

# Multivariate Regression Model to Predict Failure of Pipelines

Narain Hariharan, Mustafa Murat Cobanoglu, Hakan Sahin, Jose Rafael Menendez, and Gökhan Saygili

**Abstract**—The objective of this study is two-fold. The first is to define the significant risk factors pertinent to the corrosion and excavation failures in the natural gas pipeline industry and subsequently check for reliability of the data in hand. The second is to propose a statistical model to predict the odds of survival of the pipeline for the next ten years. The data for the proposed statistical model was obtained from the U.S. Department of Transportation-Pipeline & Hazardous Materials Safety Administration (PHMSA) database of pipeline failure over the past 25 years. A statistical model to predict probability of pipeline failure in the forthcoming decade was developed by processing the data through the JMP software to get the significant parameters contributing to failure. The proposed model predicts that the chance of survival of the pipe, according to the data evaluated, decreases gradually with age for the initial 40 years of the pipeline life and remains within acceptable limits till its nominal expected life. Between years 40 and 60, the survival rate is almost constant and just about at the threshold level, if one assumes a threshold of about 50% survival odds for the forthcoming ten years. The interesting take away from the trend observed is the increase in survival odds beyond the age of 60. The possible reasons contributing to this trend and a detailed analysis of the results based on the proposed model are discussed in the paper.

**Index Terms**—Pipeline failure, corrosion, excavation, multivariate regression model.

## I. INTRODUCTION

Pipelines are a reliable and economic option for the gas industry to transport energy. Oil-gas industry has been seeking a solution picture of the frequencies and probability of incidents. Because of this reason, U.S. Department of Transportation-Pipeline & Hazardous Materials Safety Administration (PHMSA) has started to gather incidents data from all pipeline operators. PHMSA's data provides reliable source that is used to help pipeline operators to establish failure rates and causes of failures in the gas transmission pipeline systems since 1986 [1,] [2]. In the database, there are three different categories namely, natural gas transmission system, natural gas distribution system, and natural gas gathering system.

In this study, the natural gas transmission pipelines data since 1986 has been analyzed. The U.S. Congress signed into law, the Pipeline Safety Improvement Act (PSIA) in 2002. With this law, natural gas transmission operators have had new responsibilities to develop and implement an

Integrity Management Program (IMP), so the PHMSA database has been a valuable and reliable source of information for oil-gas industry since 2002 [3]. The database provides useful information about trends which have developed over the years. The PHMSA database covers only significant incident data that lead to an unintentional gas release. Also, the database has both design parameters (pressure, diameter of pipe, depth of cover) and some incidents parameters (coordinate of incident, incident cause, and total property damage).

The objective of this paper is two-fold. First, is to define the significant risk factors pertinent to the corrosion and excavation failures in the natural gas pipeline industry and subsequently check for reliability of the data in hand. The second objective is to propose a statistical model to predict the odds of survival of the pipeline for the next ten years.

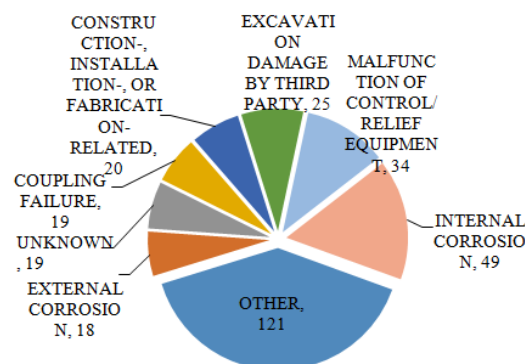


Fig. 1. Detailed representation of failure distribution of Pipelines.

## II. OVERVIEW OF THE PHMSA PIPELINE DATA

PHMSA provides three different time periods for all pipeline incidents which are 1986-2001, 2002-2009, and 2010-present. There are 2625 incidents since 1986 on natural gas transmission pipeline system. Data from all three time periods was used for analysis. A detailed representation of failure distribution of pipelines is presented in Fig. 1. There are 34 different causes reported for failure of pipelines over the 25-year period such as heavy rains/floods, external corrosion, and equipment not installed properly etc. Overall, corrosion (internal and external) is responsible for 24.3% of failures and third party damage/excavation damage is responsible for 25.5% of failure. Therefore, 49.9% of the total failure incidents reported has happened because of these two causes [4]. These are the two causes that have been considered for this study and it is reasonable to assume that the sample set would give a good representation of the entire data set. Also, 92.1% of the pipeline is made by steel, and we have considered just steel pipeline for our analysis.

Manuscript received August 1, 2016; revised December 1, 2016.

N. Hariharan and H.Sahin are with Texas A&M University, College Station, TX 77840, USA. (e-mail: narain110@tamu.edu, shnhakan@tamu.edu).

M. M. Cobanoglu is with BOTAS, Ankara, Turkey (e-mail: mmcobanoglu@gmail.com).

J. R. Menendez is with Fugro Roadware, USA (e-mail: jmenendez@fugro.com).

G. Saygili is with University of Texas at Tyler, USA (e-mail: gsaygili@uttyler.edu).

Fig. 2 shows that as of 2011 the total interstate and intrastate transmission pipeline length was around 305,000 miles. According to the 2013 PHMSA annual report, 30% of all intrastate pipelines were located in Texas, followed by California (11%) and Oklahoma (6%).

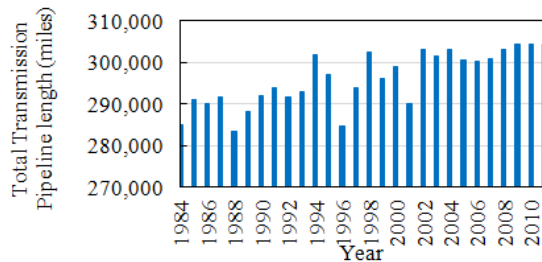


Fig. 2. Distribution of total Pipeline transmission length by year.

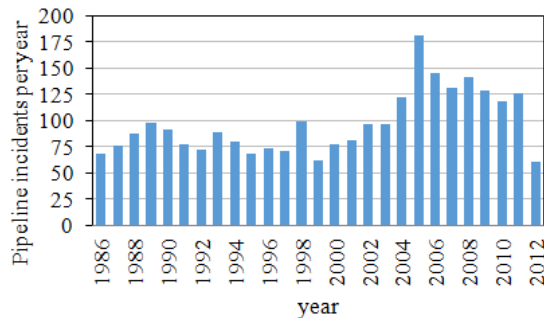


Fig. 3. Distribution of Pipeline incidents by year.

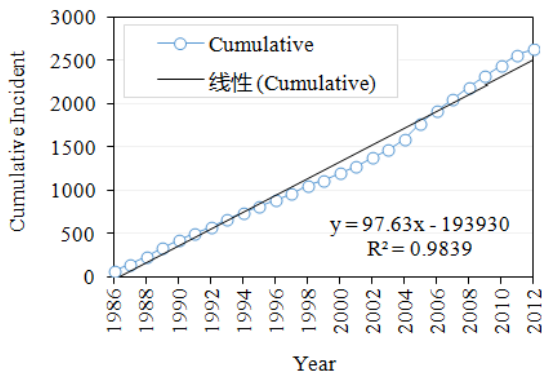


Fig. 4. Cumulative Representation of Pipeline Incidents over the 25-year analysis period.

### III. CONTENTS OF DATABASE

As described in the previous section, we have considered corrosion failure and third party damage. In the original PHMSA database, there are more than 100 columns of information to describe each incident [2], [5]. For statistical analysis, we should not need all the columns and hence many parameters which were thought to be irrelevant to this study were eliminated. We considered the following parameters for preliminary brainstorming the statistical model:

- Type of Material (Steel Pipeline)
- Incident Year
- Installation Year
- Depth of Cover
- Pipe Diameter
- Pipe Wall Thickness
- Incident Pressure

Fig. 3 and Fig. 4 display the yearly and cumulative occurrences of pipeline incidents over the 25-year analysis

period, respectively. Even though significant number incidents have been taking place, pipeline incidents involving death or injury to people dropped by more than half over the past two decades due to the gradually improving safety performance [6].

Out of a total of 2625 incidents recorded between the years 1986-2012 by PHMSA, there is an increasing trend for corrosion failure (Fig. 5) and a decreasing trend for third party damage (Fig. 6). Our opinion is those operators have been aware of cause of third party damage, so they have taken precautions for it. For example, they mark the line direction and they give free meeting for people in the region who live close the line. Also, they offer free calling system for emergency situation. Because of these precautions, there is decreasing trend.

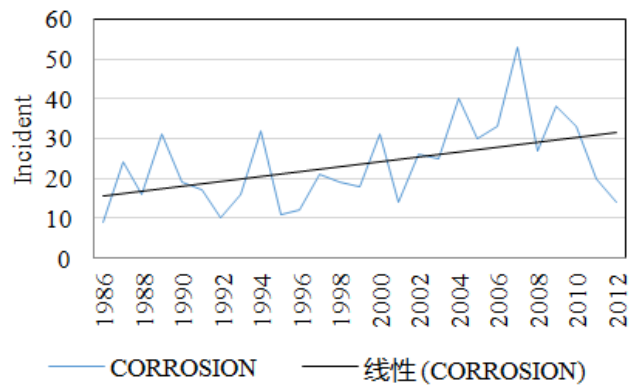


Fig. 5. Year-wise distribution of Pipeline incidents due to corrosion.

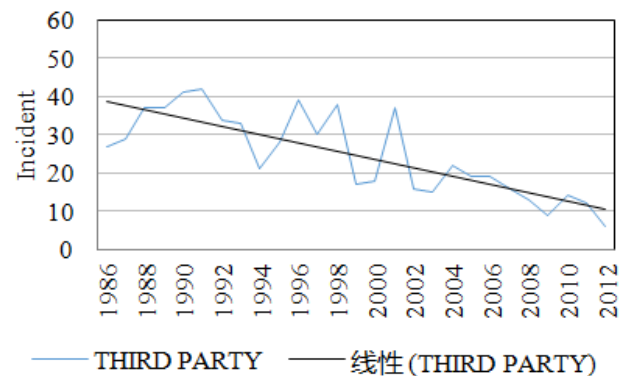


Fig. 6. Year-wise Distribution of Pipeline Incidents due to External Interference.

### IV. SAMPLING PROTOCOL

The extensive data provided by the U.S. Department of Transportation-Pipeline & Hazardous Materials Safety Administration (PHMSA), on 2625 significant pipeline incidents was carefully prioritized based on engineering judgment of the cause-effect relationship between the parameters in consideration and also through some logical assumptions to achieve the desired sample set data for analysis.

The final filtered sample set consisted of a combined 612 pipeline incidents between the years 1986 to 2012 divided in three periods as 1986-2001, 2002-2009 and 2010-2012 [5]. Another critical factor to be noted is the role of missing data in the process of sampling. This factor will be discussed in the next section of this paper.

The first step in filtering the data was based on the reason

for failure of the pipelines. Based on the report published by a European Gas Pipeline Incident Data Group (EGIG) on the study of pipeline failures, the two most common and significant reasons of pipeline failure, namely corrosion (internal and external) and excavation damage were considered for this study [7]. These two causes accounted for nearly 50% of the accidents occurring in the pipeline industry and hence it was deemed reasonable to omit the other caused for the purpose of this study [8].

Next, the data was filtered based on the type of material involved in the incident. Since failure due to corrosion is one of the two major factors in analysis and based on a reasonable assumption that excavation damage would be independent of the type of material involved, only steel pipes were considered as they would be most vulnerable to corrosion. In this process, incidents reported resulting from aluminum and carbon pipes were ignored [9], [10].

The two steps discussed thus far were the only filter implemented based on logical inferences to reduce the original data set. The rest of the elimination process was due to missing data, accounting for close to 20% of the data with certain exceptions which will be discussed.

The other important phase of sampling was to decide the set of independent variables to be considered for the statistical model. This was done on the basis of two reasons. The first was to eliminate the parameters with no useful information about the failure of pipelines and the second was filtering the parameters that could have been useful and relevant but conveyed incomplete and unclear information through the data presented. The first category included variables like area of incident, pipeline facility type, function of the pipeline, seam type, release type, leak type etc. which were deemed inappropriate to this study.

The second basis is quite an unfortunate loss since parameters such as cost and pressure data had to be ignored. In the case of cost data, it is unclear whether the number presented in millions of dollars is a result of cost spent on repairing the pipeline or if it was the cost incurred due to damage of the pipeline with variables such as cost involved in compensation of property and lives of people, thereby making it an unusable but critical variable. The pressure data provided also had similar ambiguities with several unanswered questions regarding whether the pressure reported was the maximum recommended pressure for the pipeline or if it was the pressure at the time of failure, thereby making it unfit to use.

The final independent variables considered for the model were the 'depth of cover' provided for the pipelines, the characteristic wall thickness and diameter of the pipes merged into a single component named the 'critical surface', and the 'age' of the pipelines calculated based on the year of installation.

## V. DEALING WITH MISSING DATA

A common problem that arises with a large data set is missing data. There are many reasons for this. Different pipeline operators have different ways of collecting and reporting data and in the process, while most of the essential and critical parameters would be considered, some variations in reporting methodology is bound to exist. Also,

in some cases a record could be lost or unavailable pertaining to several variables at the incident site. However, it is extremely important to deal with missing data the right way in order to arrive at sensible conclusions.

While elimination is one of the options and the most ideal one, it is impractical to eliminate all the data points with missing components, for it would result in loss of valuable data that can contribute and impact the final conclusions and could also leave very little data to analyze which becomes insignificant and could result in high degree of inaccuracy and uncertainty.

In the pipeline data set, for example, nearly 2200 of the 2600 odd incidents had at least one component missing and eliminating all of the 2200 would leave us with only 400 incidents to further filter and reduce to the meaningful sample desired. Hence, a need arises to develop a proper basis to deal with the missing data to be able to use it in statistical modeling.

A two-step process was adopted to fix the problem of missing data. First, the data points with more than one missing component were completely deleted, as previously mentioned, accounting for nearly 20% of the 2625 incidents reports. Once this was done and the data set was narrowed based on cause of failure and type of material involved, the remaining points were carefully screened for missing data. The period-wise mean of the available data points of the respective independent variables was then substituted in place of the missing data. The data points were checked for reasonable correlation with neighboring data points after implementing this approach and were found to correlate satisfactorily.

## VI. MULTIVARIATE STATISTICAL MODEL

### A. Purpose of the Model

The proposed statistical model is intended to predict the probability of survival of the pipeline given its characteristics, namely age, diameter, wall thickness, and depth of cover. The survival analysis and predicting the odds of failure/survival is based on the number of incidents reported in the 612 data point sample set, during each 10-year increment phase of the pipeline life cycle, i.e. the number of pipeline incidents during the years 0-10, 11-20, 21-30... and so on till the span covering the maximum life reported.

### B. Model Description

The multivariate model has three independent variables, namely age of the pipe in years, critical surface of the pipe in in.2 and the depth of cover, in inches. While age and depth of cover are direct input parameters, the critical surface component is a term that is calculated with the diameter and wall thickness of the pipes. The critical surface is so defined to get a more meaningful term that contributes to the failure of the pipeline, in specific failure due to corrosion. Exposure surface is an alternate term that can be used to better understand the idea behind the critical surface definition. The critical surface would account for the exposed surface of the pipeline to internal and external corrosion [11].

The pipelines, assumed to be circular in cross-section would have a critical surface defined in Eq. (1)

$$\text{Critical Surface} = 2\pi(\text{Diameter} - \text{Wall Thickness}) \quad (1)$$

The dependent parameter in the model, which is the probability of survival of the pipeline over the next ten years, is expressed in percentage and calculated by counting the frequency of incidents per decade increment of pipeline age. This would however result in a case where pipelines within a 10-year age interval would have the same probability to fail based on age and the particular survival odds would depend only on the other characteristics of the pipeline. To account for this, the failure rate parameter is weighted with respect to the maximum pipeline age in the data set.

The set of independent variables and the dependent variable are fed into the statistical analysis software 'JMP' to screen the parameters [12]. JMP runs a variety of combinations of independent parameters and presents the combinations that are expected to have the most significant impact on the dependent variable in question, the failure frequency.

The set of significant combinations are then used to make a multivariate model to predict the survival probability of the pipelines. The general form of the model is in Eq. (2).

$$\text{Probability of Survival (\%)} = 100 - \left[ \left( A + BX_1 + CX_2 + \dots + NX_n \right) \left( \frac{\text{Age}}{\text{Maximum Age}} \right) \right] \quad (2)$$

where:  $X_1, X_2, X_n$  = combinations of the independent variables (age, depth of cover and critical surface) that are significant as reported by JMP; A, B, ...N = corresponding coefficients generated upon running the regression. The maximum age of the pipeline in the sample set is 84 years.

## VII. DATA ANALYSIS METHODOLOGY

### A. JMP Software

JMP software is used to perform simple and complex statistical analyses. It connects statistics with graphics to interactively explore, understand and visualize data. For example, this feature allows tracking any point in a graph and its corresponding data. It provides a comprehensive set of statistical tools and statistical quality control techniques in a single package. Therefore, JMP has been chosen in this particular study to analyze and illustrate the data. Data analysis process is explained in the next section.

### B. Data Processing in JMP

Step-wise Procedure to perform Data Analysis in JMP:

- Go to JMP software and open a New Data Table.
- Place your data into the table as each column represents a set of data and then name each of columns accordingly. In this project columns are named as "Age", "Depth of Cover", "Exposer Surface" and "Incident".
- Next, go to Analyze tab and select Modeling and then select Screening. The screening is shown in Fig. 7.
- A new small window will open, cast select "Y" and "X" columns into roles. The columns of Age, Depth of Cover and Exposer Surface are selected as "X"

parameters, and Incident is selected as single "Y" parameter.

- Click on the OK button to start running the analysis. A screening window is shown in Fig. 8.

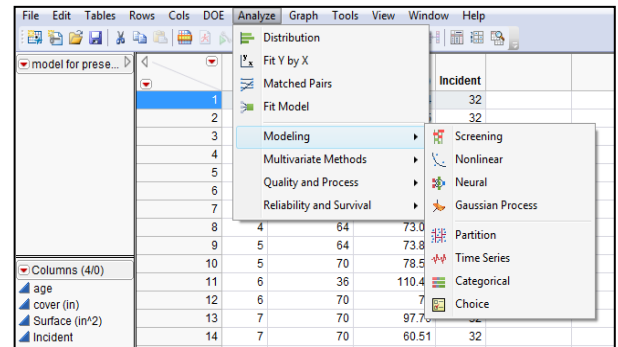


Fig. 7. Path used to perform screening in JMP.

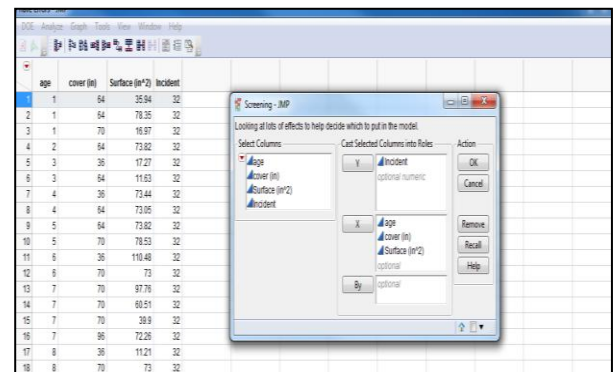


Fig. 8. Snapshot of JMP Screening input.

Term	Contrast	Lenh t-Ratio	Individual p-Value	Simultaneous p-Value
age	-9.9913	-11.09	< 0.001*	< 0.001*
Surface (in^2)	3.4936	4.00	< 0.001*	0.0441*
cover (in)	0.1799	0.21	0.8363	1.0000
age*age	-24.3095	-27.81	< 0.001*	< 0.001*
age*Surface (in^2)	0.3432	0.39	0.6951	1.0000
Surface (in^2)*Surface (in^2)	0.5372	0.61	0.5383	1.0000
age*cover (in)	-0.7722	-0.89	0.3739	1.0000
Surface (in^2)*cover (in)	-1.4364	-1.71	0.0874	1.0000
cover (in)*cover (in)	-0.8715	-1.00	0.3171	1.0000
age*age*age	13.3211	15.24	< 0.001*	< 0.001*
age*age*Surface (in^2)	-1.3043	-1.49	0.1361	1.0000
age*Surface (in^2)*Surface (in^2)	0.1422	0.16	0.8896	1.0000
Surface (in^2)*Surface (in^2)*Surface (in^2)	0.2972	0.34	0.7313	1.0000
age*age*cover (in)	-1.6200	-2.24	0.0263	1.0000
age*Surface (in^2)*cover (in)	1.2813	1.47	0.1441	1.0000
Surface (in^2)*Surface (in^2)*cover (in)	-0.2383	-0.27	0.7853	1.0000
age*cover (in)*cover (in)	-0.8028	-0.92	0.3571	1.0000
Surface (in^2)*cover (in)*cover (in)	1.2907	1.48	0.1410	1.0000
cover (in)*cover (in)*cover (in)	1.0579	1.21	0.2253	1.0000

Fig. 9. Snapshot of JMP Screening output.

The purpose of the screening is that analysis will search to determine significant weight of each parameter. Also, the analysis matches the parameters with different combination and significant combinations will be determined and screened. In the result window, parameters are shown as histograms in which the darker pink colored parameters are more significant than lighter colors. This analysis helps to eliminate some of the insignificant parameters.

Initially, the analysis process starts with a multi-variable parameter as mentioned earlier, and based on the analysis results, the lesser significant parameters have been eliminated and the subsequent analysis has been performed with higher significant parameters. Thus this process will be iterated with the significant parameters again thereby more than repeated analyze may be required for more accuracy. A set of screening results which are color coded to mark significant variables are shown in Fig. 9.

Once the screening analysis gives a significant correlation,



those parameters are utilized for modeling. The next phase is to determine a best representative mathematical model and correlation based on the selected parameters. Modeling tool is utilized for this purpose which is under analysis tab and select a “Fit Model”. Once it is selected a new window will open. Fit model path is shown in Fig. 10.

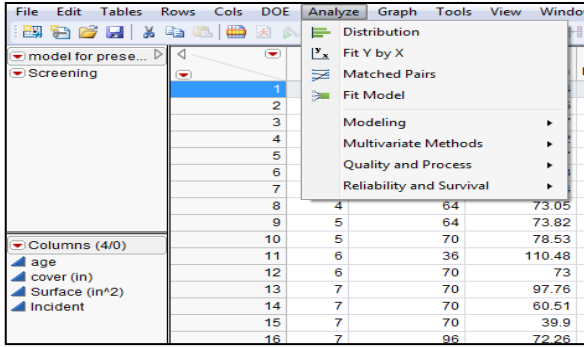


Fig. 10. Snapshot of JMP 'Fit model'.

Pick the role of the selected parameters as “Y” and select the construct model effects as “X” parameters. Modeling type can be selected from drop down menu and then model will start by clicking on Run DOE button. Fit model window is shown in Fig. 11.

The analysis results will be shown in a new window as plotted graph along with a summary of fitting data including  $R^2$ , mean, error, and coefficient of parameters.

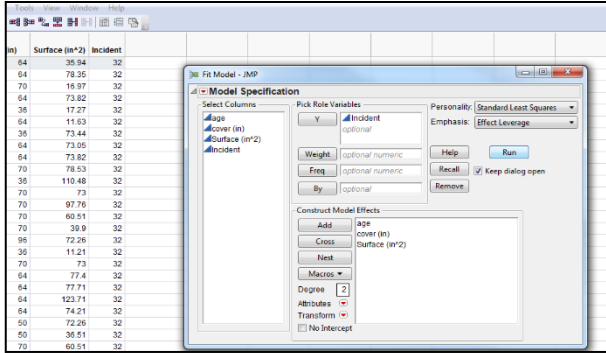


Fig. 11. Fit model window along with added X and Y parameters.

### VIII. MODEL DERIVATION

A relation for the Failure index and subsequently the Probability of Survival was established from the JMP results, giving birth to our statistical model. The model includes independently, age pipeline cover and critical surface and combinations of exposure surface and depth of cover, all of which are parameters used in the model because they have significant contribution effect on the failure prediction [1], [2].

Mathematical form of the derived model is given in Eq. 3 and Eq. 4;

$$I_F = \begin{bmatrix} (102.16 - 0.002A + 0.031E_s - 0.065(A - 34)^2) \\ -(0.005(E_s - 81.6)(D_c - 53.6)) \\ -(0.002(D_c - 53.6)^2) \end{bmatrix} \quad (3)$$

$$\text{Probability of Survival} = (100 - I_F) \left( \frac{A}{M_{age}} \right) \quad (4)$$

where  $I_F$ = Failure Index;  $A$ = Age of pipe line as year;  $E_s$ = Exposure surface of pipe include inside and outside surface (in);  $D_c$ = Depth of cover (in.); and  $M_{age}$ = Maximum survival age. A summary of the statistical analysis results of the model are tabulated in Table I.

TABLE I: SUMMARY OF FIT MODEL

R-Square	0.7385
R-Square Adj.	0.7346
Root Mean Square Error	17.77
Mean of Response	86.225
Observations	612

### IX. CORRELATION PLOTS

Fig. 12 shows the effect of ageing on the probability of survival of the pipelines. This representation of the survival analysis with the model follows an inverse ‘bathtub’ relationship which has been validated by a shape factor  $\beta$  of greater than 1.0 upon applying the Weibull distribution to the data set in consideration supporting the theory that the new product performs better than the used version [13], [14]. This result is debatable since there is expected to be some wear-in at the time of installation of the pipelines due to the conditions, compatibility issues and various other factors. However, the results in the present study point to an idealistic case where the pipeline has been designed taking into consideration all uncertainties and variability making the newly designed pipeline least susceptible to failure. The other scientific reason supporting the trend is the fact that corrosion/ chloride ingress takes significant time to penetrate the ground and affect the pipe, thereby resulting in gradual decline in survival rates as observed [15], [16].

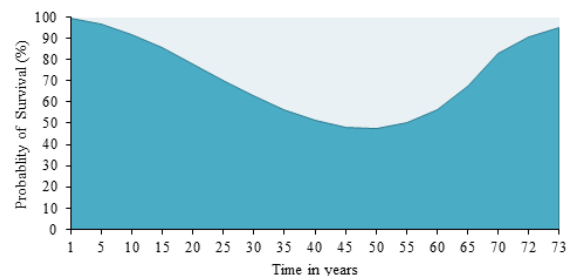


Fig. 12. Correlation graph representing the relationship between age and probability of survival.

TABLE II: PARAMETRIC ANALYSIS PERFORMED ON THE MODEL

Variable	Value	% Probability of Survival for next 10 years
Age (years)	5	97.1
	20	78.3
	35	56.4
	65	68.3
Depth of Cover (in.)	20	74.1
	35	78.3
	60	81.1
Critical surface (in.)	50	80.3
	80	78.3
	100	75.7

### X. DISCUSSION OF THE RESULTS:PARAMETRIC ANALYSIS

The results of the parametric study on the model are

presented in Table II shown below. It is to be noted that the baseline values for the variables are set at Age of 20 years, Depth of cover of 35 inches and a Critical surface of 80 inches. All the variation in each of the variables is performed with the other independent variables set at the baseline values.

## XI. CONCLUSIONS

Corrosion has been identified in previous studies as the most common cause of pipeline incidents and often leads to openings (pinholes and cracks) resulting in leaks and in more severe cases, rupture of the pipeline. The monetary loss incurred as a result of the damage is significant.

The high contribution of external interference emphasizes its importance to the pipeline operators and authorities. External interference operations are characterized by potentially severe consequences.

It is obvious from the trend that depth of cover is an important consideration in the design of pipelines and the odds of survival of the pipeline significantly increases with more depth of cover. This can be attributed to the fact that with a greater depth of cover, the chance of external interference is expected to be lesser as a result of difficulty to access.

The physical characteristics of the pipeline, namely the diameter and the wall thickness of the pipelines are partially dependent on each other and have a role to play in predicting failure. In general, it is expected that smaller diameter pipelines are more likely to be affected by third party damage since they can be easily hooked up during ground works compared to larger and bigger pipelines. Also the fact that smaller diameter pipes are associated with smaller wall thickness makes them less resistant to failure.

As a corollary, wall thickness can be said to be an extremely effective protective measure against external interference. The failure point of a thinner pipeline is reached more quickly. Corrosion in thicker pipelines takes longer before resulting in an incident and therefore there is more time to detect and resolve the issue.

Hence, the exposure or the critical surface, which is a function of the diameter and wall thickness of the pipeline, obviously plays a role in the failure of the pipe. It can be observed from the parametric study that with greater exposure, the pipeline is more prone to failure. This trend is along expected lines since greater the length of pipeline exposed to environmental changes (moisture, temperature), greater is the susceptibility to corrode.

On the whole, it can be concluded that major failures like leaks and ruptures are more likely to occur in pipelines with smaller diameters, a small depth of cover and less wall thickness.

It can be observed from the proposed model that the chance of survival of the pipe, according to the data evaluated, decreases gradually with age for the initial 40 years of the pipeline life and remains within acceptable limits till its nominal expected life. Between years 40 and 60, the survival rate is almost constant and just about at the threshold level, if one assumes a threshold of about 50% survival odds for the forthcoming ten years. The interesting take away from the trend observed is the increase in survival

odds beyond the age of 60. This leads us to believe that the pipeline, once past its expected life of about 50 years, the operators inspect the pipe and perform the required maintenance on the pipeline or dig out and replace the pipeline depending on the condition of the pipe and also the availability of funds.

The age at which the survival rate reaches a constant and no longer decreases with time can be said to be the nominal expected age of the pipeline. This age is about 40 years for the data set analyzed. It is hence recommended that beyond the age of 40, inspection/monitoring of the pipeline is required and maintenance or replacement of the pipe is needed depending on the desired threshold set.

The prediction of the survival analysis with the model follows an inverse 'bathtub' relationship which has been validated by a shape factor  $\beta$  of greater than 1.0 upon applying the Weibull distribution to the data set in consideration supporting the theory that the new product performs better than the used version. This result is debatable since there is expected to be some wear-in at the time of installation of the pipelines due to the conditions, compatibility issues and various other factors. However, the results in the present study point to an idealistic case where the pipeline has been designed taking into consideration all uncertainties and variability making the newly designed pipeline least susceptible to failure. The other scientific reason supporting the trend is the fact that corrosion/chloride ingress takes significant time to penetrate the ground and affect the pipe, thereby resulting in gradual decline in survival rates as observed.

## XII. RECOMMENDATIONS

For a more comprehensive understanding of failure of pipelines due to corrosion, data regarding the temperature, humidity and other site conditions and also the type of coating, if any involved need to be considered. Incorporating these factors into the model is expected to improve predictions.

Several protective measures have been taken by pipeline operators to try and prevent corrosion. These measures include cathode protections and the use of protective coatings. Inline inspections and pipeline surveys also help detecting corrosion at an early age.

Improvements in the prevention of external interference incidents calls for more stringent enforcements of land use planning, the application of one-call systems for the digging activities of external parties by the adoption of appropriate actions by the gas companies like supervision or marking the pipeline in the direct neighborhood of the pipeline digging activity.

The proposed model does not have the capacity to predict the time when the necessity of maintenance needs to be initiated and it is left for the user to decide this based on the desired minimum probability of survival for the next ten years.

More information regarding the nature and extent of damage due to corrosion and other external intrusions and also about the site conditions is needed for a better prediction of failure.

The model can be improved and have a much more

meaningful representation with incorporating the above said factors. Another critical variable that has been ignored due to lack of availability of sufficient information is the cost/budget required to maintain or replace the pipes. This factor is expected to have a significant bearing on the decision making process

## REFERENCES

- [1] PHMSA, "About PHMSA," 2011
  - [2] U. S. DOT. (2011). Fct sheet: High consequence areas (HCA). [Online]. Available: <http://primis.phmsa.dot.gov/comm/FactSheets/FSHCA.htm>
  - [3] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, Fifth Edition, John Wiley & Sons, Hoboken, NJ, 2011.
  - [4] G. A. Papadakis, "Major hazard Pipelines: A comparative study of onshore transmission accidents," *Journal of Loss Prevention in the Process Industries*, vol. 12, no. 1, pp. 91-107, 1999.
  - [5] PHMSA. (2011). Pipeline incident data access. <http://primis.phmsa.dot.gov/comm/reports/safety/SIDA.html?nocache=1217>
  - [6] U.S. Department of Transportation. The state of national pipeline infrastructure. [Online]. Available: [https://opsweb.phmsa.dot.gov/pipelineforum/docs/Secretarys%20Infrastructure%20Report\\_Revised%20per%20PHC\\_103111.pdf](https://opsweb.phmsa.dot.gov/pipelineforum/docs/Secretarys%20Infrastructure%20Report_Revised%20per%20PHC_103111.pdf)
  - [7] EGIG. (2015). Gas Pipeline incidents: 9th report of the European gas pipeline incident data group (period 1970 – 2013)," *European Gas Pipeline Incident Data Group*.
  - [8] W. K. Muhlbauer, *Pipeline Risk Management Manual: Ideas, Techniques, and Resources*, Gulf Professional Publishing and Elsevier, Burlington, MA, 2004.
  - [9] F. Caleyó, L. Alfonso, J. Alcántara, and J. M. Hallen, "On the estimation of failure rates of multiple Pipeline systems," *Journal of Pressure Vessel Technology*, vol. 130, no. 2, pp. 25-29, 2008.
  - [10] J. Røstum, "Statistical modelling of pipe failures in water networks," Doctoral Dissertation, Norwegian University of Science and Technology, Trondheim, Norway, 2000.
  - [11] J. A. Nachlas, *Reliability Engineering: Probabilistic Models and Maintenance Methods*, CRC Press, Taylor & Francis Group, Florida, 2005.
  - [12] SAS, "JMP statistical software," 2012.
  - [13] D. V. D. Brand and R. Kenter, "Gas pipeline incidents," *European Gas Pipeline Incident Data Group*, Netherlands, 2011.
  - [14] K. Reinschmidt, "Project risk management, CVEN 644, lecture notes fall 2012," *Texas A&M University*, 2012.
  - [15] M. Baker Jr, "Mechanical damage final report," Department of Transportation Pipeline and Hazardous Materials Safety Administration Office of Pipeline Safety, Washington, D.C, 2009.
  - [16] H. Ascher and H. Feingold, "Repairable systems reliability: Modeling, inference, misconceptions and their causes," New York, NY, 1984.
- Narain Hariharan** is currently pursuing a Ph.D. degree in civil engineering at Texas A&M University, USA. He received his bachelor's degree in civil engineering from National Institute of Technology, Tiruchirappalli, India in 2011 and his master's degree in civil engineering from Texas A&M University in 2013. Mr. Hariharan also has three years of work experience in USA during which he has worked in multiple projects in the construction and consulting industries.
- Mustafa Murat Cobanoglu** has a M.Sc. degree in Civil/Construction Engineering and Management at Texas A&M University and has over 5 years of professional experience in project control, quality control, schedule management, cost management, preparing and reviewing contract documents for major public works projects including natural gas pipelines, highway construction, highway bridges, airport terminal facilities. Murat currently works for BOTAS in Ankara, Turkey as a pipeline and civil engineer.
- Hakan Sahin** received his bachelor of science in civil engineering degree from Nigde University in 2006. He received a master-of-science degree and a doctor of philosophy in civil engineering from Texas A&M University, USA, respectively, in 2011 and 2014. Since 2014, he has been working as a project manager in geotechnical engineering department at Intertek-Professional Service Industries, Inc. (PSI), Texas, USA.
- Jose Rafael Menendez** received his Ph.D degree in civil engineering from Texas A&M University, in 2014 and has since worked as a project manager in Fugro Roadware, Inc. Mr. Menendez has over twenty years of industry work experience in several parts of South America where he has managed multiple consulting projects in geotechnical and pavement engineering. Mr. Menendez is also a registered Professional Engineer in Texas.
- Gokhan Saygili** obtained a bachelor's degree from Istanbul Technical University in 2002, a master's degree from Concordia University, Montreal, Canada in 2005, and a Ph.D. degree from the University of Texas at Austin in 2008. After working at Norwegian Geotechnical Institute for four years, he joined the faculty in the Civil Engineering Department at the University of Texas at Tyler in the Fall of 2013. He is a registered Professional Engineer in Texas.