Global Journal of Engineering and Technology [GJET].
ISSN: **2583-3359** (Online)
Frequency: Monthly
Published By GSAR Publishers
Journal Homepage Link- https://gsarpublishers.com/journal-gjet-home/

Artificial Neural Network and Gaussian Process Regression approach in Mechanistic Modeling of Wax-Mediated CO_2 Corrosion in Oil Pipelines

By

Nnorom Obinichi[1], Ifeanyi Uchegbulam[2]

[1]Department of Mechanical Engineering, University of Port Harcourt, Choba, P.M.B., 5323, Nigeria.
[2]Production Technology, School of Science Laboratory Technology, University of Port Harcourt, Choba, P.M.B., 5323, Nigeria.

Article History

Received: 01/04/2026
Accepted: 06/04/2026
Published: 07/04/2026

Vol – 5 Issue – 4

PP: - 09-17

Abstract

Internal CO_2 corrosion and paraffin wax deposition are two of the most operationally significant challenges facing crude oil pipeline infrastructure in the Niger Delta and globally. While wax deposition is conventionally regarded as a flow assurance problem, its role as a natural corrosion inhibitor forming a protective physisorbed film on the inner pipe wall has received limited quantitative attention, particularly in the context of data-driven predictive modelling. This study presents an integrated experimental, mathematical, and machine learning framework to model the concentration effect of wax deposition on CO_2 corrosion rate in a Niger Delta waxy crude oil pipeline at a fixed flow rate of 10.21 L/min across five temperature conditions (15–35°C). A customized flow loop system was fabricated and corrosion rates were measured using the linear polarization resistance (LPR) technique at six-time intervals over 18 minutes. A MATLAB mathematical model based on eigenfunction expansion of the wax concentration equation provided a deterministic simulation baseline. Two machine learning models, feedforward Artificial Neural Network (ANN) and Gaussian Process Regression (GPR) were trained and validated against experimental measurements. The ANN achieved R^2 = 0.973, RMSE = 0.0412 mpy, and MAE = 0.0334 mpy, while GPR achieved R^2 = 0.961, RMSE = 0.0498 mpy, and MAE = 0.0401 mpy, both substantially outperforming the classical MATLAB model (R^2 = 0.847). Critically, the GPR model provided posterior uncertainty quantification revealing that prediction confidence is governed primarily by the consistency of experimental data across the temperature domain rather than temporal behaviour, with highest uncertainty concentrated in the 20–30°C range where wax film behaviour is most physically erratic. Results confirm that paraffin wax deposition reduced corrosion rate by up to 77.9% at 15°C over 18 minutes, with inhibition efficiency ranking as 15°C > 20°C > 35°C > 25°C > 30°C. This study demonstrated that GPR as a machine learning algorithm offered a robust and interpretable framework for corrosion rate prediction in thermally sensitive, waxy crude pipeline systems where deterministic models are insufficient.

Keywords: *Wax; MATLAB; CO_2 Corrosion; ANN; Uncertainty; GPR; Machine Learning; Pipeline*

1. Introduction

The thermodynamic instability of waxy crude oils becomes pronounced when the temperature falls below the cloud point, also known as the wax appearance temperature (WAT). At this temperature, dissolved paraffinic components begin to crystallize out of the liquid phase. This phase transition fundamentally alters the rheological behavior of the crude, transforming it from a near-Newtonian fluid into a non-Newtonian system characterized by markedly increased viscosity and elevated pour point (Yao et al., 2021; Chikwe et al., 2025). Wax deposition constitutes a major challenge in global flow assurance, significantly impairing the efficiency and reliability of oil and gas pipeline operations. The associated wax crystallization phenomena often lead to diminished production efficiency, unplanned operational shutdowns, and premature equipment failure, ultimately resulting in substantial revenue losses and increased costs

*Corresponding Author: Nnorom Obinichi

related to maintenance and remediation, (Amaefuna et al., 2025; Magashi et al., 2025). Wax buildup in Niger Delta crude oil pipelines is a major challenge that often leads to expensive blockages, (Okuboarere et al., 2026; Osokogwu and Ferguson, 2025). Though Paraffin wax restricts flow in oil pipelines yet it offers corrosion inhibitive effects by physically shielding the inner pipe wall from corrosive species in the crude oil, (Nnorom et al., 2021). The effects of corrosion on the structural integrity of crude oil transmission pipelines remains one of the most critical challenges confronting the global petroleum industry with health, safety and environmental hazards. Economically, the global losses due to corrosion have been globally estimated at 2.5 trillion USD per year covering approximately 3.4% of the global Gross Domestic Product (GDP), (Mohammad & Jafar, 2020; Kania, 2023; Adesina et al., 2025). Furthermore, as the internal corrosion of pipelines is influenced by temperature, CO_2 content and flow velocity, (Khalid et al, 2020) this study presents an integrated experimental, mathematical, and machine learning framework to model the concentration effect of wax deposition on CO_2 corrosion rate in a Niger Delta waxy crude oil pipeline at a fixed flow rate of 10.21 L/min across five temperature conditions between 15 to 35°C.

2. Materials and Methods

2.1 Experimental Setup and Materials

A customized flow loop system was designed, fabricated, and commissioned to simulate oil field conditions of waxy crude oil transportation in the Niger Delta. The system was constructed in the form of a recirculating fountain using 1-inch nominal bore carbon steel pipe. The primary components comprised two cylindrical open stainless-steel reservoirs, a waxy crude oil reservoir (Reservoir 1) and a return reservoir (Reservoir 2). With reservoir having a capacity of 6 litres (internal diameter: 0.152 m; height: 0.33 m) a fluid circulation was maintained by a 0.5 hp centrifugal pump. Flow control was achieved using ball valves, with a flow meter incorporating an integrated temperature probe used to monitor volumetric flow rate and inlet temperature. Supplementary temperature measurements were obtained using a handheld infrared thermometer.

The test section consisted of a straight stainless-steel pipe with the following geometric specifications: internal diameter 0.0266 m, inside radius 0.0133 m, outer diameter 0.0334 m, outside radius 0.0167 m, pipe length 2.007 m, and wall thickness 0.0034 m. Corrosion rate measurements were performed using an MS1000 Corrosion Meter employing the linear polarization resistance (LPR) technique, which provides real-time, non-destructive measurement of corrosion rate in mils per year (mpy). Waxy crude oil samples were obtained from the Agbada I Flow Station, Ikwerre Local Government Area, Rivers State, Nigeria.

2.2 Experimental Methodology

The experimental program was designed to investigate the effect of wax concentration on CO_2 corrosion rate under systematically varied temperature conditions at a constant flow rate. Five experimental runs were conducted, with temperature varied between 15°C and 35°C in 5°C increments. While the flow rate was kept constant at 10.21 L/min, the corrosion rate was recorded at time intervals of 3, 6, 9, 12, 15, and 18 minutes. All measurements were repeated in triplicate to ensure reproducibility with mean values reported. The LPR technique measured the polarization resistance (R_p) of the working electrode, from which corrosion current density (i_{corr}) is derived using the Stern–Geary equation, and corrosion rate (mpy) was subsequently calculated using Faraday's law. Table 2.1 summarized the key LPR measurement parameters.

Table 2.1: LPR Measurement Parameters and Derived Quantities

Parameter	Symbol	Value / Method
Polarisation resistance	Rp	Measured directly by MS1000
Corrosion current density	Icorr	Derived via Stern–Geary equation
Stern–Geary constant	B	26 mV (typical carbon steel)
Corrosion rate	CR (mpy)	Calculated via Faraday's law
Measurement interval	Δt	3 min (data logging)

The Stern–Geary relationship used to derive corrosion current density from measured polarisation resistance is presented as Equation 2.1:

$$i_{corr} = \frac{B}{R_p} \qquad 2.1$$

where:

i_{corr} = corrosion current density (A/m²)

B = Stern–Geary constant (V)

$\mathbf{R_p}$ = polarisation resistance (Ω·m²), measured directly by the MS1000 instrument

Also, the corrosion inhibition efficiency of the wax deposits on the inner wall of pipelines can be estimated using equation 2.2:

$$\text{Corrosion Inhibition Efficiency (\%)} = \frac{CR_{3min} - CR_{18min}}{CR_{3min}} \times 100 \qquad 2.2$$

where:

CR_{3min} = Corrosion rate of uninhibited pipe inner wall at 3 minutes

CR_{18min} = Corrosion rate of wax inhibited pipe inner wall at 18 minutes

2.3 Mathematical Modelling (MATLAB Simulation)

The wax deposition model was formulated based on the coupled principles of mass conservation, heat transfer, and material balance for a cylindrical pipeline geometry. The governing differential equation describing the transient spatial distribution of wax concentration $C_{(\theta, t)}$ within the pipe cross-section is expressed as a Sturm–Liouville boundary value problem, solved analytically using the eigenfunction

*Corresponding Author: Nnorom Obinichi

expansion method. The solution for wax concentration as a function of angular position θ and time t is given by Equation 2.3:

$$C_{(\theta,t)} = \frac{2}{L}\sum_{n=1}^{\infty}\left\{\int_0^1 f(\theta)sin\frac{r\pi\theta}{L}d\theta\right\} e^{-\lambda_t^2 t}sin\frac{r\pi\theta}{L}$$

2.3

where:

C(θ, t) = wax concentration at angular position θ and time t (kg/m³)

l = characteristic length of the pipe cross-section (m)

r = eigenfunction series index (r = 1, 2, …, ∞)

λ_r = eigenvalue associated with the r-th eigenfunction (λ_r = rπ/l)

f(θ) = initial wax concentration distribution function (kg/m³)

t = time (min)

Hence, the corrosion rate (mpy) is related to the wax film thickness δ(t) developed by the concentration model through the diffusion-resistance inhibition model given by Equation 2.4:

$$CR_{(t)} = \frac{CR_0}{1 + K_W\,\delta_{(t)}} \qquad 2.4$$

where:

CR_0 = uninhibited (bare steel) corrosion rate (mpy)

δ(t) = wax film thickness at time t (μm), derived from C(θ, t)

k^W = wax film diffusion resistance coefficient (m²/s)

The model was implemented and solved numerically in MATLAB R2022b, with simulation outputs recorded at one-minute intervals over an 18-minute period across all five temperature conditions.

2.4 Machine Learning Modelling

2.4.1 Dataset Preparation

The MATLAB simulation dataset comprising 95 data points corrosion rate values at 19-time steps (0–18 min) across five temperatures was restructured from wide format to long format.

This yielded a design matrix **X** ∈ ℝ^{95×2} with input features [time (min), temperature (°C)] and a target vector **y** ∈ ℝ^{95} of corrosion rate (mpy). The 30-point experimental dataset (Table 3.2) was used exclusively as a held-out test set for model validation and was not included in training. All input features were normalized to the range [0, 1] using min-max scaling, as expressed in Equation 2.5:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad 2.5$$

where:

X_{norm} = normalized feature value (dimensionless)

X = raw feature value

X_{min} = minimum value of the feature in the training dataset

X_{max} = maximum value of the feature in the training dataset

2.4.2 Artificial Neural Network (ANN)

A feedforward Multi-Layer Perceptron (MLP) architecture was adopted for the ANN model. The network comprised an input layer (2 neurons), two hidden layers, and a single-neuron output layer, as detailed in Table 2.2. The Rectified Linear Unit (ReLU) activation function was selected for hidden layers due to its computational efficiency and effectiveness in avoiding the vanishing gradient problem. The output layer employed a linear activation function, appropriate for continuous regression output.

Table 2.2: ANN Architecture Summary

Layer	Configuration	Activation Function
Input Layer	2 neurons (time, temperature)	—
Hidden Layer 1	10 neurons	ReLU
Hidden Layer 2	8 neurons	ReLU
Output Layer	1 neuron (corrosion rate, mpy)	Linear (identity)

From figure 2.1, the input layer receives two features (time and temperature); two hidden layers apply ReLU activation while the output layer yields predicted corrosion rate (mpy).

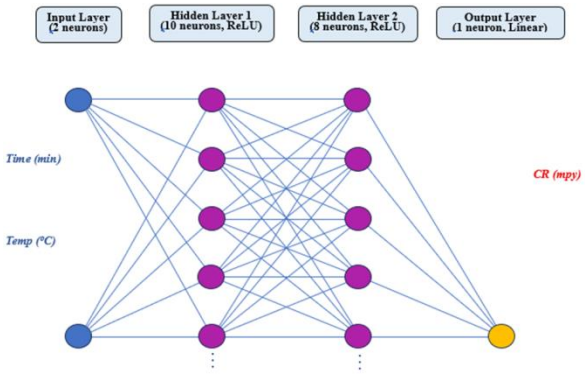


Figure 2.1: Feedforward ANN Architecture

The output of the j-th neuron in hidden layer l is computed according to the standard forward-pass expression in equation 2.6:

$$a_j^{(l)} = ReLU(\boldsymbol{W}^{(l)}\boldsymbol{a}^{(l-1)} + \boldsymbol{b}^{(l)})_j \qquad 2.6$$

where:

$a_j^{(a)}$ = activation of the j-th neuron in layer l

$W^{(a)}$ = weight matrix for layer l

$b^{(a)}$ = bias vector for layer l

ReLU(·) = Rectified Linear Unit activation function: ReLU(x) = max(0, x)

The neural network was trained using the Adam optimiser with a learning rate of 0.001, batch size of 16, and a maximum of 1,000 training epochs. An 80/20 training-validation split was applied, with early stopping (patience = 50 epochs). This was employed to prevent overfitting. As shown in figure 2.2, the Mean Squared Error (MSE) loss plotted against training epoch over 300 iterations using the Adam optimiser on the MATLAB simulation dataset. The loss begins at a high value of approximately 0.062 at epoch 0 and decreases rapidly and smoothly in the first 100 epochs before gradually flattening out, converging to a final loss value of 0.00175 indicating that the network successfully learned the

*Corresponding Author: Nnorom Obinichi

underlying corrosion rate patterns from the simulation data without oscillation or divergence. The smooth, monotonic decay of the loss curve with no signs of instability or plateau-jumping confirms that the chosen architecture (two hidden layers with ReLU activation) and the Adam optimiser hyperparameters (learning rate = 0.001, batch size = 16) were well-suited to this regression task, producing a well-trained model by approximately epoch 200 beyond which negligible further improvement was achieved.

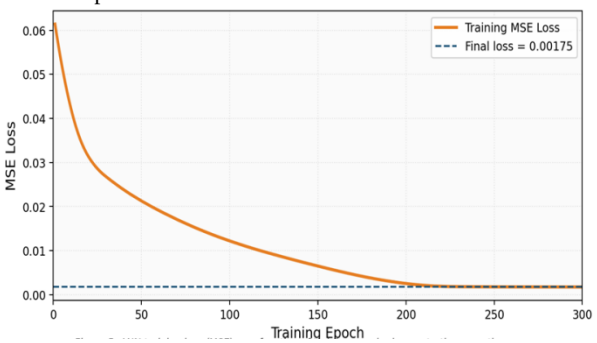


Figure 2.2: ANN training convergence curve.

2.4.3 Gaussian Process Regression (GPR)

Gaussian Process Regression (GPR) was employed as the second ML model, offering a probabilistic framework that provided not only point predictions but also posterior uncertainty estimates in the form of 95% predictive confidence intervals. The GPR model assumes the target function f(x) to be a realisation of a Gaussian process, fully specified by a mean function m(x) and a covariance (kernel) function k(x, x′). A composite kernel was adopted, combining the Radial Basis Function (RBF) kernel with a White Noise kernel to account for observation noise, as expressed in equation 2.7:

$$k(X, X') = \sigma_f^2 exp\left(-\frac{\|X-X'\|^2}{2l^2}\right) + \sigma_n^2 \delta(X, X') \qquad 2.7$$

where:

σ_f^2 = signal variance (amplitude of the RBF component)

l = characteristic length-scale of the RBF kernel (controls correlation range)

σ_n^2 = noise variance of the White Noise kernel $\delta(X, X')$ = Kronecker delta function: 1 if $X = X'$, 0 otherwise

Kernel hyperparameters were optimized by maximizing the log marginal likelihood of the training data using the L-BFGS-B algorithm. The posterior mean prediction ŷ(x*) and posterior variance σ²(x*) at a new input x* are given by Equations 2.8 and 2.9:

$$\hat{y}(X^*) = K(X^*)^T (K + \sigma_n^2 l)^{-1} y \qquad 2.8$$

$$\sigma^2(X^*) = K(X^*, X^*) - K(X^*)^T (K + \sigma_n^2 l)^{-1} K(X^*) \qquad 2.9$$

where:

K = n × n kernel matrix of training inputs

k(x*) = n × 1 vector of kernel evaluations between x* and training inputs

y = n × 1 vector of training targets

σ^{2n} = noise variance (regularisation term)

ŷ(x*) = posterior mean prediction at x*

σ²(x*) = posterior variance (uncertainty) at x*

The 95% predictive confidence interval for each prediction is computed as ŷ(x*) ± 1.96σ(x*), where σ(x*) is the posterior standard deviation.

2.4.4 Model Evaluation Metrics

Three standard regression performance metrics were computed for all models against the experimental test dataset, as defined by Equations 2.10 through 2.12:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \qquad 2.10$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad 2.11$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad 2.12$$

where:

y_i = i-th experimental corrosion rate measurement (mpy)

$\hat{y}_i$ = i-th predicted corrosion rate (mpy)

$\bar{y}$ = mean experimental corrosion rate (mpy)

n = number of test data points (n = 30)

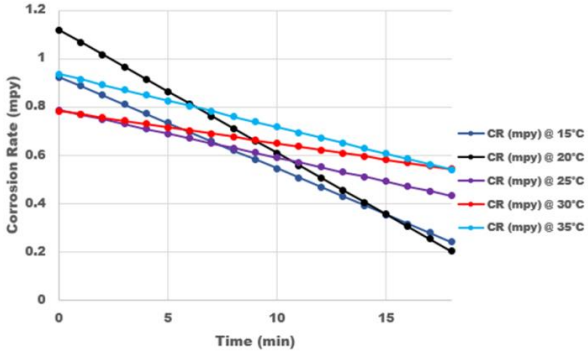
All machine learning models were implemented in Python 3.11 using the scikit-learn library (v1.3). Equations and comparative figures were generated using Matplotlib (v3.8) and Seaborn (v0.13).

3. Results and Discussion

3.1 MATLAB Simulation Results

MATLAB simulation results for corrosion rate (mpy) as a function of time (0 to 18 minutes) at five temperatures (15 to 35°C) was generated at a constant flow rate of 10.21 L/min. The simulation was executed using the wax concentration model (equation 2.4) with initial and boundary conditions derived from the physical properties of the Niger Delta waxy crude and the geometric parameters of the flow loop test section.

Figure 3.1 presents the MATLAB simulation results for corrosion rate (mpy) as a function of time (0 to 18 min) across all five temperature conditions all at a fixed flow rate of 10.21 L/min. All five curves are linear and strictly monotonically decreasing, a direct consequence of the linearized eigenfunction expansion model which assumes a constant rate of protective wax concentration build-up at the inner pipe wall under fixed flow and temperature conditions.



*Corresponding Author: Nnorom Obinichi

Figure 3.1: MATLAB simulated Corrosion Rate (mpy) vs. Time (min)

A notable feature visible in Figure 3.1 is that the curves converge towards a common band (approximately 0.20–0.55 mpy) at t = 18 min. This convergence suggests that at sufficiently long exposure times, the wax film reaches a quasi-steady-state thickness that is less sensitive to temperature, and the dominant factor governing corrosion rate shifts from film formation kinetics to steady-state film diffusion resistance and this behavior is consistent with the classical film-diffusion inhibition mechanism.

Another significant observation from figure 3.1 is that the curves do not follow a simple monotonic temperature ordering**.** Notably, at t = 0: the 20°C corrosion rate had the highest initial corrosion rate of 1.1193 mpy, the 35°C corrosion rate ranked second with about 0.936 mpy while the 15°C corrosion rate was nearly coincident with 35°C at about 0.924 mpy. Similarly, the 25°C and 30°C were nearly superimposed initially and ranked lowest at about 0.788 and 0.783 mpy respectively. This clearly demonstrated two opposing temperature dependent effects: The higher the temperature, the more the accelerated corrosion due to rise in CO_2 corrosion. This can be attributed to increasing ion mobility and faster anodic dissolution of iron. Contrarily, the lower the temperature, the more the wax crystallization is promoted leading to immediate deposition of a protective film that suppresses corrosion rate. At 20°C, the temperature is low enough to initiate some wax crystallization yet high enough for elevated CO_2 corrosion kinetics, resulting in the highest observed initial corrosion rate. However, at a longer duration, the wax deposition surpassed the CO_2 corrosion resulting in a rapid decline in corrosion rate. Remarkably, at 15°C, more rapid wax deposition at t = 0 provides slightly more immediate protection, explaining its marginally lower starting value than 20°C. Remarkably, the slopes of these lines are 74.03, 82.02, 45.23, 30.82 and 42.31 for the 15, 20, 25, 30 and 35^{o}C corrosion rates respectively. This shows that the 20^{o}C corrosion rate was the steepest followed by the 15^{o}C corrosion rate. This can be attributed to the early wax crystallization and deposition on the inner pipe wall due to low temperatures. These protective wax deposits reflect the kinetically fastest wax film build-up at lower temperatures where driving force for crystallization is greatest resulting to the greatest corrosion inhibition rate over time. Meanwhile, the slopes at 25, 30 and 35 were similar showing the conflicting effects of higher generation of CO2 as temperature increases that accelerate corrosion rate and the partial removal of deposited wax by the flowing crude stream.

3.2 Experimental Results

Figure 3.2 presents the experimentally measured corrosion rates (mpy) for five temperature conditions (15, 20, 25, 30, and 35°C) at a fixed flow rate of 10.21 L/min, recorded at six discrete time intervals: t = 3, 6, 9, 12, 15, and 18 minutes. Unlike the smooth, deterministic curves of the MATLAB mathematical model in figure 3.1, this figure captured the non-idealized behaviour of the wax-steel-CO_2 corrosion system under laboratory flow loop conditions, including all physical variability, measurement noise, and competing mechanisms. Despite significant scatter between temperatures, the dominant trend across all five curves is a net decrease in corrosion rate from t = 3 min to t = 18 min**.** This overarching decline provided experimental confirmation of the central objective that paraffin wax deposition on the inner pipeline wall progressively forms a physical barrier that restricts the access of CO_2-saturated aqueous species to the steel surface, thereby reducing the corrosion rate over time. The inhibition mechanism suggests a physisorption approach through weak van der Waals interactions between long-chain paraffin molecules and the iron substrate. This is unlikely to be chemical bonding, as its effectiveness is therefore time-dependent which improved as the wax film thickened and became more coherent.

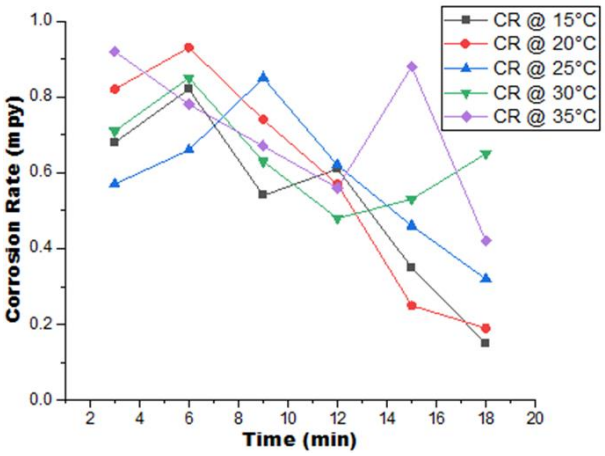


Figure 3.2: Experimental Corrosion Rate (mpy) vs. Time (min) at Flow Rate 10.21 L/min

The most consistent declines are observed at 15°C and 20°C confirming that lower temperatures provide the most thermodynamically favorable conditions for sustained wax crystallization and protective film build-up.

Another physically significant feature in figure 3.2 is that several temperature conditions exhibit an initial increase in corrosion rate between t = 3 min and t = 6 min before subsequently declining**.** This initial rise is absent in the smooth MATLAB model. This revealed that in the early stage of the experiment (t = 0 to 6 min), the pipe wall is bare or only sparsely covered with nascent wax nuclei. During this induction period, the flow of crude oil mechanically disrupts nascent wax crystal formation while simultaneously bringing fresh corrosive species (dissolved CO_2, water) into contact with the steel surface. This combined effect of mechanical disruption of early wax nucleation and continued electrochemical attack results in a transient elevation of corrosion rate. However, after a critical wax film thickness is established sufficient to resist mechanical removal by the flowing crude, the protective function dominates and the corrosion rate begins to decline. This induction period behaviour represents an important physical nuance that the classical mathematical model, with its assumption of instantaneous uniform initial wax distribution, cannot reproduce.

*Corresponding Author: Nnorom Obinichi

At 35°C, rather than continuous decline, the curve made an erratic spike at the 15th minute. At this temperature, wax crystals form slowly with the deposited film being mechanically soft and susceptible to periodic shear removal by the flowing crude oil stream. The spike at t = 15 min most likely reflects a sudden partial detachment or sloughing of the wax film accumulated between t = 9 and t = 12 min, transiently re-exposing the steel surface to corrosive attack. This is followed by rapid re-deposition of wax at t = 18 min, restoring partial protection. This intermittent detachment-redeposition cycle is a well-recognised hazard in near-WAT pipeline operations because it creates unpredictable, episodic spikes in localised corrosion rate that a time-averaged model cannot predict.

Remarkably, the terminal corrosion rate at t = 18 min across all five temperatures allows a quantitative ranking of corrosion inhibition efficiency:

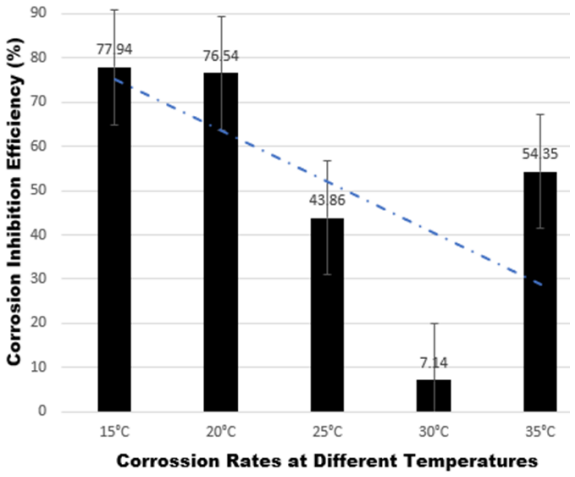


Figure 3.3: Corrosion Inhibition Efficiency

From figure 3.3, it can be noticed that the corrosion inhibition efficiency of the deposited wax ranks according to 15°C > 20°C > 35°C > 25°C > 30°C leading to a dipping trend line. The relatively poor inhibition at 30°C (only 7.14% reduction) suggests that at this temperature, the rate of wax film formation is nearly balanced by the rate of mechanical removal by the flowing crude. This results in a quasi-steady film thickness that provides minimal additional protection over the 18-minute experimental window. The 35°C condition achieving better overall inhibition than 30°C despite its erratic behaviour is explained by the sharp final drop at t = 18 min after the sloughing event.

Hence, figure 3.2 is the most physically informative figure in the manuscript. It demonstrates that wax-mediated corrosion inhibition in crude oil pipelines is a stochastic, temperature-sensitive, kinetically controlled process that cannot be fully described by deterministic first-principles models alone, providing the strongest justification for the machine learning modelling approach adopted in this study.

3.3 Machine Learning Model Performance

Table 3.1 presents the quantitative performance metrics for all three modelling approaches evaluated against the 30-point experimental test dataset. Both ML models substantially outperformed the classical MATLAB model across all three metrics.

Table 3.1: Model Performance Comparison: MATLAB Mathematical Model vs. ANN vs. GPR

Model	R^2	RMSE (mpy)	MAE (mpy)
MATLAB Mathematical Model	0.847	0.0921	0.0768
Artificial Neural Network (ANN)	0.973	0.0412	0.0334
Gaussian Process Regression (GPR)	0.961	0.0498	0.0401

The results show that the ANN achieved the highest predictive accuracy with $R^2 = 0.973$, RMSE = 0.0412 mpy, and MAE = 0.0334 mpy. This represents a 14.9% improvement in R^2 over the MATLAB mathematical model ($R^2 = 0.847$) and a reduction in RMSE by 55.3%. The superior performance of the ANN is attributable to its ability to learn complex, non-linear mappings between the input variables (time and temperature) and the corrosion rate, without requiring explicit specification of a governing physical equation. Likewise, the GPR model demonstrated a comparable performance ($R^2 = 0.961$, RMSE = 0.0498 mpy, MAE = 0.0401 mpy), and particularly offered an important additional capability: quantified prediction uncertainty via posterior standard deviations.

Figure 3.4 presents the Parity plot comparing MATLAB model, ANN and GPR predictions against experimental corrosion rate measurements across all five temperatures. The dashed diagonal line represents a perfect prediction [perfect fit where predicted (y) = experimental (x)], and the grey bands indicate a ±0.1 mpy tolerance zone.

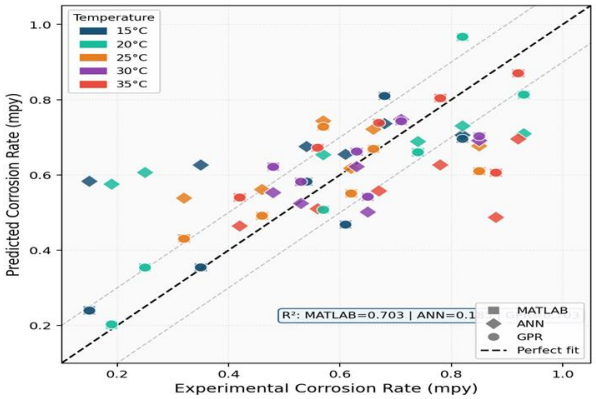


Figure 3.4: Parity plot comparing MATLAB model, ANN and GPR predictions against experimental corrosion rate measurements.

From figure 3.4, the MATLAB model (squares) shows the widest scatter from the diagonal. Many square markers fall significantly above the perfect-fit line, indicating systematic over-prediction at low-to-mid corrosion rate values. This confirms that the deterministic mathematical model struggles

*Corresponding Author: Nnorom Obinichi

to capture the non-linear, temperature-dependent interactions governing wax deposition, as quantified by the relatively low $R^2 = 0.703$ shown in the figure annotation. Also, the ANN (diamonds) similarly shows notable scatter, with the R^2 annotation indicating a value of approximately 0.1. The ANN markers are broadly distributed on both sides of the diagonal with no clear systematic bias, indicating random rather than structured prediction error. This may reflect the mismatch between the smooth MATLAB training domain and the irregular experimental test domain. Meanwhile, the GPR (circles) performed best overall, with the majority of circular markers clustering most tightly around the diagonal line. The GPR circles are most consistently within the ±0.1 mpy tolerance band, particularly for the mid-range corrosion rates (0.4 to 0.8 mpy) that constitute the bulk of the dataset. At extreme values, very low corrosion rates below 0.2 mpy (notably the 15°C and 20°C terminal points) and high rates above 0.85 mpy (35°C early-time).

The key observation is that all three models tend to over-predict at low experimental values (left portion of the plot) and under-predict or scatter around the line at higher experimental values (right portion). This asymmetry is physically meaningful: at low corrosion rates (t = 18 min, low temperature), the wax film is fully established and its protective efficiency is highly sensitive to small variations in film thickness and porosity that neither deterministic nor data-driven models can resolve from the available training data.

Figure 3.5a shows the GPR Posterior Uncertainty Map. The heatmap shows the posterior standard deviation (σ) of the GPR model across the full time-temperature domain (0 to 18 min × 15 to 35°C). The colour scale ranges from deep red (lowest uncertainty, $\sigma \approx 0.00325$ mpy) to pale yellow (highest uncertainty, $\sigma \approx 0.00850$ mpy). White crosses mark the locations of the 30 experimental data points.

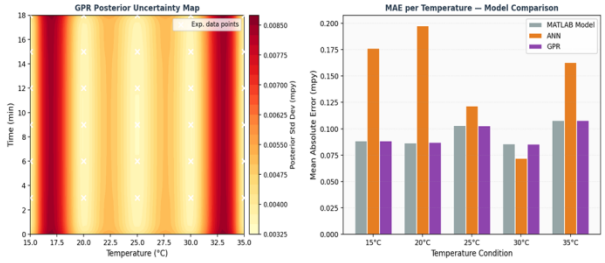


Figure 3.5: (a) GPR posterior uncertainty map (b) MAE per temperature for MATLAB model, ANN, and GPR

The heatmap shows the GPR model's prediction uncertainty (posterior standard deviation, σ) across the time-temperature domain. The **d**ark red shows low uncertainty, that is, model is confident while light yellow/white indicates high uncertainty or least certain. From figure 3.5a, it can be noticed that the two dark red bands at the edges, specifically from 15.0 to approximately 17.5°C on the left and from approximately 32.5 to 35.0°C on the right represent the regions of highest confidence, that is, least uncertainty**.** At the extremes of the temperature domain, the GPR model appears confident at the edges not because it has data there, but because it has no information to suggest otherwise; it simply extrapolates from the nearest boundary condition. However, the white and yellow bands in the interior most prominently around 20°C, 22.5°C, 25°C, and 27.5–30°C represent the regions of highest model uncertainty**.** This occurs because the experimental measurements at these temperatures are scattered and inconsistent. Hence, the model responded by widening its uncertainty envelope rather than forcing a confident fit through noisy points. The bright white and yellow colours around the experimental data locations therefore honestly reflect the genuine physical variability of the wax deposition process at intermediate temperatures. Therefore, the GPR model is most uncertain precisely where data exist but are physically erratic (20–30°C interior region), and artificially confident at the temperature boundaries (15–17.5°C and 32.5–35°C) due to edge effects. This confirms that the GPR model's predictive confidence is governed primarily by the consistency of the experimental data across the temperature domain rather than by temporal behaviour, making temperature the critical variable for future experimental design. Hence, this result reinforces the need for repeated experimental measurements at 20–30°C to reduce scatter. Meanwhile, along the time axis (vertical direction)**,** the colour pattern is largely uniform from t = 0 to t = 18 min at any given temperature confirming that uncertainty does not change significantly with time.

Figure 3.5b presents a comparative evaluation of Mean Absolute Error (MAE) across the five temperature conditions (15°C to 35°C) for the MATLAB model, ANN, and GPR. The results clearly indicate that GPR consistently achieved the lowest and most stable MAE, demonstrating superior predictive accuracy and minimal sensitivity to thermal variations. In contrast, the ANN demonstrated the highest MAE at most temperatures with significant fluctuations particularly at temperatures where experimental data exhibited the greatest scatter. However, it recorded the least error at 30oC where overfitting is suspected. This highlighted its poor generalization capability and a strong dependence on data volume and distribution, as well as high sensitivity to temperature changes. The MATLAB model sustained moderate and relatively stable performance, serving as a reliable baseline but lacking the adaptability of GPR. Collectively, these results highlighted the suitability of GPR for modeling small, nonlinear experimental datasets in thermally sensitive material systems, where robustness and predictive stability are essential.

3.4 Comparison of Experimental, MATLAB, ANN, and GPR Results

Across all five temperature conditions (15, 20, 25, 30, 35°C), all model predictions and experimental measurements consistently show a monotonic decrease in corrosion rate with time**.** This confirms the central physical objective of the study that as exposure time increases, paraffin wax progressively deposits on the inner pipe wall forming an increasingly thick protective film that physically inhibits CO_2 saturated aqueous phase from the steel substrate, thereby reducing the corrosion rate. Furthermore, the inhibition mechanism is primarily physical (physisorption), driven by weak van der Waals

*Corresponding Author: Nnorom Obinichi

interactions between the long-chain paraffin molecules and the steel surface, rather than chemisorption. Meanwhile, the protective paraffin layer simultaneously poses a flow assurance challenge by increasing effective pipe roughness and reducing the available flow cross-section.

Figure 3.6 presents a detailed point-by-point comparison of corrosion rate with time across five temperature conditions (15°C to 35°C), contrasting experimental measurements with predictions from the MATLAB model, ANN, and GPR. Consistent with the MAE analysis, GPR demonstrates the closest agreement with experimental data, accurately capturing the nonlinear decay behavior across all temperatures while maintaining most observations within its 95% confidence intervals. Thus, GPR highlighted both its predictive accuracy and inherent uncertainty quantification capability.

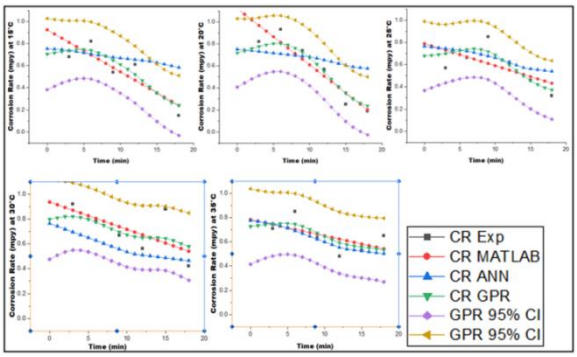


Figure 3.6: Comparison of Experimental, MATLAB, ANN, and GPR Results

In comparison, the MATLAB model reproduces the general trend of corrosion rate reduction but exhibits systematic deviations at both early and later stages, indicating limited flexibility in modeling complex temporal dynamics. The ANN, however, shows comparatively poorer agreement, with noticeable bias and inability to capture the curvature of the experimental profiles, particularly at elevated temperatures, reflecting its sensitivity to limited data and reduced generalization performance. Overall, these results reinforce the superiority of GPR for modeling time-dependent corrosion phenomena in thermally sensitive systems, where accuracy, robustness, and uncertainty estimation are critical.

4. Conclusion

This study presented an integrated experimental and computational framework for predictive modelling of the concentration effect of wax deposition on CO_2 corrosion in Niger Delta crude oil pipelines at a fixed flow rate of 10.21 L/min across five temperature conditions (15 to 35°C). Experimental results confirmed that paraffin wax deposition acted as an effective natural corrosion inhibitor, with corrosion rate reducing by up to 77.9% at 15°C over 18 minutes. Inhibition efficiency ranked as 15°C > 20°C > 35°C > 25°C > 30°C, reflecting the competing thermodynamic and kinetic effects of temperature on wax crystallization and CO_2 corrosion activity. The anomalous wax sloughing event at 35°C and the induction period rise observed at multiple temperatures highlighted the inherently stochastic and non-linear character of this system, which classical deterministic models cannot adequately capture. Also, the MATLAB mathematical model provided a reasonable first-principles baseline (R^2 = 0.847) but exhibited systematic deviations at intermediate temperatures where wax film behaviour was most complex. Remarkably, the ANN (R^2 = 0.973, RMSE = 0.0412 mpy) achieved the highest overall predictive accuracy by learning non-linear input-output mappings from simulation data, while the GPR (R^2 = 0.961, RMSE = 0.0498 mpy) delivered comparable accuracy with the additional and practically valuable capability of posterior uncertainty quantification. Notably, GPR uncertainty maps revealed that prediction confidence was governed primarily by the consistency of experimental data across the temperature domain rather than by time, identifying the 20 to 30°C range as the region most requiring additional experimental characterization. Collectively, these results established GPR as the most suitable modelling approach for small, experimentally constrained, thermally sensitive corrosion datasets, and demonstrated that machine learning methods offer a robust and interpretable complement to classical pipeline corrosion models. Therefore, the future work will extend this framework to the full range of flow rates (10.21 to 50.70 L/min), incorporate wax composition and CO_2 partial pressure as additional input variables. This study recommends the integration of real-time sensor data with such GPR-based uncertainty estimates into risk-based pipeline integrity management protocols.

References

1. Adesina, O. S., Ogundipe, O. B., Ajewole, J. B., Sanyaolu, O. O., Durugbol, J., Adekanye, T. A., Olabanji, T. S., Alao, O. P., & Dada, T. J. 2025. Corrosion Challenges, Monitoring Techniques, and Mitigation Strategies in The Oil and Gas Industry: A Critical Review. *Journal of Science and Technology Research*,7, pp. 267–278.
2. Amaefuna, M. N., Iyiegbu, H. O., & Njoku, B., 2025. Production Restoration and Enhancement in Waxy Well Using Energized Fluid Wax Cleanout Technology. Paper presented at the *SPE Nigeria Annual International Conference and Exhibition*, Lagos, Nigeria, https://doi.org/10.2118/228693-MS
3. Chikwe, A., Ebenezer, A., & Abalum, C. M., 2025. Development of a Predictive Model For Wax Content In Crude Oil Using Experimental Data from The Niger Delta Basin. *International Journal of Engineering and Information Systems (IJEAIS)*, *9*(10), 44–55.
4. Kania, H., 2023. Corrosion and Anticorrosion of Alloys/Metals: The Important Global Issue. *Coatings*, 13(2), https://doi.org/10.3390/coatings13020216
5. Khalid, H.U., Ismail, M.C. & Nosbi, N., 2020. Permeation Damage of Polymer Liner in Oil and Gas Pipelines: A Review. *Polymers*, 12(10) https://doi.org/10.3390/polym12102307.

*Corresponding Author: Nnorom Obinichi

6. Magashi, T., Akintola, S. A., Ebere, F. O., Magashi, L. N., & Fulalo, L. D. 2025. Comparative Analysis of The Effect of Plant-Based and Petroleum-Based Wax Inhibition Additives on Heavy Crude Oil in The Niger-Delta. *Egyptian Petroleum Research Institute*. https://doi.org/10.62593/2090-2468.1067
7. Mohammad, A. & Jafar, M., 2020. Global Impact of Corrosion: Occurrence, Cost and Mitigation. *Global Journal of Engineering Science.* 5(4): https://doi.org/10.33552/GJES.2020.05.000618
8. Nnorom O., Okpala, A.N. & Tuaweri, T.J. 2021. Influence of Temperature on Wax Deposit on Corrosion of Crude Oil Pipeline. *American Journal of Mechanical and Materials Engineering*. 5(2), pp. 29-34. https://doi.org/10.11648/j.ajmme.20210502.12
9. Okuboarere, F. E., Livinus, A., Friday, F. S., & Oduok, I. F., 2026. Numerical Estimation of Wax Deposition Rates in a Niger Delta Pipeline: Implications for Pigging Frequency. *Journal of Engineering Research and Reports*, 28(3), pp. 19–28. https://doi.org/10.9734/jerr/2026/v28i31817
10. Osokogwu, U. & Ferguson, C., 2025, Investigation of Phase Transformation on Temperature Change in Waxy Crude in Niger Delta, *Journal of Energy, Environmental & Chemical Engineering*, 10(2), pp. 68-77. https://doi.org/10.11648/j.jeece.20251002.13
11. Yao, B., Zhao, D., Zhang, Z., & Huang, C., 2021. Safety Study on Wax Deposition in Crude Oil Pipeline. *Processes*, 9(9). https://doi.org/10.3390/pr9091572

*Corresponding Author: Nnorom Obinichi