Subterra: An Open-Source Validated Benchmark for GNSS-Denied Tunnel Navigation

Abstract—**Underground and subway tunnels are among the hardest settings for autonomous navigation: they are Global Navigation Satellite System (GNSS) denied because no satellite signal penetrates the surrounding rock, and geometrically degenerate—long, featureless, near-uniform cross-sections offer almost no landmarks, so onboard state estimation drifts along the tunnel axis. We present Subterra, an open-source benchmark for GNSS-denied tunnel navigation whose difficulty is inherited rather than designed: its geometry is reconstructed directly from real LiDAR in the GEODE dataset and released as a MuJoCo environment with a Gymnasium-compatible POMDP interface. The reconstructed bore is a $351\,\mathrm{m}$ swept tunnel; observations come from a noisy planar LiDAR and a drifting odometry model, and a Unitree B2 quadruped is provided as a locomotion embodiment. A reproducible 18-category, 199-metric suite scores the reconstruction at a global $\mathbf{97.1\%}$ fidelity, with a $\mathbf{2.46\,cm}$ synthetic-LiDAR surface error against the real reference. As a usability check, a PPO agent trained end-to-end in Subterra reaches a mean rolling coverage of $\mathbf{90.02\%} \pm 0.01\%$ on all five seeds, confirming that the benchmark is hard yet learnable.**

I. Introduction

Autonomous robots deployed in tunnels and mines face a compounding perception failure: no satellite signal penetrates the surrounding rock, and long near-uniform cross-sections offer no landmarks for state estimation, so even LiDAR odometry drifts irreversibly along the tunnel axis. Navigation therefore reduces to a partially observable Markov decision process (POMDP) [1], the regime in which GNSS-denied odometry drift is the dominant failure mode [8].

Progress is bottlenecked by *evaluation*: synthetic benchmarks author their own difficulty, which a reviewer can dispute. **Subterra** reconstructs its geometry from real GEODE LiDAR [2], exports it to MuJoCo [6] behind a Gymnasium interface [7], and pairs it with a noisy-LiDAR-plus-drifting-odometry POMDP. Our contributions are:

- **C1** Subterra, an open-source GNSS-denied tunnel-navigation POMDP reconstructed from real GEODE LiDAR and released as a MuJoCo scene and a Gymnasium environment—furnished from the data with railway track structure and driven by a noisy-LiDAR-plus-drifting-odometry sensor model.
- **C2** A reproducible suite of 18 categories / 199 logged metrics that establishes that the reconstruction faithfully reproduces the real bore (global 97.1%; $2.46\,\mathrm{cm}$ surface error).

Reinforcement learning is *not* a contribution of this paper: we train a PPO agent only as a usability check that standard algorithms can learn to navigate the released environment (Sec. IV).

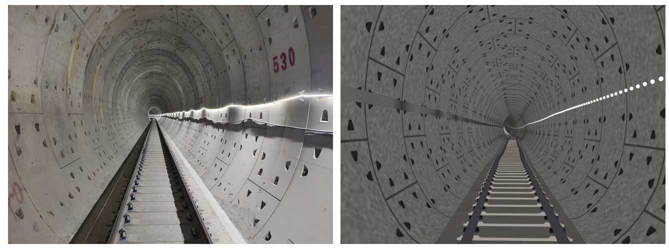


Fig. 1: **Subterra: real vs. reconstructed.** Real GEODE tunnel frame (left) and MuJoCo bore environment (right).

II. The Subterra Environment

Real2Sim reconstruction. We use an underground tunnel sequence from GEODE [2], a real Livox BMI1088 LiDAR, MTi-30 IMU and stereo dataset built to stress SLAM in geometrically degenerate environments. From the accumulated point cloud data Subterra extracts a $351\,\mathrm{m}$ centerline and sweeps a D-section bore ($\approx 5.65 \times 6.36\,\mathrm{m}$) along it, exported as a MuJoCo scene [6] (Fig. 1).

Physical tunnel structure. The bore is lined like a real bored tunnel: featureless segmental concrete walls that offer no axial landmark. The wall pattern is a grid of recessed holes $\approx 13\,\mathrm{cm}$ deep, each joined to the next cell by a connecting nail—the segment-jointing detail visible in the source environment. Subterra also reconstructs 586 concrete monoblock sleepers at $0.6\,\mathrm{m}$ spacing, paired rails at the standard $1.435\,\mathrm{m}$ gauge on raised pads, Pandrol-style fasteners at each rail seat. The gauge is fixed by the track standard and each sleeper's fit is validated against the local bore half-width, so every dimension is set by the track standard or the LiDAR bore rather than hand-tuned.

POMDP formulation. The hidden state is the robot's pose along a deep, GNSS-denied tunnel; the agent acts with continuous velocity commands $a = [v, \omega]$, $v \in [0, 2]\,\mathrm{m/s}$, $\omega \in [-1.5, 1.5]\,\mathrm{rad/s}$. A 16-beam planar LiDAR ($30\,\mathrm{m}$ range, $\pm 5\,\mathrm{cm}$ Gaussian noise) pins lateral offset and heading but is *ambiguous along the axis*. Proprioceptive odometry carries per-step noise *and* a per-episode systematic bias, so dead-reckoning *drifts*—the core failure mode of GNSS-denied operation. The Gymnasium reset/step interface exposes an 18-D observation (LiDAR + drifting odometry, no absolute pose), and a Unitree B2 quadruped is provided as a 3-D locomotion embodiment. **Reward.**[4] At each step the agent receives

$$r_t = \underbrace{r_{\text{term}}}_{\text{sparse terminal}} + \underbrace{g\,\frac{\Delta s_t}{L}}_{\text{coverage}} - \underbrace{c_d\,\frac{|d_t|}{w_t}}_{\text{lateral offset}} - \underbrace{c_s\big(|\Delta v_t| + |\Delta\omega_t|\big)}_{\text{control jerk}}, \tag{1}$$

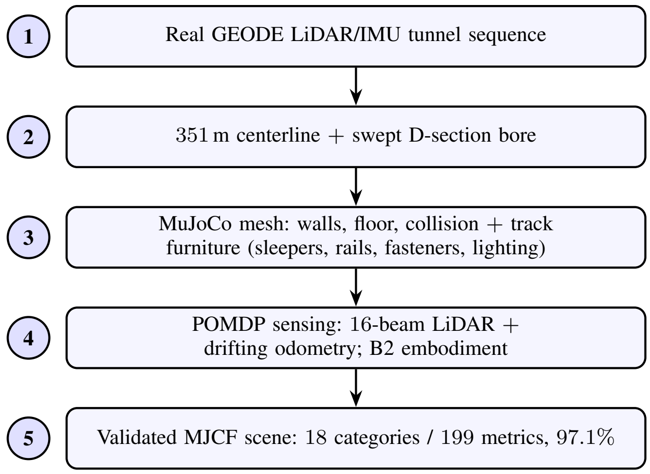


Fig. 2: Subterra construction pipeline: from the raw GEODE recording to the released, validated MJCF scene.

where r_{term} is a sparse terminal reward of $+1$ for success (reaching 90% tunnel coverage without wall contact) or 0 otherwise, granted once at episode end; $\frac{\Delta s_t}{L}$ is the coverage reward, where Δs_t is the newly covered arc-length beyond the furthest point reached so far (a ratchet mechanism preventing reward for re-covering ground) and L is the total tunnel length, scaled by gain g; $\frac{|d_t|}{w_t}$ is the lateral offset penalty, where d_t is the deviation from the centerline and w_t is the local half-width, penalizing off-centre travel with weight c_d; and $|\Delta v_t| + |\Delta \omega_t|$ is the control jerk penalty penalizing changes in linear and angular velocity with weight c_s.

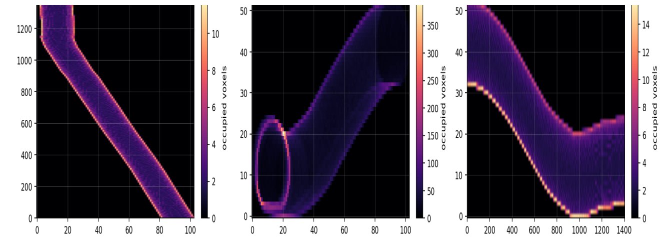
III. Sim-vs-Real Fidelity Validation

Subterra is validated quantitatively with a reproducible suite that scores the reconstruction across 18 independent categories—spanning geometry, mesh integrity, topology, 2-D/3-D free space, navigability, point-cloud agreement, occupancy, and scene assets—emitting 199 logged metrics; each category reduces to a percentage and the global score is their unweighted mean. The full suite is reproducible from a single config (MuJoCo 3.8.1, fixed RNG seed) and is shown insensitive to a $\pm 10\%$ threshold sweep and to the seed, scoring a global $\mathbf{97.1}\%$.

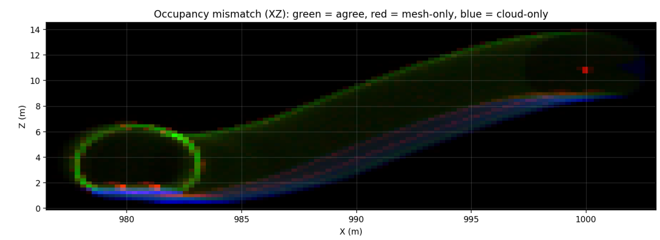
The headline geometric result is an in-frame synthetic-LiDAR cloud that agrees with the mesh to a mean point-to-surface error of $\mathbf{2.46}$ **cm** (97.5% of points within 10 cm). Figure 3 shows the spatial fidelity directly: (a) the reconstructed bore's voxel occupancy and (b) a voxel-by-voxel occupancy comparison against the real LiDAR cloud, in which agreement (green) dominates and the few disagreements lie on thin boundary voxels and the open bore mouth, not the tunnel structure.

IV. Learnability Check

We train a PPO agent [5] (Stable-Baselines3 [3], frame-stacked MLP policy) end-to-end on the partial observation, with no access to absolute pose. The task is hard for an authentic reason: the bore is reconstructed from real degenerate geometry, and the agent must close the loop through a noisy LiDAR and drifting odometry rather than a tuned difficulty knob. We run five independent seeds (0–4); coverage is reported as the mean±std across seeds (Fig. 4). All five seeds converge to the 90% coverage goal (mean rolling-100 coverage $90.02\% \pm 0.01\%$, 100% success at convergence), but only after ~ 1000 episodes of exploration—hard yet learnable, exactly the regime a useful benchmark should occupy.



(a) Bore-shell voxel occupancy of the reconstruction, summed along each axis (X–Y, X–Z, Y–Z).



(b) Occupancy mismatch (X–Z): green=mesh and cloud agree, red=mesh-only, blue=cloud-only.

Fig. 3: Sim-vs-real spatial fidelity.

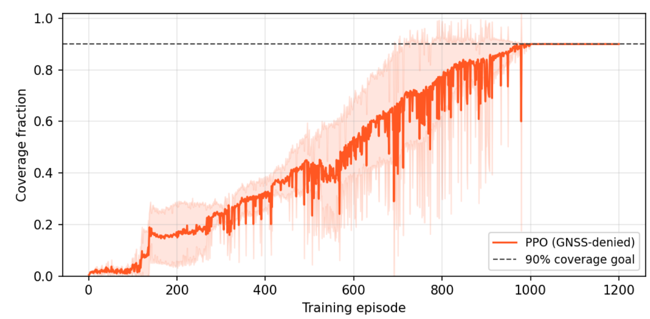


Fig. 4: Per-episode tunnel coverage. All seeds climb to and hold the 90% coverage goal (dashed), confirming the environment is learnable under partial observability.

V. Conclusion

We presented **Subterra**, an open-source benchmark for GNSS-denied tunnel navigation reconstructed from real GEODE LiDAR. We validated the reconstruction to a global 97.1% fidelity across 18 categories—a 2.46 cm surface error against the real reference—and showed that a PPO agent learns to navigate it under partial observability, reaching the 90% coverage goal on all five seeds (mean rolling coverage $90.02\% \pm 0.01\%$). Subterra, its construction pipeline, and the full validation suite are released **open-source** as a faithful, ready-to-use testbed for navigation under partial observability.

References

[1] Leslie Pack Kaelbling, Michael L. Littman, and Anthony R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101 (1–2):99–134, 1998.

[2] PengYu-Team. GEODE: A geometrically degenerate multi-sensor dataset for SLAM in challenging environments. https://github.com/PengYu-Team/GEODE_dataset, 2024. Real-world LiDAR/IMU/stereo dataset, 64 trajectories over 64 km across 7 environment types.

[3] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. Stable-baselines3: Reliable reinforcement learning implementations. *Journal of Machine Learning Research*, 22(268): 1–8, 2021.

[4] Abhishek S, Badrikanath Praharaj, and Sreeram MV. Min-explore: An open-source reinforcement learning exploration benchmark for gnss-denied underground environment. In *IEEE International Conference on Robotics and Automation (ICRA) Xplore: Cross-Disciplinary Aspects of Exploration in Robotics, Reinforcement Learning and Search*, 2026. doi: 10.48550/arXiv.2606.04569.

[5] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[6] Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5026–5033, 2012.

[7] Mark Towers et al. Gymnasium: A standard interface for reinforcement learning environments. https://github.com/Farama-Foundation/Gymnasium, 2024.

[8] Ji Zhang, Michael Kaess, and Sanjiv Singh. On degeneracy of optimization-based state estimation problems. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 809–816, 2016.