Geosteering Through the Lens of Decision Transformers: Toward Embodied Sequence Decision-Making

Hibat Errahmen Djecta[1,2], Sergey Alyaev[1], Kristian Fossum[1], Reidar B.Bratvold[2], and Dan Sui[2]

1. NORCE Research AS, Nygårdsgaten 112, 5008 Bergen, Norway

2. University of Stavanger, Kjell Arholms gate 41, 4021 Stavanger, Norway

1 Introduction

Geosteering requires real-time decisions under noisy, incomplete subsurface data. From an embodied world-model perspective, decisions shape the observations an agent receives in an uncertain environment. Traditional RL struggles with stability and long-horizon reasoning [2], so we frame geosteering as a sequence modeling problem [3,4] using a Decision Transformer (DT) [8].

Key challenges:

- Noisy and uncertain geology
- Irreversible steering decisions
- RL instability and short-horizon bias

Geosteering is challenging due to noisy measurements and hidden geology, and effective decision-making requires awareness of long-term trends. Decision Transformers capture these dependencies and offer a stable, offline alternative to Deep Reinforcement Learning (DRL) [6,7].

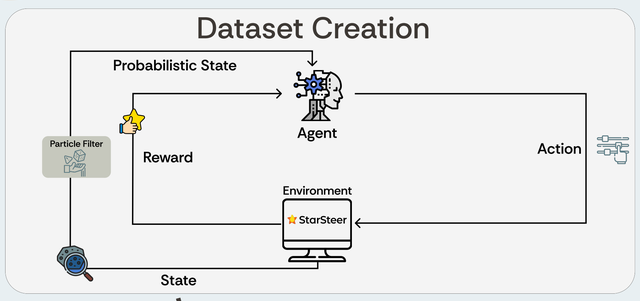
2 Formulation

In geosteering, the agent operates with uncertain and partially observed geological information, receiving noisy sensor signals while its decisions influence future measurements and well placement. This interaction naturally fits a Partially Observable Markov Decision Process (POMDP), which captures both hidden subsurface states and their evolution under action. This structure is consistent with embodied world models, where an agent must infer latent dynamics and act based on how its decisions shape future observations.

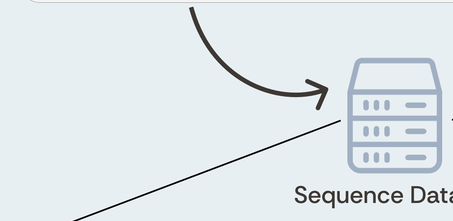
$$\mathcal{P} = (S, A, T, R, \Omega, O, \gamma)$$

States S	Actions A	Transition Model T	Reward R	Observation Ω	Observation Model O	Discount Factor γ
Reservoir top/ bottom boundaries and formation properties.	Steering commands issued during drilling	Determines how the well position evolves after an action	Measures drilling quality	Sensor readings such as gamma-ray logs	Maps latent geological states to noisy observations produced by sensors.	Encodes the importance of future rewards and long-horizon steering outcomes.

3 Approach



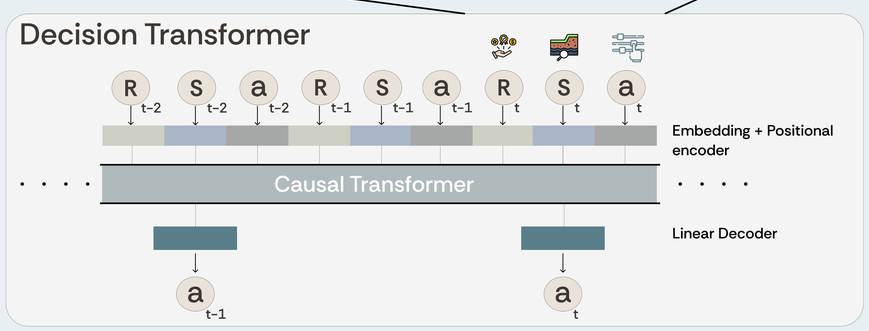
We start by creating the dataset for the Decision Transformer by collecting probabilistic state estimates from particle filters and steering actions (Δinclination, Δazimuth) generated by our previous DRL agent [1].

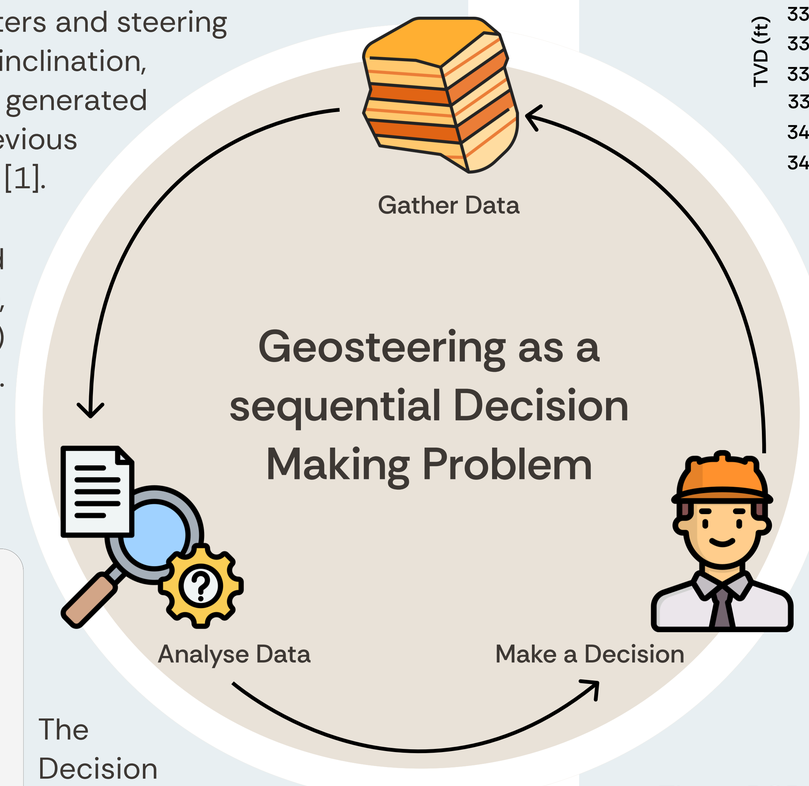


Each episode is converted into a sequence of (state, action, return-to-go) triplets.

Sequence Dataset

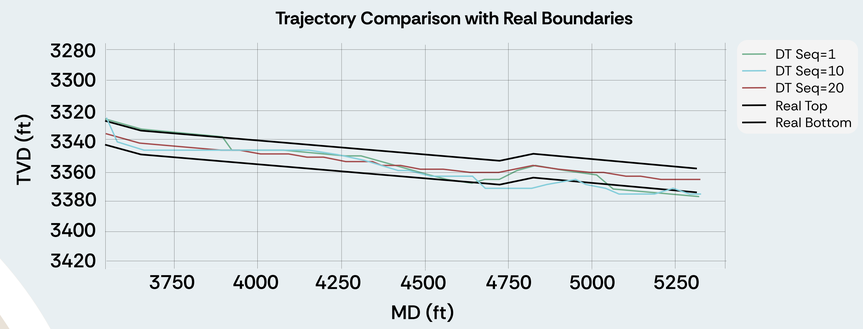
state_t	action_t	return_to_go_t	state_t+1	done



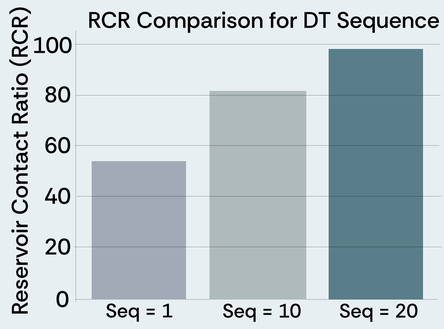


The Decision Transformer learns long horizon steering behavior over sequences.

4 Results



The trajectory comparison shows that the Decision Transformer with a short sequence length produces unstable, drifting paths, while a longer context allows it to follow the reservoir structure much more accurately.

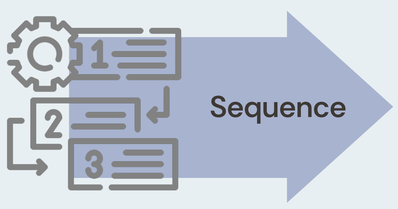


The RCR results confirm this improvement, with the Decision Transformer achieving significantly higher reservoir contact when trained with longer sequence lengths.

5 Discussion

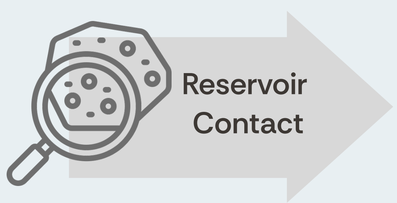
Short sequence lengths make the model react to noise and drift from the reservoir, while longer context allows the Decision Transformer to follow geological trends more consistently. The RCR results confirm that longer context improves reservoir contact and overall decision reliability. This behavior reflects the role of embodied world models, where reliable control depends on understanding how decisions influence future observations.

Key Insights



Temporal context drives reliability:

Longer sequences enable the DT to stabilize its decisions and align more closely with geological structures, whereas short sequences behave myopically.



Better reservoir contact:

Higher RCR values show that long-context models not only track boundaries more accurately but also maintain productive placement throughout the trajectory.

6 Conclusion

Decision Transformers provide a stable, offline alternative to RL for geosteering, learning long-horizon behaviors that better track geological structures and improve reservoir contact.

Key takeaways:

- Longer context windows yield more consistent and aligned trajectories
- DT captures delayed geological signals missed by short-horizon models
- Offline sequence modeling avoids RL instability in uncertain environments

References

1. Djecta, H. E., Alyaev, S., Fossum, K., B. Bratvold, R., Muhammad, R. B., & Srivastava, A. (2025, July). Uncertainty-Aware Well Placement: Simulator-Verified Dual-Network Reinforcement Learning Approach Meets Particle Filters. In International Conference on Computational Science (pp. 188-202). Cham: Springer Nature Switzerland.
2. Alyaev, S., Fossum, K., Djecta, H., Tveranger, J., Elsheikh, A. Distinguish workflow: a new paradigm of dynamic well placement using generative machine learning. In: ECMOR 2024. European Association of Geoscientists & Engineers (2024).
3. Kullawan, K., Bratvold, R., Bickel, J. A decision analytic approach to geosteering operations. SPE Drilling & Completion, 29 (2014).
4. Kullawan, K., Bratvold, R., Bickel, J. Sequential geosteering decisions for optimization of real-time well placement. Journal of Petroleum Science and Engineering, 165, 90–104 (2018).
5. Mnih, V., Kavukcuoglu, K., Silver, D., et al. Human-level control through deep reinforcement learning. Nature, 518, 529–533 (2015).
6. Muhammad, R.B., Cheraghi, Y., Alyaev, S., Srivastava, A., Bratvold, R.B. Geosteering robot powered by multiple probabilistic interpretation and artificial intelligence: Benchmarking against human experts. SPE Journal, 2025.
7. Muhammad, R.B., Alyaev, S., Bratvold, R.B. Optimal sequential decision-making in geosteering: A reinforcement learning approach. (2025).
8. Chen, L., Lu, K., Rajeswaran, A., Lee, K., Grover, A., Laskin, M., ... & Mordatch, I. (2021). Decision transformer: Reinforcement learning via sequence modeling. Advances in neural information processing systems, 34, 15084-15097.

Scan Me For Details