Proc. of the 9th International Symposium on Geotechnical Safety and Risk (ISGSR)
Edited by Zhongqiang Liu, Jian Dai and Kate Robinson
Published by Research Publishing, Singapore. ©2025 The Author(s), CC-BY 4.0 licence.
doi: 10.3850/GRF-25280825_isgsr-203-P227-cd

OFFSHORE PIPELINE ROUTING OPTIMIZATION VIA PROBABILISTIC REINFORCEMENT LEARNING FOR VARYING LANDSLIDES' STABILITY

Billy Hernawan

Civil & Environmental Engineering, Texas A&M University, United States. E-mail: hernawa2@tamu.edu

Zenon Medina-Cetina, PhD

Civil & Environmental Engineering, Texas A&M University, United States. E-mail: zenon@tamu.edu

Optimizing offshore pipeline routes is a challenging problem, especially in environments with significant uncertainty, where site characterization data is limited. Current approaches in offshore pipeline design and optimization typically involves three stages: identifying the components of relevant geohazards, creating composite maps by pixel-per-pixel aggregation techniques of geohazards, and computing likely least-cost paths between any two points of reference (i.e., origin and destination). Dijkstra's algorithm remains the algorithm of choice in pipeline optimization as it is the default least-cost path determination implemented in most GIS software. However, such implementation can be computationally intensive for a large grid problem. Recently, Q-learning, a Reinforcement Learning (RL) method, was proposed to optimize subsea pipeline route design. Nonetheless, the current approaches assume that the 'cost' map that is used as the foundation for the optimization algorithm is fixed. This study proposes the modelling of pipeline routing based on a physics-based model (i.e., infinite slope model), the Bayesian paradigm, and the use of Policy Iteration. This results in a dynamic programming algorithm which uses a probability of failure field as a prior, and then optimizes any pipeline routing given a designated origin and destination positions. Preliminary results indicate that the optimized model consistently identifies the most optimal path between the given origin and destination positions, and that the training procedure is more computationally efficient than other competing algorithms.

Keywords: Offshore, Pipeline Routing, Reinforcement Learning, Stochastic, Probability of Failure.

1. Introduction

Comprehensive risk assessment for offshore infrastructure development is necessary as billions of dollars of capital is in jeopardy due to its likely impact of natural threats. The development of offshore pipelines is critical in maintaining a consistent supply of oil and gas. This proves to be challenging as advancement in offshore is riddled with uncertainty coming from geological, hydrological and even biological threats (Haneberg 2015), which increase the potential costs due to these likely impacts during construction and operation. Comprehensive planning on pipeline routing considers both optimization on the pipeline installation, and the total distance from one point to interest to another (i.e. origin and destination), among other constraints that are intrinsic to the design and operation of pipeline's systems.

Pipeline routing is typically framed as an optimization problem, where the goal is to determine the most efficient path across a given 'cost' surface between its origin and destination. The cost surface is generally constructed by integrating multiple feature maps into a composite cost map (Haneberg, 2015; Devine et al., 2016; Devine and Haneberg, 2016). Despite the significant uncertainty associated with offshore features, most composite cost maps are deterministic in nature. Moreover, these composite geo-cost maps rely on a singular weighted index to represent features (e.g., slope) and potential hazards. A key limitation of this approach is that the index is not directly linked to a physical model describing submarine landslide vulnerability (or fragility), which is known to be one of the most unpredictable risk components. Consequently, the inherent unpredictability of submarine landslides underscores the importance of incorporating uncertainty into pipeline route optimization.

Most GIS-based software incorporates least-cost path optimization algorithms rooted in Dijkstra's algorithm (ESRI, 2024). While Dijkstra's algorithm is effective for identifying optimal routes between any given two points, it can be computationally intensive for large-scale systems and is not well-suited for simulating multiple realizations of potential optimal routes. In contrast, recent advancements have utilized Reinforcement Learning to address engineering challenges, including optimal control (Diveev et al., 2023), mechanical and structural design (Brown et al., 2022; Yang et al., 2022), and offshore pipeline route optimization (Bhowmik, 2021).

A comprehensive methodology for conducting Bayesian inference on the parameters of a physical model to derive their posterior distributions in a remote offshore setting, was introduced in previous research by Varela et al.(2025) and Das et al. (2019). This approach was later extended to map the probability of failure (PoF) derived from the posterior distributions by Alvarado-Franco (2023). In this study, the joint posterior distribution, representing the PoF of the remote offshore location was used to generate an ensemble of PoF maps and identify optimal pipeline routes for each map using Policy Iteration (Sutton & Barto, 2018). Preliminary results of the pipeline routing are presented in this paper, demonstrating that the proposed approach is computationally more efficient than current least-cost path routing methods.

2. Methodology

2.1. *Joint Posterior Distribution*

Varela et al. (2025), Das et al. (2019), and Alvarado-Franco et al. (2023) integrated probabilistic calibration of the infinite block model using Markov Chain Monte Carlo (MCMC), to prove the use of randomly sampled observations of geotechnical slope stability design parameters and retrieved posterior distribution of each design parameter (i.e., γ_{sat}, γ_{sub}, α, H, k). They showed as well that the posterior distribution of the infinite block model dynamically adjusts depending on the set of observations. Through multiple calibrations with varying observation sets, the probability of failure for a given slope angle can be computed through Monte Carlo sampling.

In this study, soil sediments between depths of 0 to 5 meters areassumed to belong to the same statistical populations, and as a result pipeline optimization does not account for depth as a factor. Although Varela et al. (2025) proposed a probabilistic calibration of the infinite slope model by dividing the soil depth into 1-meter subsets, this approach lies outside the scope of this paper. The specific geographical and geological characteristics of the offshore location remain undisclosed but further discussion can be found in previous research (Varela et al., 2025). Figure 1 highlights the specific point of interest from the entire map used for the probabilistic of the slope model calibration. The specific point was selected due to its topographical features that shows two peaks, a distinguishable valley and part of a linear fault.

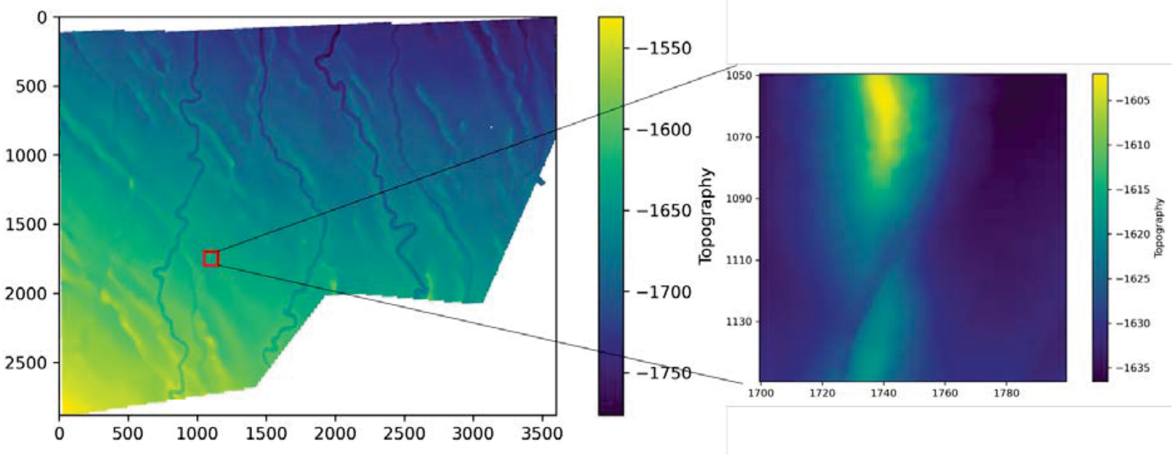
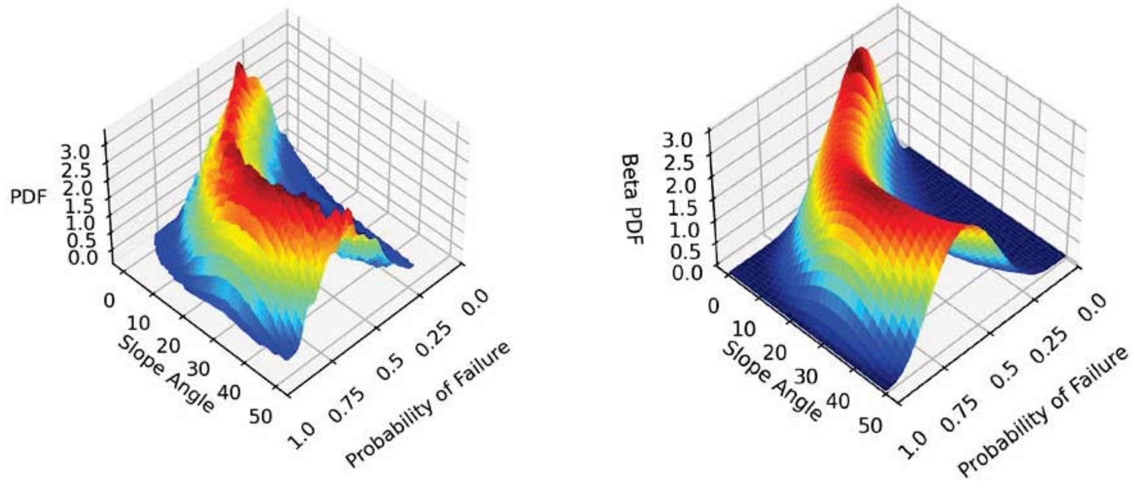


Figure 1. Offshore topography used in this study and the specific area of interest.

Figure 2a depicts the raw joint posterior distribution of a slope angle and its corresponding probability of failure, while Figure 2b presents the associated beta distribution-fitted probability density function (PDF). The distribution's shape reflects the behavior of the infinite slope model: the probability of failure (PoF) increases monotonically up to a threshold slope angle and then decreases as the slope becomes steeper beyond the threshold. Further details and analysis of this phenomenon are provided in prior work by Varela et al. (2025), Das (2019) and Alvarado-Franco (2023).



(a)

(b)

Figure 2. Joint posterior distribution of slope angle and probability of failure retrieved from probabilistic calibration (a) raw surface and (b) smoothened surface.

2.2. *Policy Iteration*

Pipeline routing on a gridded surface can be reformulated as a Markov Decision Process (MDP) problem (Chen et al., 2023). MDP is a mathematical framework to model decision-making under uncertain events,and is defined by the variables States (S), Actions (A), Transition Probability ($P(s^{'}|s,a)$), Reward ($R(s,a)$) and Discount Factor (γ). The overall goal of MDP is to find an optimal policy π^* that maximizes the cumulative expected reward (or minimizes cumulative PoF) and the cumulative reward can be expressed as

$$G_t = \Sigma_{k=0}^{\infty} \gamma^k R_{t+k} \qquad (1)$$

Policy Iteration, pioneered by Howard (1960), seeks for π^*(i.e., the best policy) by alternating between policy evaluation and improvement until convergence to optimality. For a given policy π, the state-value function $V^{\pi}(s)$ satisfies the Bellman equation (Sutton & Barto, 2018) below:

$$V^{\pi}(s) = \Sigma_{s' \in S} P\left(s^{'} \middle| s, \pi(s)\right)\left[R(s, \pi(s), s^{'}) + \gamma V^{\pi}(s^{'})\right] \qquad (2)$$

Eq. 2 is computed during the policy evaluation and can be reduced to linear system of equations to accelerate convergence. The policy iteration algorithm is given below:

Algorithm 1: Policy Iteration

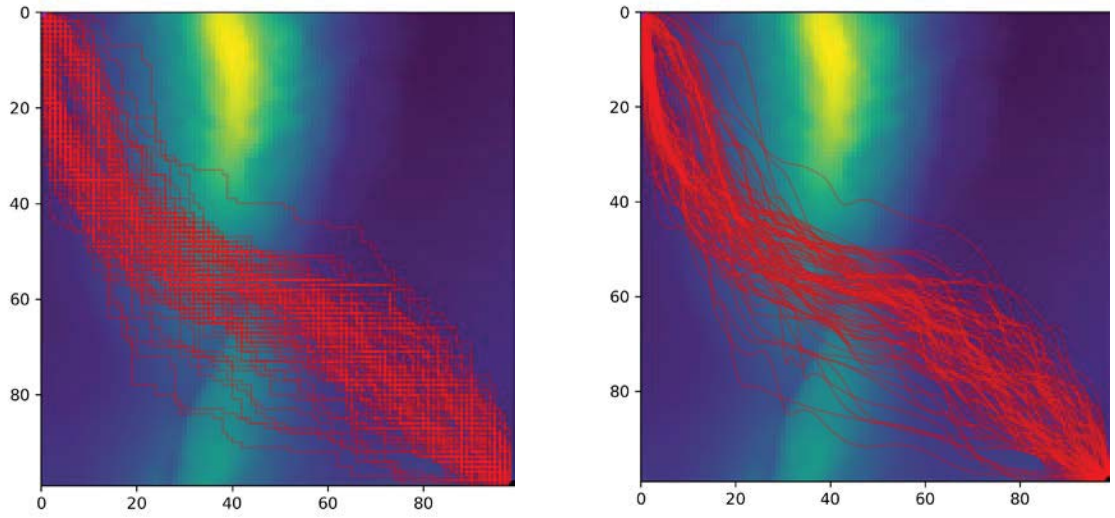
```
For every iteration less than training steps
    Do Policy Evaluation: Solve $(\boldsymbol{I} - \gamma \boldsymbol{P^{\pi}})\boldsymbol{V^{\pi}} = \boldsymbol{R^{\pi}}$ for $\boldsymbol{V^{\pi}}$
    Do Policy Improvement: $\pi(s) \leftarrow \mathrm{argmax}_a \sum_{s' \in S} P(s^{'}|s,a)\,[R(s,a,s^{'}) + \gamma V^{\pi}(s^{'})]$
    If policy stops changing, terminate process
```

In policy iteration, γ plays a critical factor in the importance of long-term future awards. For a large-grid surface, using a high value of γ emphasizes long-term rewards to ensure convergence and reaches the terminal point. For this case study, a value of 0.99 was considered to account for long-term rewards for a large grid problem. For this study, the reward is PoF for each pixel.

3. Results

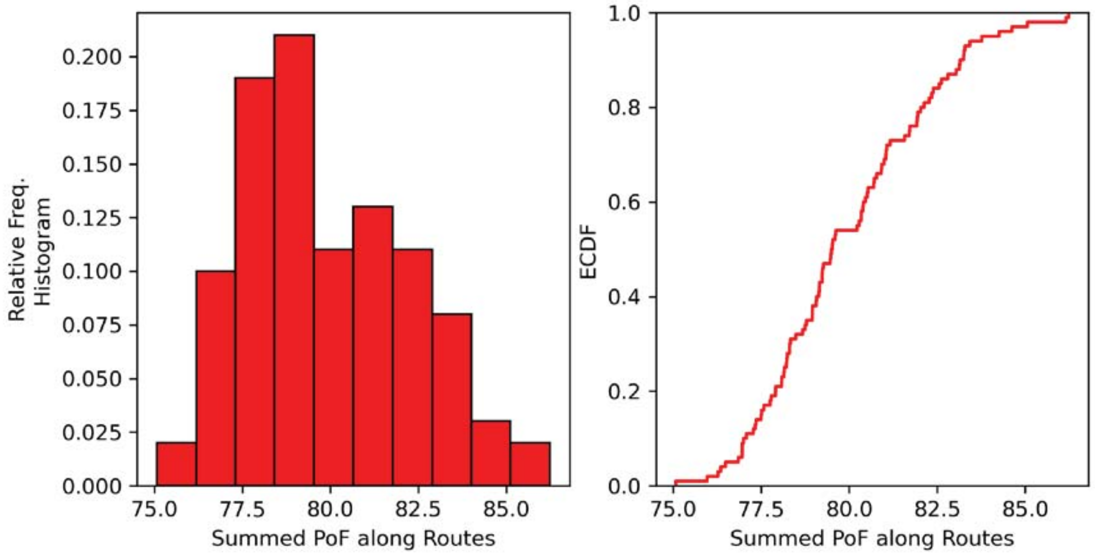
Monte Carlo sampling from the joint posterior distribution was performed to generate multiple realizations of the 'cost surface' or probability of failure maps(PoF), which were subsequently used for policy iteration. The starting and terminal points were set hypothetically located at the northwesternmost and southeasternmost points of the map, respectively. To illustrate the applicability of the proposed methodology, a total of 100 PoF map realizations were sampled, with the raw optimized paths shown in Figure 3a. Figure 3b illustrates the smoothed pipeline routes, accounting for maximum curvature limits imposed by pipe geometry, material constraints, and geological factors. Smoothing was achieved by applying a univariate spline to each pipeline route realization.

The optimization process aimed to minimize the PoF along the shortest path from the starting to the ending point. The distribution of the cumulative PoF along the optimized paths is presented in Figure 4which shows a right-skeweddistribution. This can serve to assess the degree of confidence in risk assessment and decision making when defining the pipeline route. While most of the summed PoF are clustered around lower values (i.e., 76 to 79), the tail on the right of this distributionsuggests presence of a few realizations that lead to significantly higher values of summed PoF.



(a) (b)

Figure 3. Topographical map with (a) raw optimized pipeline route and (b) smoothed pipeline route.



(a) (b)

Figure 4. Summed PoF distribution (a) relative frequency histogram and (b) empirical cumulative distribution function.

4. Conclusion

This study presents preliminary results of a probabilisticreinforcement learning (RL)-based offshore pipeline optimization methodology. For a comprehensive risk assessment in pipeline routing, additional geological, economic, and environmental factors can be integrated into a composite cost map, as outlined in previous works (Hanneberg, 2015; Makrakis et al., 2020). Policy iteration optimization remains applicable for any cost map configuration. Future implementation can integrate other relevant factors to compute a composite reward.

The area of interest was selected to minimize computational costs and facilitate the creation ofmultiple optimized routes.A notable challenge encountered in this study was scaling the grid size for optimization over larger surfaces. Although faster than the least-cost approach, policy iteration faces convergence issues when applied to excessively large grids. Future work will explore alternative Reinforcement Learning techniques to handle larger grids effectively while incorporating feasible pipeline geometry and material constraints to enhance the smoothing criteria (Lucena et al., 2014). Additionally, implementing the Bayesian framework to produce joint posterior distributions for each discrete depth, as previously proposed (Varela et al., 2025), could enable the integration of depth considerations into the pipeline routing optimization process.

References

Alvarado-Franco, J.P. (2023, 12). Bayesian Methods for Landslide Risk Assessment. PhD Thesis, Texas A&M University.

Alvarado-Franco, J.P., B. Hernawan, Z. Medina-Cetina, P. Varela-Gonzalez (2023). Probabilistic Offshore Landslides Hazard Mapping: Integrating Observations, Models and Experts' Knowledge through a Bayesian Framework.*AGU Fall Meeting 2023*, Session: Natural Hazards / Empowering Resilience to Climatic and Environmental Hazards Through Geospatial Data Science II Poster.

Bhowmik, S. (2021). Machine Learning-Based Optimization for Subsea Pipeline Route Design.*Offshore Technology Conference*, OTC-31031-MS.

Brown, N.K., A.P. Garland, G.M. Fadel, G. Li (2022). Deep Reinforcement Learning for Engineering Design through Topology Optimization of Elementally Discretized Design Domains.*Materials & Design*, 218:110672.

Chen, Y.J., B.G. Jhong, M.Y. Chen (2023). A Real-Time Path Planning Algorithm Based on the Markov Decision Process in a Dynamic Environment for Wheeled Mobile Robots. *Actuators*, 12(4).

Das, R., P. Varela-Gonzalez, Z. Medina-Cetina (2019). The Effect of Bayesian Updating in the Hazard Assessment of Submarine Landslides.*Offshore Technology Conference*, OTC-29669-MS.

Devine, C.A. and W.C. Haneberg (2016). Optimization Methods for Arctic Pipeline Route Selection.*Offshore Technology Conference*, OTC-27391-MS.

Devine, C.A., W.C. Haneberg, H. Lee, M.-L. Liu, G.-A. Chang. (2016). A Sensible Approach to Subsea Pipeline Route Determination – Moving from Hand-Drawn Routes to Geologically-Constrained, Least-Cost Optimized Paths.*Offshore Technology Conference*, OTC-26940-MS.

Diveev, A., E. Sofronova, S. Konstantinov, V. Moiseenko (2023). Reinforcement Learning for Solving Control Problems in Robotics.*Engineering Proceedings*, 33:1.

ESRI (2024). Algorithms used by Network Analyst. ArcGIS Pro 3.4 Documentation.

Haneberg, W.C. (2015). Stochastic Incorporation of Uncertainty and Subjectivity in Deepwater Pipeline Route Optimization. *Offshore Technology Conference*, OTC-25785-MS.

Howard, R.A. (1960). *Dynamic Programming and Markov Processes*. MIT Press.

Lucena, R.R.d., J.S. Baioco, B.S.L.P.d. Lima, C.H. Albrecht, B.P. Jacob (2014). Optimal Design of Submarine Pipeline Routes by Genetic Algorithm with Different Constraint Handling Techniques.*Advances in Engineering Software*, 110-124.

Makrakis, N., P.N. Psarropoulos, D. Chatzidakis, Y. Tsompanakis(2020). Route Optimization of Offshore Lifelines taking into Account Potential Earthquake-Related Geohazards.*Frontiers in Built Environment*, 6:112.

Sutton, R.S. and Barto, A.G. (2018).*Reinforcement Learning: An Introduction*. MIT Press.

Varela, P., Medina-Cetina, Z., Hernawan, B. (2025). Bayesian Model Calibration of Submarine Landslides. Landslides.https://doi.org/10.1007/s10346-025-02486-y

Yang, J., Z. Cheng, G. Xiao, X. Xu, Y. Wang, H. Ding, D. Zhou (2022). Engineering Design Optimization using Reinforcement Learning with Episodic Controllers.Cognitive Computation and Systems.