outcome parameters***	***Correlation to the field interpretation***
Resistivity of the soil EIS and measured magnitude	Soil conditions	Resistivity vs. IR

OCP and EIS Potentiostat and EIS	Surface mechanisms, corrosion rate	Severity and criterion Ranking for correlations
EIS Bias potentiostatic EIS	Corrosion rate	Potential gradient vs severity
Surface/electrolyte chemical analysis Potentisotatic	Influential parameters for severity	Direct and indirect parameters to categorize severity

Table 2 Experimental techniques and outcomes for the field interpretation with laboratory results.

Deterministic approaches based on constitutive equations and continuous electrical elements have been used to describe the current distribution under corrosive conditions when there is a coating defect on a metallic substrate at the interface when an electrochemical cell is formed. We previously had developed a transmission line model (TLM) in 1D and 2D models to characterize and quantify each element in the electrochemical cell, as illustrated in Fig. 1(a) and 2(b). EIS technique is able to capture several magnitudes characterizing not only the magnitude of each element but the surface condition. The generated information is able to add another layer in the machine learning model framework. The elements can be related to the severity of the surface by means of corrosion rates or parameters that are indirectly measured and estimated via model inference.

The multidimensional approach covers mechanistic processes that occur when a metallic material includes an anomaly or different conditions of the coating and immersed in electrolyte and different parameters characterized the electrochemical elements. In addition, the influence of AC signal flowing through external transmission lines on local pipeline defects can be represented as that of a transmission line, as shown in Fig. 6



(a)

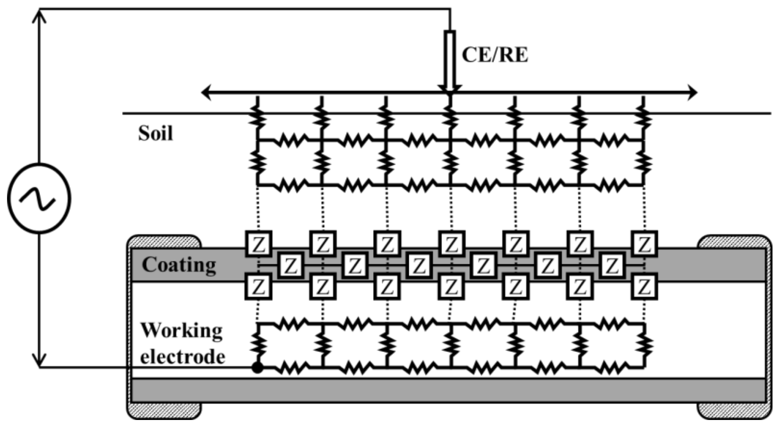


Figure 1. TLM in 1 D (a) and 2D for the coated pipeline (b)

Task 3: Bayesian machine learning to bridge the gaps in uncertainty quantification.

Clustering is a major and widely accepted technique which can be used to separate the data into distinct homogeneous groups based on their similarity. These have been previously employed and found to be an effective tool to assess the soil corrosivity for external integrity management of a pipeline structure. Hidden Markov random field (HMRF) is used to identify hidden patterns in the data (e.g., different soil categories in the current context) when a response variable is not explicitly given, which in this project is the corrosion levels at each pipeline segment due to the spatial variation of the soil environment and regional environmental factors. Based on the clustering results and the corrosivity assessment, different maintenance strategies can be applied to different segments of a pipeline structure. More specific, the results of clustering along with direct and indirect inspection data can be used to determine the severity of corrosion corresponding to each soil category. The clustering result for the 110km pipeline based on 7 principal component features is as shown in Figure 2 and Figure 3 shows the cluster plot with two of the principal component overlaid with contour plot showing the probabilistic distribution of the two principle components.

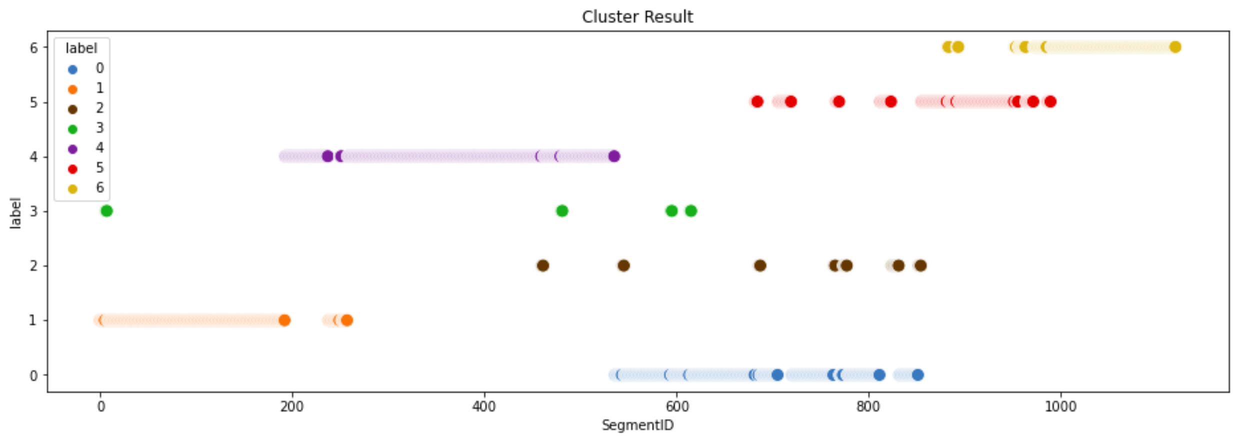


Figure 2. Clustering Result along the right-of-way

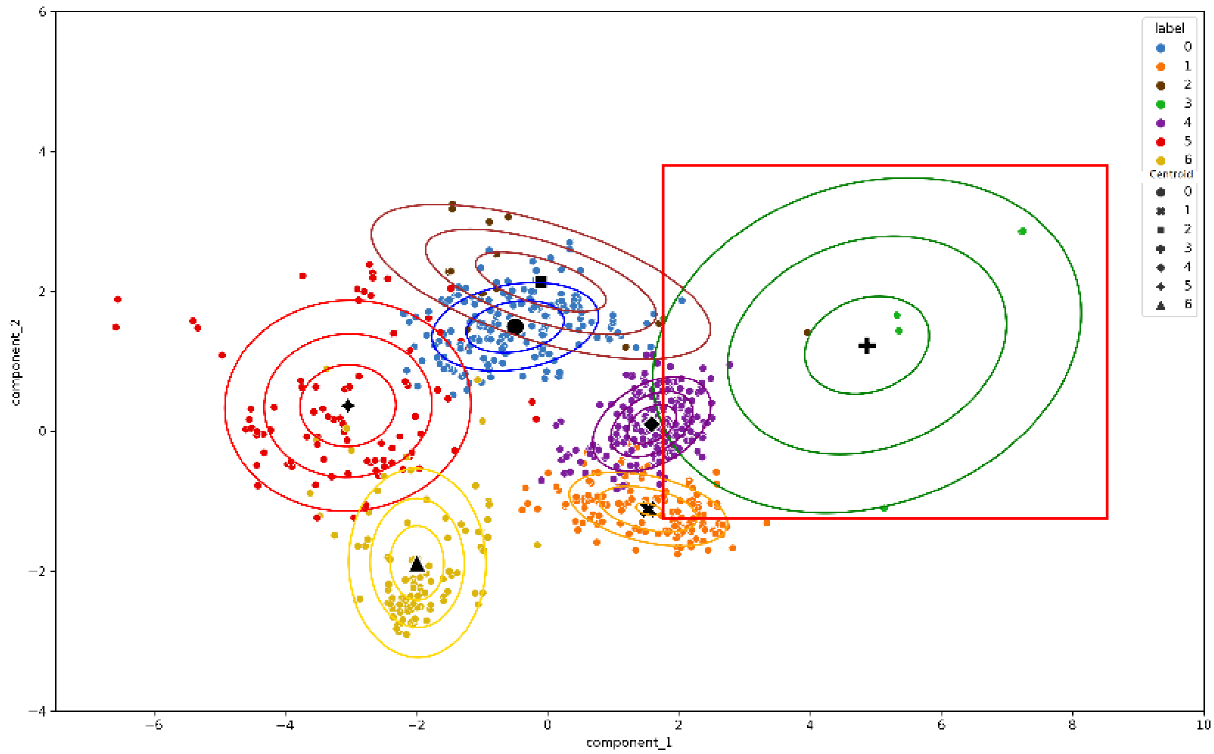


Figure 3. Scatter plots of two principal components with centroid and contour for 7 clusters

As marked in Figure 3, cluster 3 data points are outlier points since the outlier points do not form a dense cluster and they are sparsely distributed along the right-of-way. These segments were removed from the current analysis and will be investigated individually in future. The best model fit was re-evaluated using BIC criteria as shown by Figure 4. From the BIC Vs k (number of components) plot we get optimal number of clusters as 6.

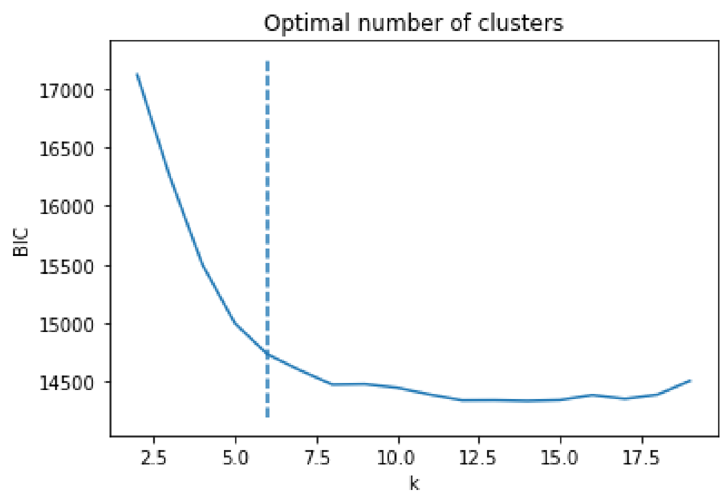


Figure 4. BIC Vs k (number of components)

The new cluster analysis result is as shown in Figure 5. Figure 6 shows the plot between the first two principal components with centroid and contour overlaid with cluster distribution. This clustering result is used to group features which is shown in Figure 7.

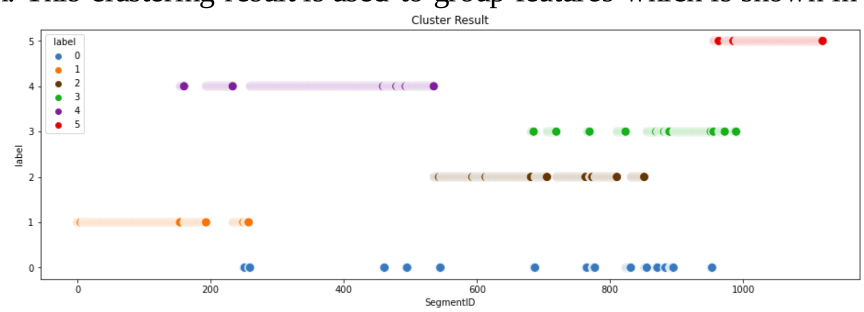


Figure 5. Cluster Result with six clusters

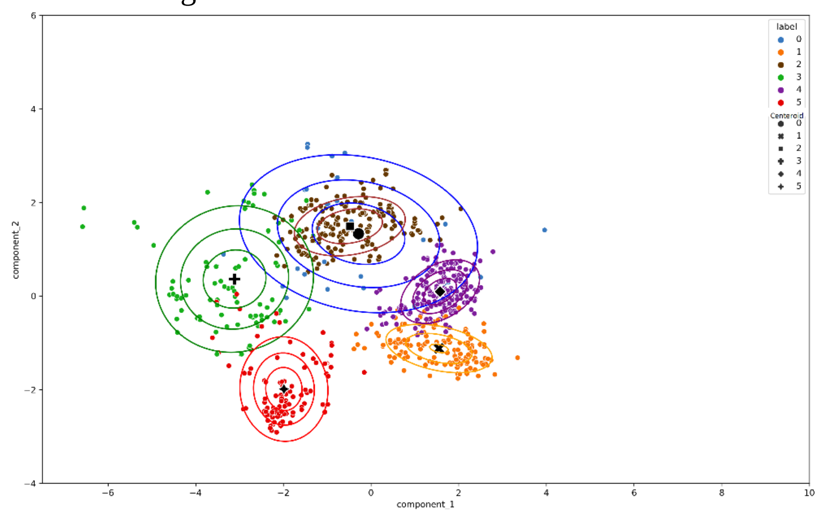


Figure 6. Scatter plots of two principal components with centroid for 6 clusters

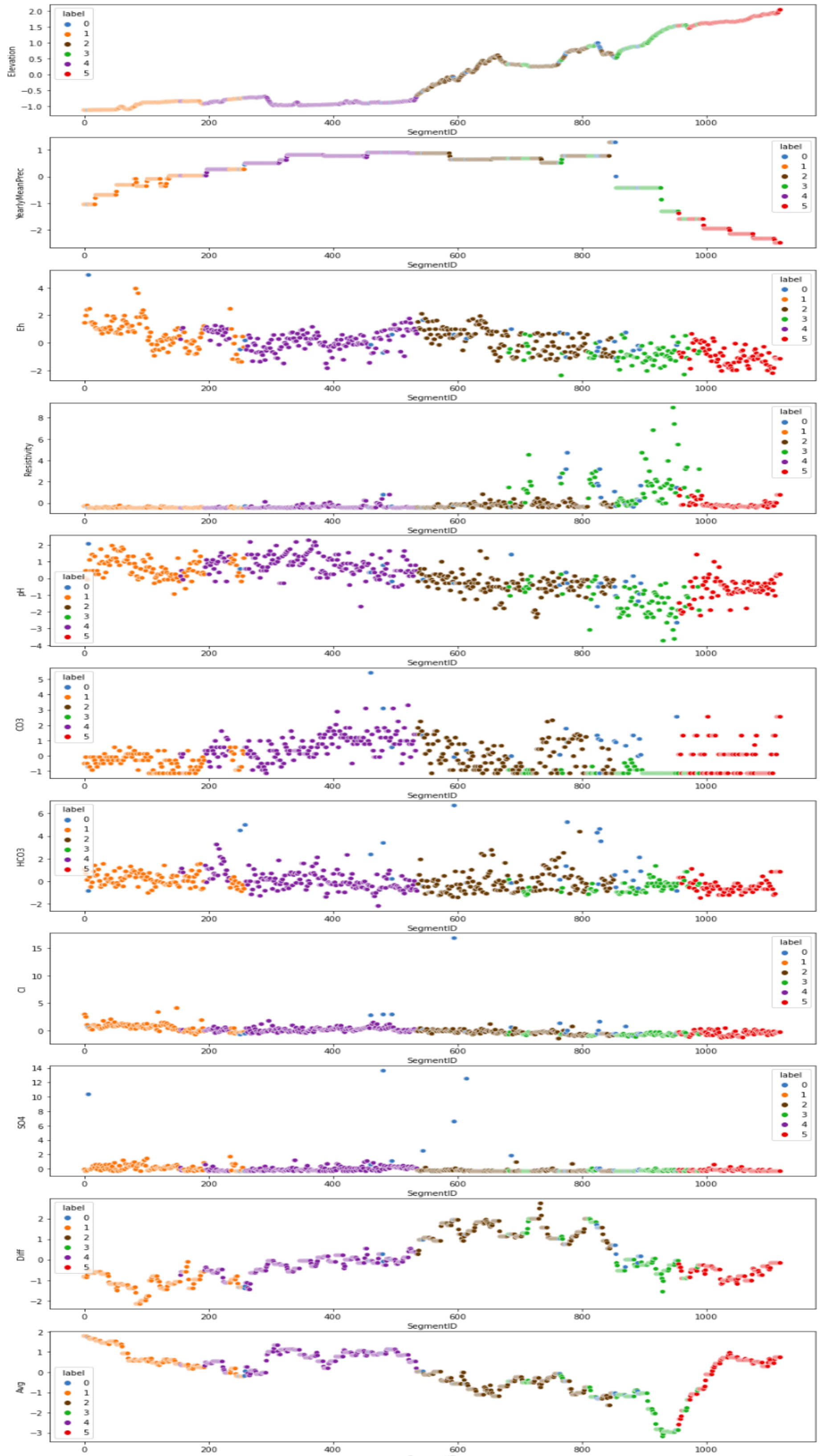


Figure 7. Cluster groups of soil features

Figure 8 shows the distribution of maximum corrosion depth in each cluster for 2005 ILI data and 2010 ILI data respectively. The progression (in a statistical means) of corrosion over the years is seen clearly for different clusters.

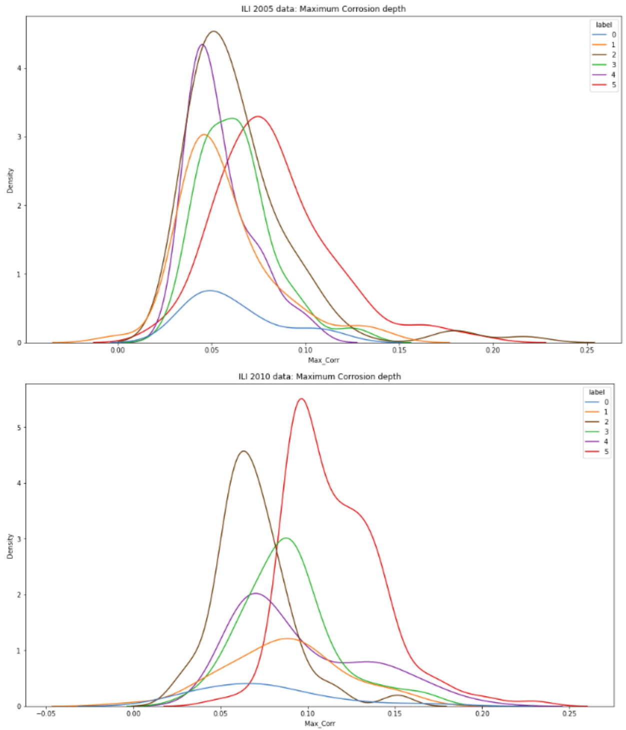


Figure 8. Probability distribution of corrosion points with maximum depth in each cluster

Figure 9 shows the distribution of number of corrosion defects_in each cluster for 2005 ILI data and 2010 ILI data respectively. The increase in number of corrosion spots can be identified for different clusters. Maximum number of corrosion defects can be observed in the cluster 5 corresponding to the sections with segment ID from 950 to the end of pipeline.

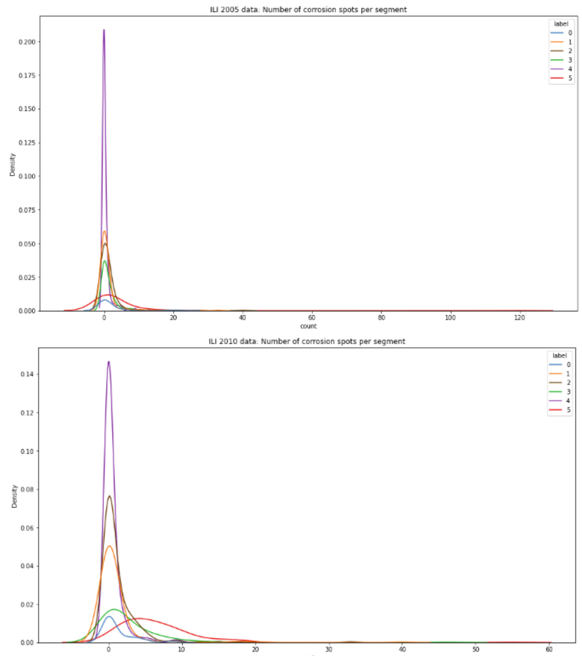


Figure 9. Probability distribution of number of corrosion points in each cluster

Figure 10 and 11 shows the indirect inspection data assigned to corresponding clusters. The break seen in DCVG data is due to the presence of a river in the pipeline right-of-way.

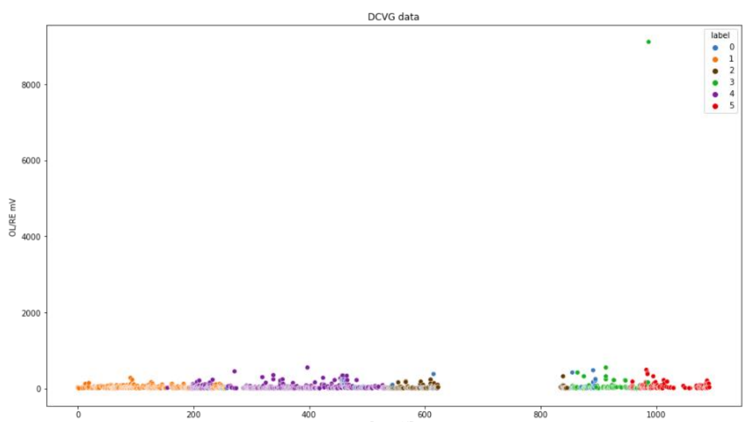


Figure 10. DCVG voltage along the right of way

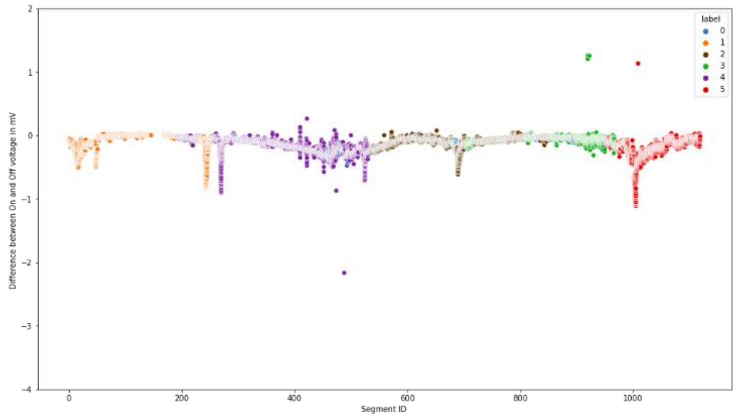


Figure 11. Difference between CIS ON and OFF potential along pipeline right of way

Now the data has been clustered along the pipeline segment and the next step is to analyze the severity of corrosion in each cluster/ pipeline segment by linking the direct (in-line inspection) and indirect (DCVG and CIPS) inspection data and combining the findings from the lab results. Detailed work is currently on going.

5: Project Schedule –

The project is on-schedule as originally-proposed.

During the following report, the team will perform Extract basic corrosion model and embed into the previously developed stochastic corrosion rate model framework, the correlations between field measurements and parameters founded in laboratory conditions to extract the information required for the machine learning method. Supervised and unsupervised machine learning analysis (to be incorporated into the final report).

6. Publication

On September 30th, 2020, the conference manuscript was accepted to the NACE 2021 Conference to be held in Salt Lake City Utah USA. A patent will be filed within few weeks for the sensing techniques used for surface severity of buried pipelines.

Observations: The experimental set up and procedure was delayed due to the restrictions existed in the laboratory. The planned experiments were completed. There is analysis that should be done to validate the experimental testing, this analysis has been delayed for the next activity.