Joint CSCE Construction Specialty & CRC Conference 2025
Conférence conjointe spécialisée en construction de la SCGC et CRC-2025

Montreal, Quebec
July 28-31, 2025 / *28-31 juillet 2025*

REGIONAL VARIATIONS IN PIPELINE LONGEVITY: INSIGHTS FROM A SPATIAL AND SURVIVAL ANALYSIS APPROACH

Asaye, L.[1], Le, C.[2*], Huang, Y.[1], Le, T.[3], Yadav, O.P.[4], Pirim, H.[5], and Le, T.[6]

[1] Department of Civil, Construction, and Environmental Engineering, North Dakota State University, Fargo, ND 58102, USA
[2] Department of Engineering Technology and Construction Management, University of North Carolina at Charlotte, Charlotte, NC 28223, USA (Corresponding Author)
[3] Department of Industrial and Management Systems Engineering, University of South Florida, Tampa, FL 33620, USA
[4] Department of Industrial and Systems Engineering, North Carolina Agricultural and Technical State University, Greensboro, NC 27411, USA
[5] Department of Industrial and Manufacturing Engineering, North Dakota State University, Fargo, ND, USA
[6] Glenn Department of Civil Engineering, Clemson University, Clemson, SC 29634, USA

ABSTRACT: Pipelines face significant challenges due to unique environmental, operational, and regulatory variations across different geographic regions, which critically influence their longevity. While previous studies have examined factors influencing pipeline failure, they have predominantly offered generalized survival models, giving limited attention to the role of regional variations in pipeline survival rates. This study addresses this critical gap by developing a novel approach that combines spatial analysis (Getis-Ord Gi*) to identify areas with varying concentrations of incidents and survival analysis (Prentice-Williams-Peterson models and Cox Proportional Hazards) to model time-dependent failure risks, offering insight into how geographic factors influence the likelihood and timing of failures. The analysis leverages oil incident data from the Pipeline and Hazardous Materials Safety Administration and network data from the U.S. Energy Information Administration, focusing on carbon steel pipelines installed between 2010 and 2022. The analysis identifies hotspot areas where pipelines experience accelerated failure rates. Furthermore, the Cox PH model, with a concordance index of 0.73, quantifies the hazard ratios associated with the risk factors. The hotspot pipeline segment, for instance, shows a notably high risk of failure, with a hazard ratio of 2.089, indicating that pipelines in these areas are twice as likely to fail compared to cold spot areas. Furthermore, regionally, the Midwest exhibits significantly lower survival probabilities and faces a fourfold increase in failure risk. This study contributes to understanding the spatial dynamics of pipeline performance, offering insights for regionally tailored maintenance strategies and material selection to improve pipeline longevity and infrastructure reliability.

1. INTRODUCTION

The longevity and reliability of pipelines are critical components of the energy infrastructure, ensuring the safe and efficient transportation of oil. In the United States, pipelines form the backbone of energy distribution, spanning vast and diverse geographic regions with unique environmental, operational, and regulatory challenges (Martínez et al., 2025). These variations can significantly influence pipeline survival rates, presenting a complex risk landscape that requires detailed investigation and tailored solutions. While pipelines are designed to operate for decades, localized factors such as weather conditions, maintenance practices, and regulatory enforcement can accelerate their degradation and lead to pipeline failures (Asaye, Le, Huang, et al., 2025; Moriyani et al., 2024). Regional variations in pipeline performance are particularly noticeable in the United States, where disparate climates and operational practices introduce variability in risk factors and failure rates (Asaye et al., 2024; Asaye, Le, Le, et al., 2025).

Previous studies on pipeline failure have primarily focused on identifying failure risks and understanding the degradation mechanisms that influence the lifespan of pipelines (Wasim & Djukic, 2022; Zha et al., 2022). For instance, Ezuber et al., (2021) has revealed that environmental conditions, including moisture levels, soil chemistry, and temperature, can accelerate pipeline degradation, especially in regions with harsh climates. Similarly, mechanical stress from operational pressures and fluctuating flow rates often exacerbate pipeline aging (Amaya-Gómez et al., 2019). Additionally, Bai et al., (2017) have pointed out the impact of external factors such as flooding, landslides, and third-party interference on pipeline integrity. Naik & Kiran, (2018) highlight regional disparities in pipeline failures, with higher accident densities and varying failure types in the southern, northeastern, and midwestern U.S. On the other hand, while previous studies have applied survival analysis to investigate pipeline failure risks, they have often overlooked the influence of geographic variability on pipeline longevity (Xiao et al., 2024). Existing studies often provide generalized findings without addressing how these factors vary regionally. The understanding of pipeline failure risks and their geographic variability remains underdeveloped, and regional disparities in pipeline longevity remain largely unexplored.

This study addresses a critical gap in literature by assessing differences in pipeline survival rates across diverse geographic regions, using the Prentice-Williams-Peterson (PWP) model to analyze recurrent failures on the same pipeline segment and the Cox Proportional Hazards (Cox PH) model to estimate hazard ratios for pipeline characteristics and spatial factors. By leveraging comprehensive data from the Pipeline and Hazardous Materials Safety Administration (PHMSA), this research identifies and compares hotspot and non-hotspot areas across the U.S., providing actionable insights into localized degradation patterns and risk factors. The findings reveal significant disparities in survival rates and underscore the need for region-specific strategies in pipeline design, material selection, and maintenance. These insights are crucial for optimizing resource allocation and improving the safety, reliability, and overall performance of pipeline systems in high-risk regions. By examining regional variations in pipeline longevity, this paper contributes to the growing body of knowledge on infrastructure resilience and risk management, providing a foundation for policymakers, engineers, and stakeholders to address the unique challenges of different geographic contexts.

2. METHODOLOGY

Figure 1 illustrates the novel framework proposed in this study that combines spatial and survival analysis to capture spatial concentration and assess cumulative failure risks, offering insights into the regional variations in pipeline longevity. Initially, data on pipeline incidents were collected from the PHMSA to capture information on failure events. Additionally, characteristics of pipeline segments, such as pipeline length, were gathered from the U.S. Energy Information Administration (USEIA). During the data preprocessing phase, incidents were grouped by pipeline segment to ensure that each segment's failure history was considered comprehensively. Missing values were imputed using the median for continuous variables and the mode for categorical variables. Outliers were identified and removed using the interquartile range method to maintain data integrity. Following data preparation, spatial analysis was conducted to identify pipeline segments, with certain locations showing a significantly higher concentration of incidents, known as hotspots, while others exhibited lower concentrations, referred to as cold spots. Subsequently, events were sorted chronologically to track the sequence of failures, with hotspots and cold spots identified through spatial analysis and incorporated as categorical variables in the survival models. The study employed the PWP model to determine survival probability by accounting for recurrent failures within the same pipeline segment and the Cox PH model to quantify the impact of factors on failure risk. Statistical techniques such as the Log-Rank Test were used to compare survival probabilities across different groups, such as geographic regions and states. Hazard ratios were calculated to assess the relative risks of each factor for failure, and the Concordance Index was used to evaluate the model's predictive accuracy. The results of these analyses provide a comprehensive understanding of risk factors and regional variations in pipeline longevity. Insights from the hazard ratios and survival probability estimates will help inform future maintenance strategies, as well as long-term planning. This methodology lays the foundation for a data-driven approach to improving the reliability and safety of pipeline systems across diverse geographies.

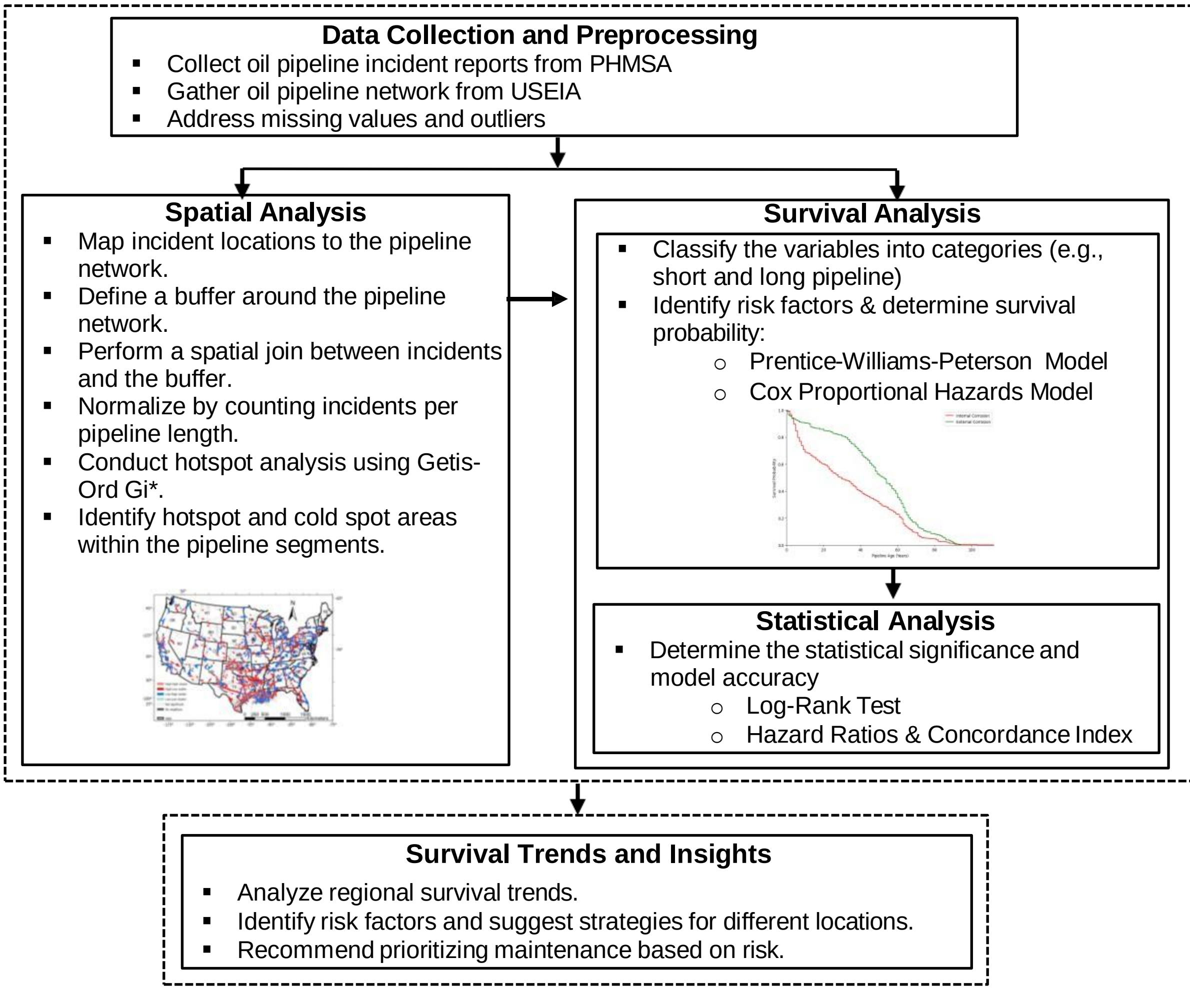

Figure 1: Framework for Analyzing Spatial Variation in Oil Pipeline Longevity

2.1 Data Collection and Preprocessing

This study analyzes data on oil pipeline segments installed between 2010 and 2022 in the US, along with incidents occurring during the same period, sourced from the PHMSA database. This data helps to characterize the survival times of pipeline segments and their likelihood of failure over time, allowing for the calculation of survival probabilities. To ensure the robustness of the analysis, outliers were identified and removed using the interquartile range method, with any data points located more than 3.44 km from the nearest pipeline segment being excluded, accounting for 2.16% of the data. This threshold was determined using the standard IQR rule, where distances exceeding 1.5 times the interquartile range above the third quartile were classified as spatial outliers. For the remaining records, missing values were imputed using the median for continuous variables and the mode for categorical variables. Following the data cleaning process, the analysis includes detailed records of 1,329 incidents, which capture both pipeline characteristics and spatial factors. The analysis specifically targets carbon steel pipeline segments, which represent over 78.16% of the data, to ensure consistency when comparing survival probabilities across similar materials. To complement this data, pipeline network attributes, including geographic information, are obtained from the USEIA and extracted using Geographic Information System layers. Table 1 shows a sample of the raw input data, including installation and incident occurrence times. Each pipeline segment is uniquely identified by pipe ID, enabling detailed segment-level analysis of failure patterns and the calculation of survival probabilities over time.

Table 1: Sample of key pipeline incident data attributes

Pipe ID	Installation year	Incident identified datetime	Pipe diameter (cm)	Pipe wall thickness (mm)	…	Pipe coating type
1	2014	7/9/2017 7:00	60.96	9.525	…	Fusion bonded epoxy
1	2014	8/11/2021 10:30	60.96	9.525	…	Fusion bonded epoxy
…	…	…	…	…	…	…
556	2016	9/16/2021 9:30	91.44	12.7	…	Paint
556	2016	9/25/2021 10:30	91.44	12.7	…	Paint

2.2 Spatial Analysis

Spatial analysis refers to the process of examining the distribution and patterns of events within a defined geographic area. In the context of pipeline incident analysis, it involves mapping pipeline incidents to the pipeline network using GIS tools and defining a buffer around the network. A spatial join is then performed to connect incidents to the buffer and count the number of incidents on each pipeline segment, with incident data normalized by calculating the number of incidents per unit length of the pipeline. Hotspot analysis is then conducted using the Getis-Ord Gi* statistic to identify segments classified as hotspots ($z \geq 1.96$, $p < 0.05$) or cold spots ($z \leq -1.96$, $p < 0.05$), which serve as input variables for subsequent survival analysis.

2.3 Survival Analysis

Survival analysis is a statistical method used to analyze the time to an event—in this study, the time to pipeline failure. This study applies two complementary models: the PWP model and Cox PH models, applied sequentially to assess failure trends and influencing factors. These models were selected for their ability to handle recurrent failure events while providing interpretable outputs, such as hazard ratios. While alternatives like parametric models or machine learning approaches were considered, they were not adopted due to stronger distributional assumptions and limited interpretability for understanding risk factors.
PWP Model*:* The PWP model is used to identify survival trends over time and to account for recurrent failure events within the same pipeline segment. It incorporates installation time up to the event to assess failure timing and long-term degradation. The total time for the i^{th} event is calculated as shown in Equation (1).

[1] $t_i^{Total} = t_i - t_0$

Where:

t_i^{Total} : Total time for the i^{th} event

t_i : Time of the i^{th} event

t_0 : Installation date (start time of observation)

The survival probability for the i^{th} incident at time t is given by Equation (2):

[2] $S_k(t) = \exp(-\int_0^t \lambda_0(u)\exp(\beta X_i, i)\, du$

Where:

$S_k(t)$: Survival probability for the i^{th} incident at time t.

$\lambda_0(t)$: Baseline hazard function for cumulative time.

$\exp(\beta X_i, i)$: Effect of covariates X_i for the i^{th} failure.

To model recurrent failures using the PWP approach, failure events were chronologically ordered for each pipeline segment to capture the sequence and timing of degradation. Time intervals were calculated from the installation date to each failure event, allowing the model to account for temporal progression. To maintain consistency with the recurrent-event structure of the PWP model, only segments with multiple recorded failures were included in the analysis.

The analysis considers the following factors:

1. *Pipeline Segment Status*: Hotspot or cold spot pipeline segment identified through spatial analysis, influencing the geographic distribution of failures.

2. Pipeline length: Pipelines were classified as "Short" (≤550 m) or "Long" (>550 m) using the median, selected for its balanced split and resistance to outliers. Other approaches were considered, but the median provided the most interpretable and statistically stable grouping for this analysis.
3. *Geographical Regions*: Incidents are further classified into four regions based on U.S. Census Bureau divisions (U.S. Census Bureau, 2024)—Northeast, South, Midwest, and West-with pipeline density(Total Pipeline Length (km) / Total Area (km^2)) calculated for each region to evaluate how density variations influence pipeline performance.
4. *State*: The study considers differences between states, including local regulations, maintenance practices, environmental conditions, and calculated pipeline density at the state level.

Cox PH Model: After the PWP model identifies survival trends, the Cox PH model estimates hazard ratios for various factors (e.g., geographical location and pipeline characteristics) that influence pipeline failure risks. This model provides a detailed understanding of how each factor impacts the likelihood of failure.

2.3.1 Statistical Methods

Several statistical methods were employed to assess the significance and interpret the results:

- *Log-Rank Test:* A non-parametric test used to compare survival distributions between groups, such as between hot and cold spot areas. It was selected for its ability to detect differences without assuming a specific survival time distribution.
- *Hazard Ratios:* Reflects the risk of failure associated with a one-unit change in a covariate. An HR greater than 1 indicates a higher risk of failure, while a ratio less than 1 suggests a reduced risk.
- *Concordance Index:* The Concordance Index is a measure of how well the survival model predicts the order of survival times. A C-index of 1.0 indicates perfect prediction.

2.4 Survival Trends and Insights

The PWP model analyzes recurrent failure events and identifies regional degradation patterns, highlighting the cumulative effects of repeated failures. The Cox PH model quantifies the impact of various risk factors, including regional variations, on pipeline failure risk. The combined application of these models provides valuable insights into pipeline longevity, emphasizing the importance of region-specific strategies for pipeline design, material selection, and maintenance to enhance infrastructure resilience.

3. RESULT AND DISCUSSION

This section provides insights into regional variations in pipeline longevity across the U.S., highlighting the pipeline characteristics and spatial factors that significantly influence pipeline degradation. It highlights major risk factors and offers recommendations to enhance pipeline resilience.

3.1 Spatial Analysis

The spatial analysis identified hotspot and cold spot pipeline segments. While cold spots have fewer incidents, they still require monitoring, underscoring the need for region-specific maintenance.

3.2 Survival Analysis

The survival analysis provided critical insights into geographical variations in pipeline longevity. It revealed that pipeline failure rates vary significantly depending on spatial factors such as state and region and location-specific strategies to manage pipeline risks effectively and ensure long-term system reliability.

Hotspot and Cold spot Pipeline Segments.

Figure 2 illustrates the survival probability curves for pipeline segments in Hot Spot and Cold Spot areas. The Log-Rank test reveals a p-value of 0.0279, suggesting a statistically significant difference in survival rates between the two groups. The proposed recurrent failure event survival probability curve is important for capturing the cumulative effect of repeated failures in hotspot regions. Pipelines in Hot Spot areas exhibit a steeper decline in survival probability compared to those in Cold Spot areas. For example, at 1500 days, the cumulative time from pipeline installation to the incident, the survival probability for pipelines in Hot Spot areas is approximately 0.2, significantly lower than the 0.4 survival probability for Cold Spot pipelines, indicating a higher failure rate for the Hot Spot segments. This trend may be attributed to factors such as more aggressive environmental conditions, operational stresses, or regional variations in maintenance practices. The findings underscore the need for region-specific maintenance strategies. Hot Spot areas, with their higher rate of failure, may face heightened operational challenges and require more focused monitoring and maintenance to mitigate the risk of accelerated degradation, while Cold Spot areas, though exhibiting fewer failures, still necessitate ongoing monitoring due to the potential for less frequent but more catastrophic incidents.

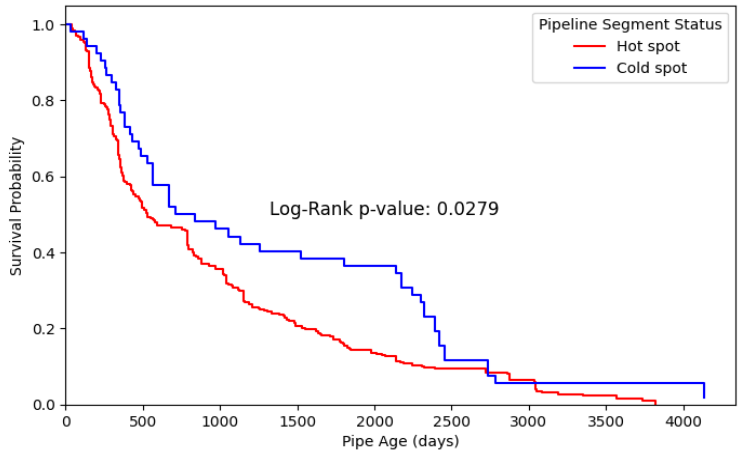

Figure 2: Survival probability curves for pipeline segments in Hot Spot and Cold Spot areas

Effect of Pipeline Length

Figure 3 compares the survival probabilities of short and long pipelines over time. The analysis reveals that short pipelines tend to have a slightly higher survival probability than long pipelines, with a marginally significant difference (p-value = 0.0546), highlighting the potential relationship between pipeline length and survival probability. This suggests that pipeline length, along with other risk factors, should be considered in future studies to better understand failure risks. Potential reasons for the observed trend include shorter pipelines that might experience less stress due to fewer connections or operational loads, leading to fewer failure events. Conversely, longer pipelines may experience increased pressure and wear from extended operational loads, and their larger size can make monitoring and maintenance more challenging, potentially exacerbating the risk of failure and lowering survival probabilities.

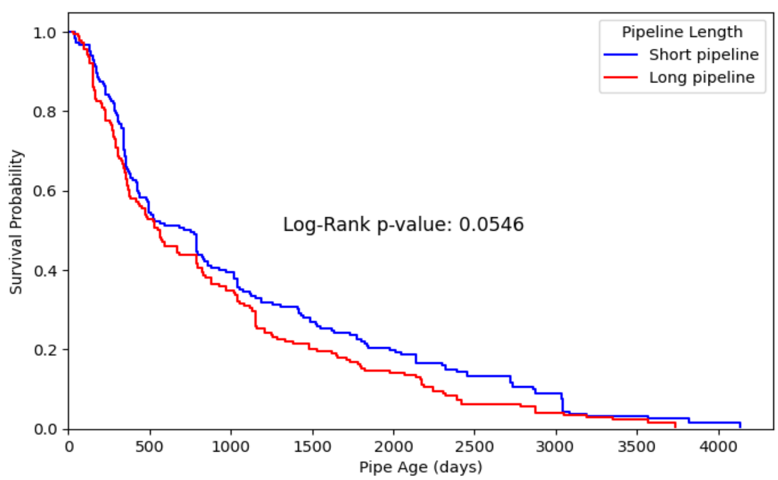

Figure 3: Survival probability curves for short and long pipeline segments

Geographical region. Figure 4 shows survival probability curves for pipeline segments across various U.S. regions, revealing significant disparities in pipeline longevity. Log-Rank test p-values in Table 2 indicate significant differences between certain regions, such as the Midwest and Southeast (p = 1.65E-08). Specifically, pipelines in the Midwest exhibit the lowest survival probabilities, suggesting that systems in this region might be exposed to more severe degradation or poor maintenance practices due to extreme harsh weather conditions. In contrast, the South regions show higher survival probabilities, attributed to milder climates and better maintenance practices. However, not all region comparisons show significant differences. For example, the Northeast and West comparisons (p = 0.894) reveal no significant difference in survival probabilities, indicating that some regions may not exhibit distinct variations in pipeline longevity. These findings contribute to the growing body of literature on regional disparities in pipeline performance.
An analysis of pipeline density distribution across regions indicates that areas with higher infrastructure concentration, such as the Midwest, tend to exhibit lower survival probabilities, suggesting that density may contribute to increased risk. This is likely due to the operational pressures in dense areas, where higher wear and tear, and more frequent incidents can exacerbate pipeline degradation. For instance, Naik & Kiran, (2018) identified regional differences in pipeline accident rates, noting higher accident densities in the southern U.S., with varying failure types across the south and Midwest. Unlike previous studies that primarily focused on failure causes without accounting for regional variations, this study provides a quantitative analysis of survival differences across U.S. regions using spatial survival analysis. These findings underscore the importance of region-specific maintenance strategies, such as increased inspection frequency and weather-resistant coatings, to address varying environmental and operational factors.

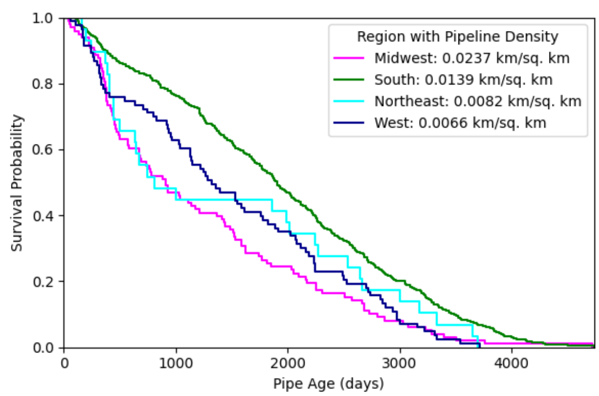

Figure 4: Survival probability curves for pipeline segments in different U.S. regions

Table 2: Log-Rank test p-values for region comparisons

Region 1	Region 2	p-value
Midwest	South	1.65E-08
Midwest	Northeast	0.386
Midwest	West	0.108
South	Northeast	0.041
South	West	0.001
Northeast	West	0.894

State Comparison. When comparing different states, variations in pipeline longevity are evident, particularly when examining Texas and North Dakota, as shown in Figure 5. The Log-Rank Test (p = 0.005) indicates a statistically significant difference between the two states, with a hazard ratio of 0.52 suggesting that pipelines in Texas have a lower risk of failure. This finding is significant, as HR quantifies the relative risk of failure, providing a measurable metric to assess the influence of environmental and operational factors on pipeline performance. Consequently, this information enables more efficient prioritization of maintenance resources, with North Dakota pipelines potentially requiring more frequent inspections and facilitating the implementation of targeted maintenance strategies that address specific seasonal needs. Pipelines in Texas exhibit higher survival probabilities, likely due to milder climates. In contrast, North Dakota pipelines experience lower survival probabilities, possibly due to harsh winters, extreme temperature fluctuations, and fewer resources for maintenance. While the regional analysis highlights broader trends—such as the overall low survival in the Midwest—the state-level comparison between Texas and North Dakota adds critical detail granularity, revealing that not all states within a region follow the same pattern**.** Even though hotspots were detected in both states, the survival probability remains higher in Texas compared to North Dakota. This outcome, despite Texas having greater pipeline density, suggests that density alone does not account for survival differences at the state level. Variations may also result from environmental conditions, infrastructure practices, or maintenance strategies, highlighting the complex factors influencing pipeline longevity.

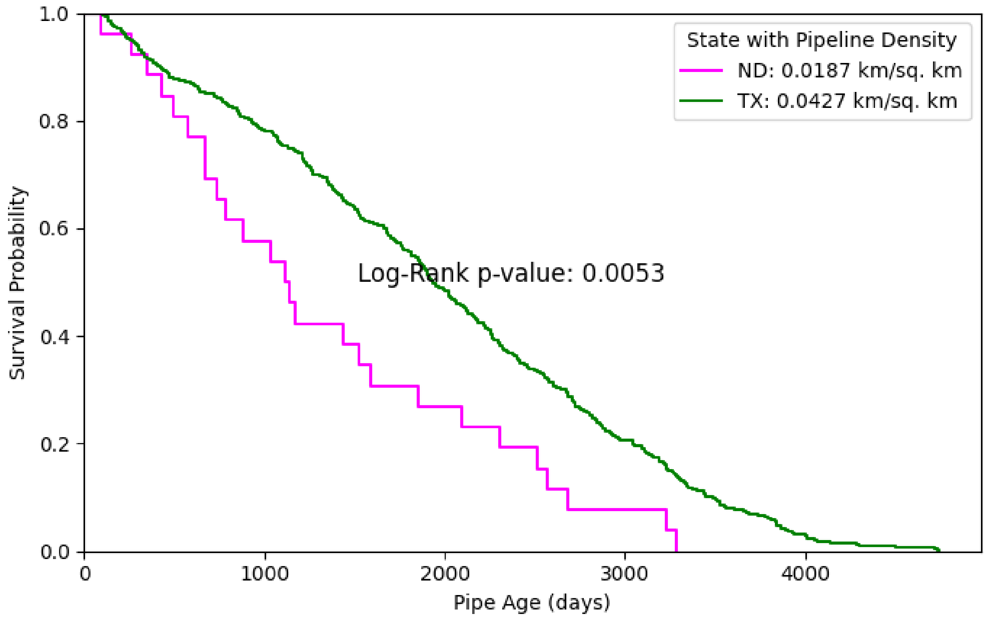

Figure 5: Survival probability curves for pipeline segments in Texas and North Dakota

Cox Proportional Hazards Model

The Cox PH model provided valuable insights into the factors influencing pipeline failure risk. As shown in Table 3, several significant covariates were identified, with a concordance index of 0.73, indicating that the model has a strong ability to correctly rank individuals by their risk, which is considered a good fit for survival analysis. Key findings highlight significant factors, including geographic location and pipeline characteristics. As shown in Figure 6, pipelines in the Midwest region exhibited a hazard ratio of 4.003, suggesting that these pipelines are more than four times as likely to fail compared to pipelines in other regions, likely due to extreme weather conditions. Furthermore, pipelines in hotspot areas (HR = 2.089) exhibited a higher risk of failure, which could be attributed to more frequent external pressures or environmental factors. Pipeline length also demonstrated a significant positive association with failure risk, with an HR of 1.192 (p = 0.012), suggesting that longer pipelines face greater operational or maintenance challenges, making them more susceptible to failure. In addition, certain pipe coatings, such as applied epoxy, were associated with higher hazard ratios, indicating that pipelines with these coatings are at increased risk of failure. However, not all results were statistically strong. The depth of cover factor (p = 0.074) was weakly significant, suggesting the relationship may be due to random variation and requiring further investigation. In contrast, pipelines located in cold spot areas exhibited a significantly lower risk of failure, with a hazard ratio of 0.192 and a coefficient of -1.650, which indicates that these pipelines are substantially less likely to fail. This aligns with the results from the PWP analysis, which also showed lower failure risks in cold spot areas, and the Cox PH model quantified this risk with a low hazard ratio (HR = 0.192). The findings underscore the importance of considering both operational factors and geographic context when developing maintenance strategies for pipelines. While some covariates show weaker significance, the highly significant results offer critical insights for prioritizing risk mitigation, ultimately improving pipeline longevity and reducing failure risks through tailored maintenance strategies.

Table 3: Cox PH Model coefficient results for pipeline failure risk covariates

Covariate	coef	exp(coef)	p
Region_midwest	1.387	4.003	0.029
Pipeline segment status_hot spot	0.737	2.089	0.012
Pipeline_length	0.176	1.192	0.012
Incident_area_type_underground	0.806	2.239	0.213
Depth_of_cover	0.002	1.002	0.074
Pipeline segment status_cold spot	-1.650	0.192	5.36E-05

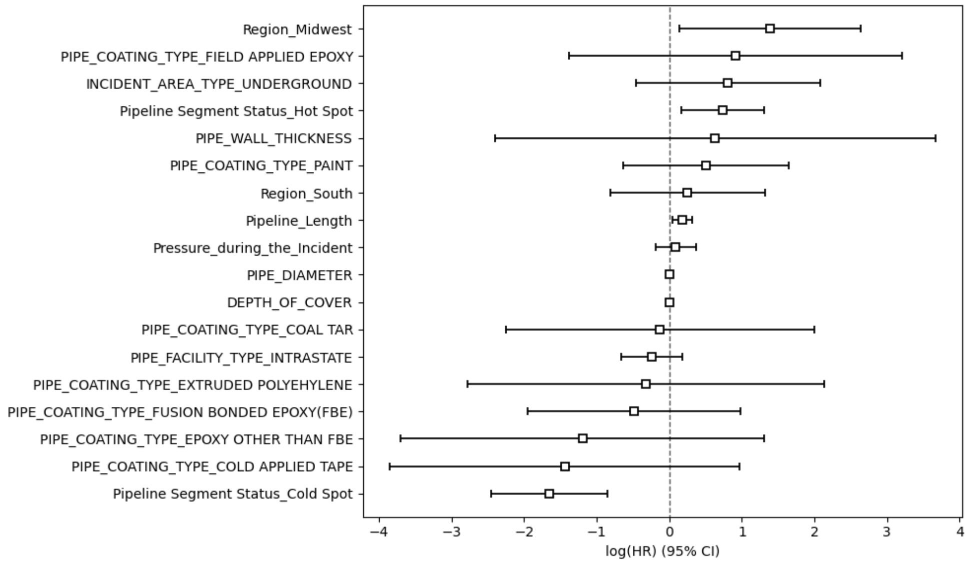

Figure 6: Hazard ratios with 95% confidence intervals for factors influencing pipeline failure risk

3.3 Survival Trends and Insights

The findings of this study emphasize the importance of region-specific strategies to address the unique challenges faced by pipelines in different geographic contexts. Key insights include:

- *Material Selection*: The analysis revealed a significant relationship between coating types and pipeline failure risks. Specifically, pipelines coated with Field Applied Epoxy and Paint showed higher failure risks, with hazard ratios of 2.49 and 1.65, respectively. In contrast, coatings like Fusion Bonded Epoxy and Cold Applied Tape coatings exhibited lower hazard ratios (0.62 and 0.24, respectively), indicating better resistance to corrosion and environmental factors. This underscores the importance of selecting durable coatings, particularly in hotspot area.
- *Region-Specific Strategies:* Hotspot areas should be prioritized in maintenance schedules to reduce the cumulative failure effects, especially those with harsh environmental conditions (like the Midwest, where pipelines exhibit a 4.003 hazard ratio), require more intensive monitoring and maintenance strategies to mitigate accelerated degradation and pipeline failure risks.
- *Targeted Maintenance*: The survival analysis demonstrated that long pipelines (e.g., those over 550 meters) exhibit higher failure risks (HR = 1.192) than short pipelines. This suggests that extended pipeline segments, especially in hotspot areas, face greater operational stresses and environmental exposure. These pipelines should be prioritized for more frequent inspections.
- *Enhanced Regulation*: The significant regional differences in survival rates (e.g., Midwest vs. South, p-value = 1.65E-08) emphasize the need for region-tailored inspection and maintenance policies. For example, in the Midwest, stricter regulations for winterization and frost protection could help prevent failure caused by temperature fluctuations.

These insights provide a foundation for more effective pipeline management, encouraging regionally tailored approaches in material selection, maintenance, and regulation.

4. CONCLUSION

This study integrates spatial and survival analyses to explore regional disparities in pipeline longevity, highlighting significant variations in failure rates potentially driven by geographic, environmental, and operational factors. The survival analysis reveals that pipelines in hotspot areas exhibit accelerated degradation and significantly reduced survival times compared to cold spot areas, highlighting the cumulative effects of recurrent failures. Additionally, this study identified key risk factors that contribute to pipeline degradation, with hazard ratios providing a quantitative understanding of the relative risk in different contexts. The Midwest, for example, demonstrated a notably higher risk of failure (HR= 4.003), likely due to severe environmental conditions. Furthermore, pipelines of shorter lengths were observed to have slightly better survival probabilities. These findings highlight the need for region-specific strategies in pipeline maintenance, including targeted schedules, material selection, and adapted regulatory frameworks. Despite these contributions, the study does not fully explore the distinct degradation patterns and localized risk factors of pipelines. Further research is needed to examine these factors in greater detail, as well as to account for additional variables influencing pipeline failure. Additionally, climate variables, such as temperature, freeze-thaw cycles, and precipitation, were not considered, but future work will integrate them to enhance explanatory power and provide actionable insights. Future research can build on these findings by incorporating more advanced survival analysis techniques, such as the Geographically Weighted Cox Model, to better capture spatially varying risk factors. Additionally, addressing left-censoring in survival data would enhance the accuracy of failure time predictions.

ACKNOWLEDGMENTS

The authors express their gratitude for the funding provided to support this study from the National Science Foundation (NSF) EPSCoR RII Track-2 Program under grant number OIA-2119691. The findings and opinions expressed in this article are those of the authors only and do not necessarily reflect the views of the sponsors.

REFERENCES

Amaya-Gómez, R., Sánchez-Silva, M., Bastidas-Arteaga, E., Schoefs, F., & Muñoz, F. (2019). Reliability assessments of corroded pipelines based on internal pressure – A review. *Engineering Failure Analysis*, *98*, 190–214. https://doi.org/https://doi.org/10.1016/j.engfailanal.2019.01.064

Asaye, L., Ali Moriyani, M., Le, C., Le, T., & Prakash Yadav, O. (2024). Insights from Applying Association Rule Mining to Pipeline Incident Report Data. *Computing in Civil Engineering 2023*, *39*(1), 763–771.

Asaye, L., Le, C., Huang, Y., Le, T. Q., Yadav, O. P., & Le, T. (2025). Predicting and Understanding Emergency Shutdown Durations Level of Pipeline Incidents Using Machine Learning Models and Explainable AI. *Processes*.

Asaye, L., Le, C., Le, T., Yadav, O. P., & Le, T. (2025). Insights into the Interactions of Pipeline Risk Factors and Consequences Using Association Rule Mining. *Journal of Performance of Constructed Facilities*. https://ascelibrary.org/doi/10.1061/9780784485248.092

Bai, M., Du, Y., Chen, Y., Xing, Y., & Zhao, P. (2017). Risk Assessment of Long Gas and Oil Pipeline Projects Inducing Landslide Disasters during Construction. *Journal of Performance of Constructed Facilities*, *31*(5), 1–7. https://doi.org/10.1061/(asce)cf.1943-5509.0000986

Ezuber, H. M., Alshater, A., Hossain, S. M. Z., & El-Basir, A. (2021). Impact of Soil Characteristics and Moisture Content on the Corrosion of Underground Steel Pipelines. *Arabian Journal for Science and Engineering*, *46*(7), 6177–6188. https://doi.org/10.1007/s13369-020-04887-8

Martínez, D. T. S., Kreuz, T., Ridens, B. L., Rahman, R. K., Vasu Sumathi, S., Ross, S., Underwood, J., Iyer, R., & Smith, N. R. (2025). *Chapter 2 - Energy transport is a cornerstone of the energy supply chain* (pp. 7–43). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-443-21893-4.00009-X

Moriyani, M. A., Asaye, L., Le, C., Le, T., & Le, T. (2024). Natural Language Processing for Infrastructure Resilience to Natural Disasters: A Scientometric Review. *Proc. 7th Int. Conf. Geotechnics*.

Naik, D. L., & Kiran, R. (2018). Data Mining and Equi-Accident Zones for US Pipeline Accidents. *Journal of Pipeline Systems Engineering and Practice*. https://doi.org/10.1061/(asce)ps.1949-1204.0000340

U.S. Census Bureau. (2024). *U.S. census regions and divisions. Retrieved February 10, 2025, from*. https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf

Wasim, M., & Djukic, M. B. (2022). External corrosion of oil and gas pipelines: A review of failure mechanisms and predictive preventions. *Journal of Natural Gas Science and Engineering*.

Xiao, R., Zayed, T., Meguid, M. A., & Sushama, L. (2024). Improving failure modeling for gas transmission pipelines: A survival analysis and machine learning integrated approach. *Reliability Engineering and System Safety*, *241*(September 2023), 109672. https://doi.org/10.1016/j.ress.2023.109672

Zha, S., Lan, H., & Huang, H. (2022). Review on lifetime predictions of polyethylene pipes: Limitations and trends. *International Journal of Pressure Vessels and Piping*, *198*, 104663.