Proceedings of ISFOG 2025
5TH INTERNATIONAL SYMPOSIUM ON
FRONTIERS IN OFFSHORE GEOTECHNICS
Nantes, France | June 9-13 2025

ISBN 978-2-85782-758-0

A Comparison of Machine Learning Algorithms for Predicting Fatigue Life of Steel Catenary Risers

Lusheng Jia
National Energy Deepwater Oil and Gas Engineering Technology Research and Development Center, Beijing, China

Wanchao Wu*, Cheng Zhang, Zhichao Shen
South China University of Technology, Guangzhou, Guangdong, China

Yi Liu
National Energy Deepwater Oil and Gas Engineering Technology Research and Development Center, Beijing, China

**202310193763@mail.scut.edu.cn* (*corresponding author*)

ABSTRACT: Steel Catenary Risers (SCR) are essential components in offshore oil and gas development, yet their fatigue life assessment remains challenging due to the combined effect of vessel motion, hydrodynamic loads, and seabed soil conditions. Current dynamic methods for flexible structures, such as the lumped mass method in OrcaFlex, struggle with large deformation analysis. This study employs the Absolute Nodal Coordinate Formulation to model SCR, integrating environmental loads, vessel motion, and nonlinear pipe-soil interactions. Based on the numerical model, a database of 10,000 fatigue scenarios were constructed to train and test machine learning models for quick prediction of SCR fatigue life. In this study, traditional machine learning models, including Linear Regression, Polynomial Regression, and Ridge Regression, were compared with deep learning models, including Multilayer Perceptron, Deep Neural Networks and Transformer for predicting SCR fatigue life. The results demonstrate that the deep learning models outperform traditional machine learning methods, offering an effective tool for fatigue life prediction in offshore engineering.

Keywords: Steel Catenary Risers; Fatigue life prediction; Machine Learning; Offshore engineering

1 INTRODUCTION

Steel Catenary Risers (SCR), which are steel pipes suspended in a catenary shape from floating platforms to subsea infrastructure, serve as vital conduits in offshore oil and gas development. Their fatigue life assessment is challenging due to the combined effect of multiple factors, such as vessel motion, hydrodynamic loads, and pipe-soil interaction (Liu et al., 2024). The prediction of SCR fatigue life often relies on experimental data or finite element method such as OrcaFlex, which struggles to deal with SCR large deformation effectively (Zhang et al., 2019). By comparison, the Absolute Nodal Coordinate Formulation (ANCF) is particularly well-suited for large deformation analysis for SCR (Gerstmayr & Shabana, 2006; Zhang et al., 2019).

Traditional numerical method requires extensive simulations to quantify SCR fatigue, leading to very high computational costs (Hejazi et al., 2022). To address these challenges, machine learning models are increasingly used to assess SCR fatigue life (Sivaprasad et al., 2023). In this study, ANCF is utilised to construct the database of SCR fatigue scenarios, considering the combined effect of hydrodynamic loads, vessel motion, and nonlinear pipe-soil interaction. Based on the constructed database, various machine learning algorithms are evaluated regarding the prediction accuracy and efficiency. Both traditional machine learning approaches and deep learning models are considered to demonstrate their potential as effective tools for fatigue life prediction in offshore engineering applications.

2 MACHINE LEARNING ALGORITHMS

2.1 Traditional Machine Learning Methods

2.1.1 Linear Regression (LR)

Linear regression assumes a linear relationship between dependent variable y and the independent variables $x = [x_1, x_2, \dots, x_p]^T$, expressed as:

$$y = \beta_0 + \sum_{j=1}^{p} \beta_j x_j + \varepsilon \quad (1)$$

where β_0 is the intercept term. β_j are the regression coefficients, reflecting the influence of the j-th independent variable on the dependent variable. ε is the error term, assumed to follow a normal distribution $\varepsilon \sim N(0, \sigma^2)$ (Montgomery et al., 2021).

2.1.2 Polynomial Regression (PR)

Polynomial regression extends linear regression by introducing higher-order terms of the independent variable to capture nonlinear relationships (Aiken & West, 1991). The general form of a polynomial regression model of degree d can be represented as:

$$y = \beta_0 + \sum_{j=1}^{d} \beta_j x^j + \varepsilon \quad (2)$$

It is important to note that the model remains linear in parameters β, allowing Ordinary Least Squares (OLS) estimation (Draper & Smith, 1998). Increasing polynomial degree enhances flexibility but also increases the risk of overfitting, which can be mitigated by cross-validation (James et al., 2013).

2.1.3 Ridge Regression (RR)

Ridge regression prevents overfitting by adding an L_2 (the squared magnitude of coefficients) regularization term to linear or polynomial regression, penalizing large regression coefficients and thereby controlling model complexity and improving generalization (Hoerl & Kennard, 2000). The objective function for ridge regression is:

$$\min_{\beta} J(\beta) = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \quad (3)$$

where λ is the regularization parameter that controls the strength of the penalty. A higher λ shrinks coefficients, reduces multicollinearity, and improves generalization.

2.2 Deep Learning Models

2.2.1 Multilayer Perceptron (MLP)

Multilayer Perceptron (MLP) is a type of feedforward neural network that typically consists of an input layer (x), 1 to 2 hidden layers (h_1,h_2), and an output layer (y), with weight coefficients (w_1, w_2) connecting the layers, as shown in Figure 1 (Heaton et al., 2018). MLP is one of the earliest neural network models, well-suited for simple tasks due to its straightforward architecture.

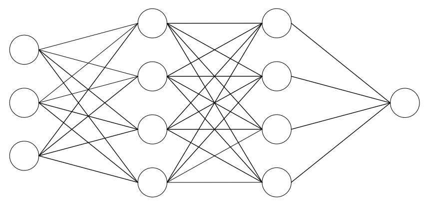


Figure 1. General structure of MLP (Heaton et al., 2018)

2.2.2 Deep Neural Network (DNN)

Deep Neural Network (DNN) is an extension of the Multilayer Perceptron (MLP), introduced by Hinton et al. in 2006 (Hinton et al., 2006). DNN feature multiple hidden layers, allowing them to learn complex features and capture nonlinear relationships, offering capabilities beyond MLP.

DNN commonly use the Rectified Linear Unit (ReLU) as the activation function in their hidden layers, as it mitigates the vanishing gradient problem and enhances training efficiency (Nair & Hinton, 2010). This makes ReLU superior to classical activation functions, such as Sigmoid and Tanh for deep networks.

2.2.3 Transformer

The Transformer model, proposed by Vaswani et al. (2017), captures global dependencies in sequences using self-attention instead of recurrent or convolutional structures. It has achieved significant success in Natural Language Processing, computer vision, and time-series forecasting (Devlin et al., 2019). Despite its effectiveness, its use in regression analysis remains limited, offering opportunities for further exploration. The architecture of transformer comprises an Encoder and Decoder with key components, such as Multi-Head Self-Attention, Feed-Forward Networks, and Residual Connections, which enhance learning capacity, stability, and convergence (Brown et al., 2020).

3 METHODOLOGY

This section describes methodologies for modeling and analysing SCR fatigue behavior. The proposed framework employs a multi-stage computational workflow structured into four interdependent modules (Table 1):

Table 1. System Overview of the Framework

Module	Components
1. Preprocessing Module	-Riser parameters -Seabed/wave/platform parameters
2. Macro Element & Load Module	-Pipe-soil interaction (macro-element model) -Pipe-water interaction (hydrodynamics)
3. Core Module	-Parameter description of ANCF -Riser nonlinear deformation model -Cyclic stress statistics
4. Riser Performance Module	-Riser motion/shape calculation -Strain/curvature/stress distribution -Fatigue damage quantification

3.1 Modeling SCR Based on ANCF

ANCF directly maps the configuration of structure to the global coordinate system, avoiding complex transformations, and uses slope coordinates to describe geometric nonlinearities (Gerstmayr & Shabana, 2006). SCR in this study are discretized into beam elements with nodes having six degrees of freedom, and Hermite cubic polynomials is used as shape functions (Zhang et al., 2019). Internal forces include elastic, generalised external, and inertial forces, with elastic forces derived from strain and curvature. Hydrodynamic force is modeled using the Morison equation, which calculates the dynamic hydrodynamic resistance by considering the relative velocity between the structure and the fluid, the force area, and the drag coefficient. The upper vessel moves on the water surface with the waves, and its motion is simulated by altering the absolute coordinates of the top point in the time domain (Zhang et al., 2019). This method effectively models SCR nonlinear dynamics and has been validated in practical applications (Wang et al., 2017). The Newton-Raphson method is used for static analysis, with numerical integration for stiffness calculations (Matikainen et al., 2009).

3.2 Pipe-soil Interaction Model

The vertical pipe-soil interaction is modeled using the RQ model (Randolph & Quiggin, 2009) to simulate SCR-seabed contact with a nonlinear hysteretic approach. The model captures seabed hysteresis and cumulative penetration across four modes: Initial penetration, Uplift, Repenetration, and Not in Contact, which is validated by lab and field tests (Randolph & Quiggin, 2009; Shiri et al., 2014). For lateral pipe-soil interaction, a modified Coulomb friction model is employed, and the lateral pipe-soil model parameters are not discussed in this paper.

3.3 Fatigue Analysis

The *S-N* curve method is used for fatigue analysis:

$$N = aS^{-m} \tag{4}$$

where N is the number of stress cycles, a is a material constant, S is the stress amplitude, and m are material-specific parameters. Miner's theory for cumulative fatigue damage is:

$$D = \sum_{i=1}^{k} \frac{n_i}{N_i} \tag{5}$$

where n_i is the number of cycles for a specific stress range, and N_i is the number of cycles required for fatigue failure at that stress range. Rainflow counting is employed to calculate stress cycles and amplitudes, and $D \geq 1$ indicates failure (Liu et al., 2024).

4 VALIDATION

Figure 2 shows the annual fatigue damage (fatigue life = 1 / annual fatigue damage) results at the riser sections of 0°, 45°, 90°, 135°, and 180° near the touchdown point under Cross-current conditions, calculated using both the ANCF and OrcaFlex methods (SCR parameters are shown in Table 2). The maximum fatigue damage calculated by ANCF at each angle is slightly higher than that calculated by OrcaFlex, primarily because ANCF more comprehensively accounts for the large deformation characteristics of the structure, resulting in higher computed stress levels. Overall, the fatigue damage results calculated by both methods are relatively close, verifying the accuracy of the ANCF algorithm in fatigue analysis.

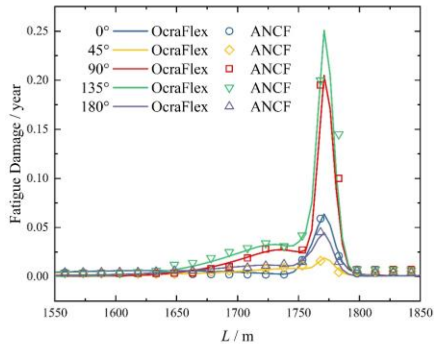


Figure 2. Validation of ANCF model

5 APPLICATION OF MACHINE LEARNING ALGORITHMS

5.1 Database Preparation

The Response Amplitude Operator (RAO) is used to represent the dynamic response of the floating structure, as shown in Figure 3, which is obtained through frequency-domain potential flow analysis under irregular sea states. The sway, surge, and heave data of the floating structure for different wave periods are obtained and subsequently utilised as inputs for fatigue analysis.

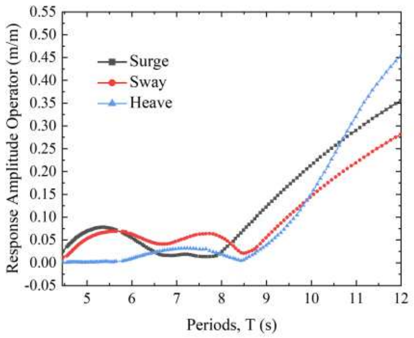


Figure 3. Response Amplitude Operator of Vessel

The SCR parameters for the whole database remained constant, as shown in Table 2.

Table 2. The main parameters for SCR.

Parameters	Values	Parameters	Values
Total Length (m)	1733	Tensile Stiffness (N)	7.54E+09
Outer Diameter (m)	0.324	Drag Coefficient	1.2
Wall Thickness (m)	0.041	Added Mass Coefficient	1
Submerged Weight (kg/m)	94.39	Fatigue S-N *a*	1.05E+12
Yang's module (N/m²)	2.07E+11	Fatigue S-N *m*	3

Table 3 summarises the range of parameters for environmental and seabed conditions considered in this study. Key environmental parameters include wave height, wave angle, and sea surface velocity, while seabed properties such as stiffness, shear strength gradient, suction resistance coefficient, suction decay parameter, and re-penetration parameter were considered. Random values were generated by sampling according to a standard normal distribution within the specified ranges, to create a diverse database that captures a wide range of potential environmental and seabed conditions.

A total of 10,000 groups of input parameters were supplied to the ANCF model to calculate the SCR fatigue life. This provided a robust database for subsequent SCR fatigue life prediction, with 80% of data used for training and the remaining 20% for testing.

Table 3 The range of parameters for environmental and seabed conditions

Parameters	Min Value	Max Value
Wave height (m)	1	18
Wave angle (deg)	0	360
Velocity at sea surface (m/s)	0.92	2.3
The seabed stiffness (N/m²)	0	3500
Shear Strength Gradient (N/m²/m)	1300	6000
Suction Resistance Ratio	0	0.3
Suction Decay Parameter	0.2	0.6
Repenetration Parameter	0.1	0.5

5.2 Evaluation Indicators

In this study, R-squared (R^2) and Root Mean Squared Error (*RMSE*) are used to evaluate the performance of the regression models.

The R^2 formula is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{6}$$

The *RMSE* is calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

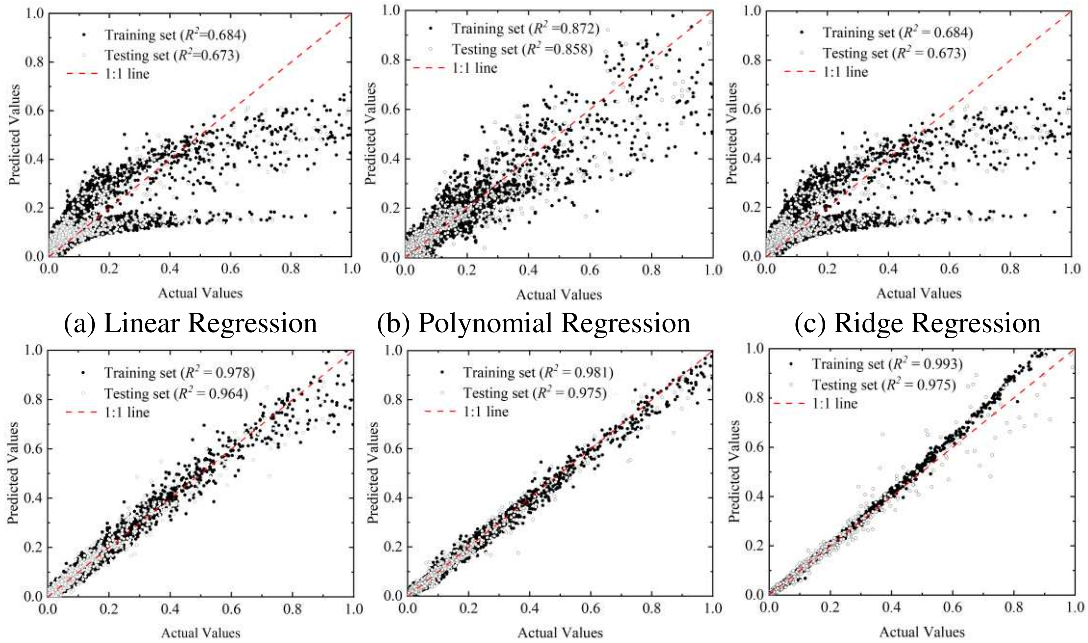
where, y_i is actual value, $\hat{y}_i$ is predicted value, $\bar{y}_i$ is mean of actual values, n is number of observations. R^2 measures the proportion of variance explained by the model, ranging from 0 to 1, with 1 indicating perfect prediction. *RMSE* quantifies prediction error as the square root of the average squared differences between predicted and actual values, and the smaller the value, the better the performance.

5.3 Results and Analysis

Figure 4 and Table 4 compare the prediction accuracy, efficiency, and hyperparameters of several machine learning models, including LR, PR, RR, MLP, DNN, and Transformer, for predicting fatigue life based on the constructed database. Scatter plots show the comparison between actual (calculated by

ANCF) and predicted (predicted by machine learning models) values of the annual fatigue damage (fatigue life =1 / annual fatigue damage) of the SCR, with the ideal 1:1 line in red.

As summarised in Table 4, LR shows poor prediction capability due to its simplicity with R^2 for training set less than 0.7. PR improves accuracy by accounting for non-linearities. RR achieves the same prediction performance as LR, suggesting regularization is unnecessary here. Compared to traditional machine learning models, MLP and DNN significantly improve prediction accuracy, with R^2 exceeding 0.9, though computation time also increases. The Transformer achieves the highest prediction accuracy, which requires the most computational time.



(a) Linear Regression (b) Polynomial Regression (c) Ridge Regression

(d) Multilayer Perceptron (e) Deep Neural Network (f) Transformer

Figure 4. Comparison of prediction accuracy of machine learning algorithms

Table 4 Hyperparameter Settings and evaluation indicators

Algorithms	**Hyperparameter**	**Training set**		**Testing set**		**Time ratio**[a]
		RMSE	***R^2***	***RMSE***	***R^2***	
LR	/	0.078	0.684	0.078	0.673	1
PR	degree=2	0.049	0.872	0.052	0.858	0.43
RR	alpha=0.001	0.078	0.684	0.078	0.673	0.38
MLP	alpha = 0.001; activation = ReLU; hidden layers = 1; Hidden layer size = 100; learning rate = 0.001; shuffle = True	0.021	0.978	0.026	0.964	2.73
DNN	activation = ReLU; hidden layers = 2; hidden layer size = 64; learning rate = 0.001; epochs = 100; batch size = 32; shuffle = True	0.019	0.981	0.021	0.975	11.29
Transformer	activation = ReLU; hidden layers = 1; hidden layer size = 64; nheads = 4; learning rate = 0.0001; epochs = 100; batch size = 32; dropout = 0.1; shuffle = True	0.012	0.993	0.022	0.975	36.18

Notes: [a] is the computational time of each model relative to the computational time of LR. (LR required 39 seconds on an Intel Core i7-13700 CPU, 2.1 GHz, 8 cores).

6 CONCLUSIONS

In this study, the ANCF was used to construct a database of SCR fatigue analysis for training machine learning models. Traditional machine learning models, including LR, PR, and RR, were compared with deep learning models, including MLP, DNN and Transformer to predict SCR fatigue life. This study demonstrates the superior prediction accuracy of deep learning models. However, the computational cost of the Transformer model is significant, making DNN or MLP more practical for time-sensitive applications.

AUTHOR CONTRIBUTION STATEMENT

Lusheng Jia: Investigation, Writing- Original draft. **Wanchao Wu**: Software, Script Development. **Cheng Zhang**: Methodology. **Zhichao Shen**: Supervision, Writing-Reviewing and Editing. **Yi Liu**: Conceptualization.

ACKNOWLEDGEMENTS

Zhichao Shen is supported by Guangdong Basic and Applied Basic Research Foundation (2023A1515240026) in China.

REFERENCES

Aiken, L. S., & West, S. G. (1991). Multiple Regression: Testing and Interpreting Interactions. Sage Publications, Inc.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P. et al,. (2020). Language Models are Few-Shot Learners. Advances in Neural Information Processing Systems, 33. https://doi.org/10.48550/arXiv.2005.14165

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT, 1, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Draper, N. R., & Smith, H. (1998). Applied Regression Analysis (3rd ed.). Wiley.

Gerstmayr, J., & Shabana, A. A. (2006). Analysis of Thin Beams and Cables Using the Absolute Nodal Co-ordinate Formulation. Nonlinear Dynamics, 45, 109–130. https://doi.org/10.1007/s11071-006-1856-1

Hejazi, R., Grime, A., Randolph, M., & Efthymiou, M. (2022). An efficient probabilistic framework for the long-term fatigue assessment of large diameter steel risers. Applied Ocean Research, 118, 102941. https://doi.org/10.1016/j.apor.2021.102941

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. Neural Computation, 18(7), 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527

Hoerl, A. E., & Kennard, R. W. (2000). Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics, 42(1), 80–86. https://doi.org/10.1080/00401706.2000.10485983

Heaton, J., Goodfellow, I., Bengio, Y., & Courville, A. (2018). Deep learning. Genet Program Evolvable Mach, 19, 305–307. https://doi.org/10.1007/s10710-017-9314-z

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: With Applications in R. Springer.

Liu, D., Li, S., Ai, S., Sun, L., & Guedes Soares, C. (2024). Fatigue analysis of steel catenary risers under coupled cross-flow and in-line vortex-induced vibrations with oblique incoming flow. Marine Structures, 95, 103578. https://doi.org/10.1016/j.marstruc.2024.103578

Matikainen, M. (2009). Development of beam and plate finite elements based on the absolute nodal coordinate formulation. Transplantation Proceedings, 43(1), 291-295.

Mekha, B. B. (2001). New frontiers in the design of steel catenary risers for floating production systems. Journal of Offshore Mechanics and Arctic Engineering, 123(3), 153. https://doi.org/10.1115/1.1410101

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). Introduction to Linear Regression Analysis (6th ed.). Wiley.

Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning. 807–814. https://doi.org/10.5555/3104322.3104425

Randolph M., Quiggin P. (2009). Non-linear hysteretic seabed model for catenary pipeline contact. Proceedings of the ASME 2009 28th International Conference on Ocean, Offshore and Arctic Engineering, Honolulu, HI, United states: American Society of Mechanical Engineers, 145-154. https://doi.org/10.1115/OMAE2009-79259

Shiri, H. (2014). Response of steel catenary risers on hysteretic non-linear seabed. Applied Ocean Research, 44, 20-28. https://doi.org/10.1016/j.apor.2013.10.006

Sivaprasad, H., Lekkala, M.R., Latheef, M., Seo, J., Yoo, K.H., Jin, C., & Kim, D.K. (2023). Fatigue damage prediction of top tensioned riser subjected to vortex-induced vibrations using artificial neural networks. Ocean Engineering. https://doi.org/10.1016/j.oceaneng.2022.113393

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. Advances in Neural Information Processing Systems, 30. https://doi.org/10.48550/arXiv.1706.03762

Wang, L., Currao, G., Han, F., et al. (2017). An immersed boundary method for fluid–structure interaction with compressible multiphase flows.

Journal of Computational Physics, 346, 131-151. https://doi.org/10.1016/j.jcp.2017.06.008

Zhang, C., Kang, Z., Ma, G., & Xu, X. (2019). Mechanical modeling of deepwater flexible structures with large deformation based on absolute nodal coordinate formulation. Journal of Marine Science and Technology, 24, 1241-1255. https://doi.org/10.1007/s00773-018-00621-0