

Indoor Benchmark of 3D LiDAR SLAM at iilab – Industry and Innovation Laboratory

Journal Title
XX(X):1–20
©The Author(s) 2025
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/



JORGE DIOGO RIBEIRO^{1,2}, RICARDO B. SOUSA^{1,2}, JOÃO G. MARTINS^{1,2}, ANDRÉ S. AGUIAR², FILIPE N. SANTOS², and HÉBER MIGUEL SOBREIRA²

Abstract

This paper presents the IILABS 3D dataset, a novel and publicly available resource designed to address current limitations in indoor benchmarking of 3D LiDAR-based Simultaneous Localization and Mapping (SLAM) algorithms. Existing SLAM datasets often focus on outdoor environments, rely on a single type of LiDAR sensor, or lack ground-truth data suitable for evaluating diverse indoor conditions. IILABS 3D fills this gap by providing a sensor-rich, indoor-exclusive dataset recorded in a controlled laboratory environment using a wheeled mobile robot platform. It includes four heterogeneous 3D LiDAR sensors – Velodyne VLP-16, Ouster OS1-64, RoboSense RS-Helios-5515, and Livox Mid-360 – featuring both mechanical spinning and non-repetitive scanning patterns, as well as an IMU and wheel odometry. The dataset also features calibration sequences, challenging benchmark trajectories, and high-precision ground-truth poses captured with a Motion Capture (MoCap) system. By combining diverse sensor technologies, extensive calibration data, and carefully designed indoor scenarios, IILABS 3D enables more comprehensive and reproducible evaluation of LiDAR-based SLAM algorithms, fostering innovation in autonomous navigation within complex indoor environments. The dataset information and associated tools are available on the project webpage: jorgedfr.github.io/3d_lidar_slam_benchmark_at_iilab.

Keywords

Dataset, Ground Mobile Robot, Indoor Environment, Light Detection And Ranging (LiDAR), Simultaneous Localization and Mapping (SLAM).

1 Introduction

Due to their advantageous properties, mobile robots, autonomous vehicles, and unmanned aerial vehicles frequently employ Light Detection And Ranging (LiDAR) sensors (Bresson et al. 2017; Nguyen et al. 2022). LiDAR provides accurate ranging data across a wide Field of View (FoV), maintains performance under varying lighting conditions, and requires relatively low computational resources to generate 3D dense point clouds (Roriz et al. 2022). While vision and radar sensors offer valuable capabilities and can be critical for specific tasks, integrating LiDAR with complementary technologies, such as Inertial Measurement Unit (IMU)s and wheel odometry, can enhance overall system robustness and reliability (Shan et al. 2020; Guadagnino et al. 2024).

Recent advances in LiDAR-based Simultaneous Localization and Mapping (SLAM) have driven significant progress in robotics and autonomous navigation (Siegwart et al. 2011; Durrant-Whyte and Bailey 2006; Bailey and Durrant-Whyte 2006). Nevertheless, many real-world indoor applications still face challenges in achieving robust, accurate localization and mapping. Existing datasets and benchmarking frameworks are often limited by their narrow focus on outdoor scenarios or by their reliance on a single LiDAR sensor, which restricts the exploration of sensor performance under diverse indoor conditions (Cadena et al. 2016). This gap is especially critical in industrial and service robotics, where

the complex geometries and dynamic environments of indoor spaces demand high-precision solutions (Xu et al. 2022; Sousa et al. 2023).

In response, we introduce the IILABS 3D dataset – iilab Indoor LiDAR-based SLAM 3D. This dataset is a novel, publicly available resource designed to address the need for datasets specially focused on the benchmark of LiDAR-based SLAM algorithms in indoor environments. This dataset was captured using a mobile robot at the iilab – Industry and Innovation Laboratory – and features multiple sensor sources, including four distinct 3D LiDAR sensors (Velodyne VLP-16, Ouster OS1-64, RoboSense RS-Helios-5515, and Livox Mid-360) with varying FoVs and scanning patterns, an IMU (Xsens MTi-630 AHRS), and wheel odometry. Furthermore, it comprises both calibration sequences and challenging benchmark trajectories, which makes it a good testbed for evaluating and comparing state-of-the-art SLAM algorithms in indoor environments.

¹ Faculty of Engineering, University of Porto (FEUP), 4200-465 Porto, Portugal

² INESC TEC – Institute for Systems and Computer Engineering, Technology and Science, 4200-465 Porto, Portugal

Corresponding author:

Jorge Diogo Ribeiro, Faculty of Engineering, University of Porto (FEUP), 4200-465 Porto, Portugal.

Email: jorge.d.ribeiro@inesctec.pt

In parallel, we present a detailed benchmark analysis of nine 3D LiDAR-based SLAM algorithms (LOAM (Zhang and Singh 2014), LeGO-LOAM (Shan and Englot 2018), LIO-SAM (Shan et al. 2020), DLIO (Chen et al. 2023), VineSLAM (Aguiar et al. 2022), KISS-ICP (Vizzo et al. 2023), GLIM (Koide et al. 2024), Kinematic-ICP (Guadagnino et al. 2024), and MOLA-LO (Blanco-Claraco 2025)) evaluated over six sequences and four distinct LiDAR configurations. In our analysis, we assess the accuracy of the odometry trajectories produced by the SLAM algorithms by comparing them against the ground-truth data from a Motion Capture (MoCap) system. We employ the Absolute Trajectory Error (ATE), Relative Translational Error (RTE), and Relative Rotational Error (RRE) metrics for a comprehensive evaluation. Our findings reveal that algorithms integrating IMU data consistently achieve superior accuracy on the IILABS 3D dataset¹, while those relying solely on wheel odometry tend to exhibit more significant errors. This highlights a potential area for future research focused on improving wheel odometry integration in indoor SLAM systems.

Therefore, the principal contributions of this work are:

- **IILABS 3D Dataset:** We introduce a novel indoor LiDAR-based SLAM dataset featuring four diverse 3D LiDAR sensors, an IMU, and wheel odometry, complemented by calibration sequences and challenging benchmark trajectories, along with ground-truth data acquired using a MoCap system;
- **Comprehensive Benchmark Analysis:** We provide a evaluation of nine state-of-the-art 3D LiDAR-based SLAM algorithms over multiple sequences and sensor configurations, offering detailed insights into their accuracy under diverse indoor conditions;
- **Open-Source Resources:** We release all benchmark scripts², Dockerfiles, Robot Operating System (ROS) packages, and the iilabs3d-toolkit³ for dataset download and evaluation metric computation, thereby promoting reproducibility and promoting further research in indoor LiDAR-based SLAM.

The rest of this article is organized as follows. Sec. 2 reviews previous works closely related to the presented contributions. Next, Sec. 3 introduces the proposed dataset, detailing its structure and key characteristics. A comprehensive benchmark analysis is presented in Sec. 4, evaluating nine different 3D LiDAR-based SLAM algorithms using the IILABS 3D dataset. Finally, Sec. 5 summarizes the findings and discusses potential directions for future research.

2 Related Work

This paper presents an indoor LiDAR-based SLAM dataset. As a result, the related work discussion focuses on LiDAR-based SLAM approaches, benchmarks, and datasets. First, we review state-of-the-art 2D and 3D LiDAR-based SLAM algorithms. Next, we explore existing works and frameworks designed for benchmarking SLAM algorithms. Finally, we provide an overview of indoor SLAM datasets, highlighting their key characteristics and contributions. An exhaustive review of LiDAR-based SLAM algorithms is beyond the scope of this work. For a comprehensive analysis of the

fundamentals of SLAM, the reader should refer to the works of Durrant-Whyte and Bailey (2006); Bailey and Durrant-Whyte (2006) and Grisetti et al. (2010). For existing surveys on SLAM algorithms, readers may refer to Cadena et al. (2016), Bresson et al. (2017), and Xu et al. (2022).

2.1 LiDAR-based SLAM algorithms

Typically, SLAM generates a static map, which can be used subsequently for localization. In contrast, long-term SLAM involves a dynamic update of the environment representation while accounting for changes over time (Sousa et al. 2023). Despite advancements in lifelong SLAM, many industrial applications still rely on just creating static maps due to the robustness and simplicity of static mapping solutions in environments where frequent updates are unnecessary (Roriz et al. 2022; Cadena et al. 2016).

Early contributions for 2D LiDAR SLAM, such as GMapping (Grisetti et al. 2007) and HectorSLAM (Kohlbrecher et al. 2011), generated static 2D occupancy grid maps. GMapping (Grisetti et al. 2007) introduced the Rao-Blackwellized particle filter, where each particle has its own map and robot pose. In contrast, HectorSLAM (Kohlbrecher et al. 2011) relied singularly on scan-matching with scan-to-map alignment but without loop closure. Subsequent advancements in 2D SLAM shifted towards pose-graph implementations, improving scalability and facilitating loop closure detection. Cartographer (Hess et al. 2016) introduced submaps represented as probability grids within a pose graph framework, with scan matcher-based loop closure. Later, SLAM Toolbox (Macenski and Jambrecic 2021) used sparse bundle adjustment to optimize scan-based pose graphs. An experimental feature implements lifelong mapping by removing extraneous or outdated information, bounding the computational resources over extended mapping periods.

While 2D SLAM remains popular in industrial applications for its simplicity and robustness, 3D LiDAR data provides richer spatial information than 2D lasers, improving dynamic object tracking, obstacle detection, and non-planar surface mapping. As a result, 3D LiDAR-based SLAM is being increasingly adopted for scenarios where 2D approaches are inadequate for dynamic and non-structured environments (Roriz et al. 2022).

One of the earliest algorithms to achieve reliable performance with feature-based 3D LiDAR SLAM was LOAM (Zhang and Singh 2014), which divided point clouds into two subsets – edge (corner) points and planar features –, enhancing both efficiency and accuracy in scan matching. Subsequent variations, such as LeGO-LOAM (Shan and Englot 2018) and F-LOAM (Wang et al. 2021), built upon LOAM to optimize performance. LeGO-LOAM introduced ground segmentation, improving computational efficiency and reducing noise on real-time applications for ground-based mobile robots. F-LOAM enhanced robustness in dynamic environments through refined feature extraction and motion compensation. As for LOAM-Livox (Lin and Zhang 2020), the authors adapted the LOAM algorithm to accommodate non-uniform and non-repetitive scanning patterns (e.g., Livox sensors) that originate in less structured point clouds.

Significant advancements in 3D SLAM parametrization were introduced by KISS-ICP (Vizzo et al. 2023). The algorithm automatically adjusted parameters to accommodate environment changes or motion profile, reducing manual tuning. Kinematic-ICP (Guadagnino et al. 2024) built on KISS-ICP for wheeled mobile robots on planar surfaces, using wheel odometry for scan de-skewing and integrating kinematic constraints into a point-to-point Iterative Closest Point (ICP) optimization scheme (Besl and McKay 1992). Additionally, MOLA-LO (Blanco-Claraco 2025) introduced a modular system for localization and mapping, further demonstrating the versatility of adaptive parametrization in SLAM systems.

An alternative to wheeled odometry for correcting the skewed point clouds – due to the rotational mechanism of conventional LiDAR sensors and platform motion during data acquisition – is using IMU sensors. Loosely coupled sensor fusion methods, such as those employed in LOAM (Zhang and Singh 2014) and LeGO-LOAM (Shan and Englot 2018), incorporate IMU data to provide motion priors for scan de-skewing but do not integrate the IMU measurements into the optimization process, limiting their accuracy. In contrast, recent approaches, such as LIO-SAM (Shan et al. 2020), DLIO (Chen et al. 2023) and FAST-LIO (Xu and Zhang 2021), implement tight coupling of LiDAR and IMU measurements into the factor graph optimization process, improving the accuracy of pose estimation and map consistency.

Overall, the SLAM field still continues to be actively researched with further contributions. Recent works by Aguiar et al. (2022), Koide et al. (2024), Ferrari et al. (2024), and Pan et al. (2024) introduced innovative techniques for SLAM. VineSLAM (Aguiar et al. 2022) employed a novel particle filter to integrate point and semi-plane features extracted from 3D LiDAR data, demonstrating its effectiveness in localizing robots in agricultural vineyards. GLIM (Koide et al. 2024) is built upon the HDL Graph SLAM (Koide et al. 2019) framework to incorporate GPU acceleration for improving performance in large-scale environments. Furthermore, MAD-ICP (Ferrari et al. 2024) proposed a dynamic uncertainty-driven model update alongside efficient kd-tree data structures, enabling robust and adaptative LiDAR odometry. As for PIN-SLAM (Pan et al. 2024), the authors employed a point-based implicit neural representation to maintain global map consistency, efficiently processing large-scale 3D point cloud data.

2.2 Benchmarking Frameworks

Benchmarks have been widely used in SLAM research to ensure fair comparisons between algorithms by providing normalized testing conditions. Although benchmarks are commonly used in SLAM, some studies rely on proprietary datasets, selectively choosing sequences from well-known public datasets, such as KITTI (Geiger et al. 2013), or comparing only to a limited subset of state-of-the-art algorithms. These evaluation practices may introduce bias in the results, either through fine-tuned parameters that favor the proposed algorithm or by selecting dataset sequences that highlight its strengths. Consequently, there is a growing need for benchmarks that promote more transparent and reproducible evaluations.

Tab. 1 presents open-source benchmarking frameworks. Frameworks that provide their own dataset, including TUM RGB-D (Sturm et al. 2012), KITTI Vision (Geiger et al. 2012), ETH3D (Schöps et al. 2017), KITTI 360 (Liao et al. 2023), and VBR (Brizi et al. 2024), offer multiple sequences to test algorithms under different conditions. The datasets are often divided into training and test sequences. Training data includes accessible ground truth for algorithm tuning and preliminary evaluation, while testing sequences assess the algorithm with non-public ground-truth data, allowing independent and unbiased evaluation by the framework. Furthermore, benchmarking frameworks typically offer a Command Line Interface (CLI) toolkit for researchers to download dataset sequences and compute performance metrics. Two primary metrics are widely used in SLAM benchmarking: ATE in meters and Relative Pose Error (RPE) as a percentage. ATE quantifies the overall global alignment between the estimated and ground-truth trajectories, while RPE assesses local consistency between consecutive poses. Some frameworks further subdivide RPE into RTE and RRE, which measure translational and rotational discrepancies, respectively. Lastly, some frameworks maintain web-based ranking systems, allowing researchers to publicly submit results based on standardized metrics.

In contrast to only supporting a single dataset, frameworks such as SLAMBench (Nardi et al. 2015; Bodin et al. 2018; Bujanca et al. 2019) and SLAM HIVE (Yang et al. 2023; Liu et al. 2024) support several publicly available datasets, enabling cross-domain comparisons. These frameworks are capable of evaluating various SLAM algorithms on multiple datasets, while also monitoring additional performance metrics such as CPU and RAM usage. SLAMBench supports more than 16 Visual SLAM algorithms and 9 datasets. In contrast, SLAM HIVE currently supports 8 SLAM algorithms and sequences from 3 datasets. Nevertheless, its modular design allows for the integration of additional algorithms and dataset sequences. Additionally, Liu et al. (2024) introduced a Docker-based framework to enable graphical access to the benchmarking framework via cloud-based web pages.

Another widely used tool in SLAM benchmarking is EVO (Grupp 2017). Inspired by TUM RGB-D (Sturm et al. 2012), EVO provides a CLI toolkit with similar syntax and configuration, supports several data formats, and computes accuracy metrics such as ATE and RPE. The supported formats are TUM trajectory files (Sturm et al. 2012), KITTI pose files (Geiger et al. 2013), EuRoC MAV CSV files (Burri et al. 2016), and ROS 1 and ROS 2 bag files, including TF messages, pose, transform, and odometry topics. Furthermore, the toolkit generalizes trajectory plots, performing alignment, and converting between different trajectory representations. Subsequently, EVO’s comprehensive functionality has led to widespread adoption. DLIO (Chen et al. 2023) used EVO to calculate the metrics in their results. Similarly, Hilti 2022 (Zhang et al. 2023) dataset utilized EVO to compute the metrics in their SLAM challenge rankings. KITTI 360 (Liao et al. 2023) employed EVO to calculate metrics and generate the plots presented on its webpage. Also, MOLA (Blanco-Claraco 2025) used EVO to produce quantitative metrics

Table 1. State-of-the-art benchmarking frameworks for SLAM algorithms.

| Benchmarking Framework | Supported Datasets | Sensor Data Types | Metrics | User Interface | Year |
|--------------------------------------|---------------------------------------|--------------------------------------|---|--|------|
| KITTI Vision (Geiger et al. 2012) | KITTI dataset (Geiger et al. 2013) | Monocular, Stereo, LiDAR | RTE (%), RRE (deg/m) | CLI toolkit, Public web server rankings | 2012 |
| TUM RGB-D (Sturm et al. 2012) | TUM RGB-D dataset | RGB-D, IMU | ATE (m), RPE (%) | CLI toolkit | 2012 |
| SLAMBench (Nardi et al. 2015) | Various public Visual SLAM datasets | Monocular, Stereo, RGB-D, IMU | ATE (m), Execution time (s), RAM usage (MB) | CLI framework | 2015 |
| ETH3D (Schöps et al. 2017) | ETH3D dataset | Monocular, Stereo, RGB-D, IMU | ATE (m), RTE (%), RRE (deg/m) | CLI toolkit, Public web server rankings | 2017 |
| EVO (Grupp 2017) | — ⁽¹⁾ | — ⁽¹⁾ | ATE (m), RTE (%), RRE (deg/m) | CLI toolkit | 2017 |
| KITTI 360 (Liao et al. 2023) | KITTI 360 dataset | Monocular, Stereo, LiDAR | ATE (m), RPE (%) | CLI toolkit, Public web server rankings | 2023 |
| SLAM Hive (Yang et al. 2023) | Various public SLAM datasets | Monocular, Stereo, RGB-D, LiDAR, IMU | ATE (m), RTE (%), CPU usage (%), RAM usage (MB) | Web-based graphical interface, Docker-based system | 2023 |
| VBR (Brizi et al. 2024) | VBR dataset | Monocular, Stereo, LiDAR, IMU | ATE (m), RPE (%) | CLI toolkit, Public web server rankings | 2024 |

¹ EVO (Grupp 2017) is a toolkit to evaluate the trajectory output by the SLAM algorithms. The supported trajectory formats are: TUM trajectory files (Sturm et al. 2012), KITTI pose files (Geiger et al. 2013), EuRoC MAV CSV files (Burri et al. 2016), and ROS 1 and ROS 2 bag files, including TF messages, pose, transform, and odometry topics.

Abbreviations: Absolute Trajectory Error (ATE), Command Line Interface (CLI), Inertial Measurement Unit (IMU), Light Detection And Ranging (LiDAR), Robot Operating System (ROS), Relative Pose Error (RPE), Relative Rotational Error (RRE), Relative Translational Error (RTE)

and visualizations when benchmarking its LiDAR Odometry (LO) pipeline on several public datasets.

2.3 Datasets

While datasets have been developed for a wide range of SLAM applications, including autonomous driving (Geiger et al. 2013; Kim et al. 2020), long-term SLAM (Sousa et al. 2023; Carlevaris-Bianco et al. 2015; Maddern et al. 2017), semantic SLAM (Liao et al. 2023), and collaborative SLAM (Feng et al. 2024), there are few datasets focusing on indoor sequences. Moreover, many datasets that include indoor often also feature outdoor sequences (Carlevaris-Bianco et al. 2015; Jiao et al. 2022; Yin et al. 2022; Qingqing et al. 2022; Chen et al. 2024; Feng et al. 2024; Brizi et al. 2024). Narrowing the scope further to ground robots with 3D LiDAR sensors, the availability of suitable datasets becomes even more limited (Carlevaris-Bianco et al. 2015; Jiao et al. 2022; Yin et al. 2022, 2023; Chen et al. 2024; Feng et al. 2024), revealing a gap in resources to evaluate indoor 3D LiDAR-based SLAM. Tab. 2 provides an overview of state-of-the-art SLAM datasets that include indoor sequences and 3D LiDAR data.

A common characteristic among 3D LiDAR datasets with indoor sequences is the reliance on mechanical spinning sensors. Most employ Velodyne or Ouster sensors with 16 to 128 beams. Only a few incorporate more than one 3D LiDAR. Among those, only four datasets (Hilti 2021 (Helberger et al. 2022), TIERS (Qingqing et al. 2022), A Multi-LiDAR Multi-UAV Dataset (Catalano et al. 2023), and GODE (Chen et al. 2024)) employ non-repetitive scanning sensors. However, considering different 3D LiDAR with various FoV and scanning pattern configurations may provide valuable insights into which configurations best suit indoor environments.

Regarding additional sensors to 3D LiDAR, nearly all datasets include an external IMU. The only exceptions are TIERS (Qingqing et al. 2022) and A Multi-LiDAR Multi-UAV Dataset (Catalano et al. 2023), where the internal IMU of the 3D LiDAR is used instead. In the case of ground-based platforms, only NCLT (Carlevaris-Bianco et al. 2015), FusionPortable (Jiao et al. 2022; Wei et al. 2024), and Ground-Challenge (Yin et al. 2023) provide wheel odometry data, which provides a cost-effective source of motion information in industrial and indoor applications from measuring the wheels angular speed.

Furthermore, ground-truth data varies among the datasets. Several rely on Global Navigation Satellite System (GNSS), though this method is most effective in outdoor scenarios or indoors with ample window visibility. Many datasets employ laser scanning techniques to obtain ground-truth data, using devices such as Leica MS60 or adopting MoCap systems like OptiTrack or Vicon Vero cameras. Notably, only three datasets (FusionPortable (Jiao et al. 2022; Wei et al. 2024), Hilti 2022 (Zhang et al. 2023), and GODE (Chen et al. 2024)) also provide ground-truth maps, typically generated from high-precision lasers (e.g., Leica RTC360).

In summary, current datasets reveal possible gaps for indoor-oriented SLAM research. First, many do not have indoor-exclusive sequences. Next, most datasets rely on a limited selection of mechanical spinning LiDARs, primarily Velodyne and Ouster models, with similar beam configurations (symmetric and uniform FoV). This limitation does not allow the evaluation of SLAM algorithms for various LiDAR configurations in terms of FoV characteristics (symmetry, beam distribution) and scanning patterns. Moreover, non-repetitive scanning LiDARs are scarcely represented despite their potential to generate denser point cloud data (Xie et al. 2024). Finally, wheel odometry data remains underutilized in the datasets despite

Table 2. State-of-the-art SLAM datasets with 3D LiDAR sensors and indoor sequences.

| Dataset | Platforms | Sensors | | | | Ground-Truth | | Year |
|---|-----------------------------|--|--------------------------------|-----|----------------|-------------------------|-----|------|
| | | Spinning LiDARs | Non-repetitive scanning LiDARs | IMU | Wheel Odometry | Trajectory | Map | |
| NCLT (Carlevaris-Bianco et al. 2015) | UGV | Velodyne HDL-32E | – | Yes | Yes | GNSS, SLAM | No | 2015 |
| Hilti 2021 (Helmberger et al. 2022) | Handheld, UAV | Ouster OS0-64 | Livox Mid-70 | Yes | – | MoCap | No | 2021 |
| FusionPortable (Jiao et al. 2022; Wei et al. 2024) | Handheld, Legged robot, UGV | Ouster OS1-128 | – | Yes | Yes | Laser scan, GNSS | Yes | 2022 |
| Hilti 2022 (Zhang et al. 2023) | Handheld | Hesai PandarXT-32 | – | Yes | – | Laser scan | Yes | 2022 |
| M2DGR (Yin et al. 2022) | UGV | Velodyne VLP-32C | – | Yes | No | Laser scan, MoCap, GNSS | No | 2022 |
| NTU Viral (Nguyen et al. 2022) | UAV | 2× Ouster OS1-16 | – | Yes | – | Laser scan | No | 2022 |
| TIERS (Qingqing et al. 2022) | Mobile platform | Velodyne VLP-16, Ouster OS1-64, Ouster OS0-128 | Livox Horizon, Livox Avia | No | No | MoCap, SLAM | No | 2022 |
| A Multi-LiDAR Multi-UAV Dataset (Catalano et al. 2023) | UAV | Ouster OS1-64 | Livox Avia, Livox Mid-360 | No | – | MoCap | No | 2023 |
| Ground-Challenge (Yin et al. 2023) | UGV | Velodyne VLP-16 | – | Yes | Yes | – | No | 2023 |
| GEODE (Chen et al. 2024) | Handheld, UGV, Car, Boat | Velodyne VLP-16, Ouster OS1-64 | Livox Avia | Yes | No | Laser scan, MoCap, GNSS | Yes | 2024 |
| S3E (Feng et al. 2024) | UGVs | Velodyne VLP-16 | – | Yes | No | MoCap, GNSS | No | 2024 |
| VBR (Brizi et al. 2024) | Handheld | Ouster OS1-64, Ouster OS0-128 | – | Yes | – | GNSS | No | 2024 |

Abbreviations: Global Navigation Satellite System (GNSS), Inertial Measurement Unit (IMU), Motion Capture (MoCap), Unmanned Aerial Vehicle (UAV), Unmanned Ground Vehicle (UGV)

being a cost-effective alternative to IMU sensors for estimating relative motion data (Korayem et al. 2006). As a result, addressing these shortcomings by developing diverse, sensor-rich datasets is crucial for advancing robust indoor SLAM.

3 IILABS 3D Dataset

In order to address the limitations of existing datasets and provide a more comprehensive resource for evaluating indoor LiDAR-based SLAM on ground-wheeled robots, we developed a new dataset named IILABS 3D – iilab Indoor LiDAR-based SLAM 3D (Ribeiro et al. 2025). This dataset is publicly available in the INESC TEC – Institute for Systems and Computer Engineering, Technology and Science research data repository¹. Unlike existing datasets, which often lack components such as diverse 3D LiDAR configurations or wheel odometry data, IILABS 3D is specifically designed to fill these gaps and support the benchmarking of LiDAR-based SLAM algorithms in indoor environments.

IILABS 3D features four different 3D LiDARs (Velodyne VLP-16, Ouster OS1-64, RoboSense RS-Helios-5515, and Livox Mid-360) with varying characteristics and design applications, offering a diverse range of sensor configurations for testing. The dataset also includes pre-processed wheel odometry data and IMU data to enhance sensor fusion capabilities. Additionally, the dataset contains calibration sequences to ensure transparency and

reproducibility in benchmarking studies and challenge sequences to evaluate the robustness of algorithms under various conditions.

The following subsections provide detailed descriptions of the dataset’s key components: the mobile robot and sensors used (Sec. 3.1), the ground-truth system employed (Sec. 3.3), and the dataset sequences designed to present a range of challenges and benchmarking scenarios (Sec. 3.2).

3.1 Mobile robot and sensors

The dataset was collected using a mobile robot platform equipped with 3D sensor perception, as shown in Fig. 1. This mobile robot, which builds on the Hangfa Discovery Q2 platform, was developed at CRIIS – Centre for Robotics in Industry and Intelligent Systems from INESC TEC. Detailed information about the robot is available in the work by Sousa et al. (2025) and on its public webpage⁴.

The Hangfa Discovery Q2 is a compact four-wheeled omnidirectional robot developed by Hangfa Robotics⁵. This robot has a coaxial pendulum suspension system on its rear wheels, ensuring ground contact of the four wheels while reducing vibrations on uneven surfaces. Also, the platform features QMA10 mecanum wheels with a 101.6 mm diameter, a load capacity of 30 kg, a 0.65 m/s maximum translational, and a 140°/s maximum rotation speed. However, the original platform lacks access to wheel odometry data and support for advanced sensing devices. As a result, the work of Sousa et al. (2025) addressed



Figure 1. Revised Hangfa Discovery Q2 mobile platform with 3D sensor perception

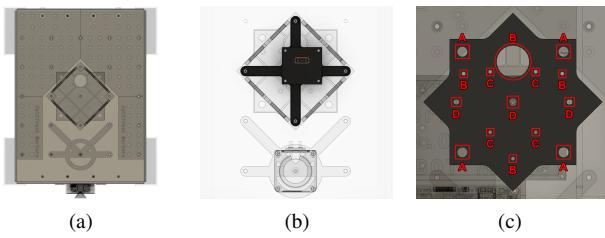


Figure 2. 3D model views on AutoDesk Fusion 360 for the integration of the sensors into the Hangfa Discovery Q2 platform: (a) top view of the bend metal sheet; (b) support on the bottom of the bend metal sheet for the IMU; (c) support on the top of the bend metal sheet for 3D LiDARs (A: Ouster OS1-64; B: RoboSense RS-Helios-5515; C: Livox Mid-360; D: Velodyne VLP-16).

those issues, while integrating multimodal perception into the platform.

Regarding sensor integration, the dataset acquisition considered three types of sensors equipped on the robot: four 3D LiDARs (Velodyne VLP-16, Ouster OS1-64, RoboSense RS-Helios-5515, and Livox Mid-360), a 2D laser scanner (Hokuyo UST-10LX), and an IMU (Xsens MTi-630 AHRS). These sensors are fixated on the platform using the 3D-printed supports illustrated in Fig. 2 and designed in AutoDesk Fusion 360. While the 3D LiDAR and IMU sensors are centered in the robot's geometric center, the 2D laser scanner is positioned forward without interfering with the data from any of the 3D LiDARs considered for the dataset. Indeed, the generic support for 3D LiDAR from Sousa et al. (2025) ensures the optical point alignment with the robot's geometric center and minimal occlusion by the chassis while accommodating a range of sensors – including Livox Mid-360, Ouster OS, RoboSense Helios, and Velodyne Puck series.

Furthermore, the four 3D LiDARs selected for the dataset offer various technologies and configurations, including mechanical spinning (Velodyne, Ouster, RoboSense) and non-repetitive scanning (Livox) data, a diverse FoV angles (see Fig. 3), and different beam configurations (16, 32, and 64 beams). This selection enriches the dataset's applicability across different SLAM scenarios (ground versus ceiling-based perception, symmetrical versus asymmetrical FoV of the sensor). Therefore, each 3D LiDAR used in the study contributes uniquely to the dataset, as follows:

- **Velodyne VLP-16:** a widely used sensor in recent datasets due to its robustness and compatibility with SLAM algorithms. It features 16 beams and a symmetrical vertical FoV from -15° to $+15^\circ$, serving as a reliable baseline for benchmarking;
- **Ouster OS1-64:** common in autonomous driving and indoor datasets, this sensor provides high-resolution data with 64 beams and a vertical FoV from -21.1° to $+21.27^\circ$. It supports configurable horizontal resolutions (512, 1024, or 2048) and operational frequencies (10 Hz or 20 Hz). An integrated IMU enhances SLAM compatibility and reduces reliance on external sensors;
- **RoboSense RS-Helios-5515:** a cost-effective sensor with a vertical FoV from -55° to $+15^\circ$. Its design concentrates lasers densely in the center while sparsely covering the edges, improving near-field accuracy and blind-spot detection. This configuration supports simultaneous long-range perception, making it ideal for mobile robots in dynamic environments;
- **Livox Mid-360:** a non-repetitive scanning LiDAR with 360° horizontal coverage and a vertical FoV from -7° to $+52^\circ$. With a range of approximately 40 meters, it is well-suited for indoor applications like warehouses, where it captures semi-static features such as ceilings and shelves, aiding SLAM algorithms in maintaining consistent localization.

In addition to 3D LiDARs, the robot is equipped with a Hokuyo UST-10LX 2D laser and an Xsens MTi-630 AHRS IMU. The 2D laser enables real-time 2D localization and parametric-based trajectory navigation (Sobreira et al. 2015) during the dataset acquisition. This approach ensures consistent trajectory reproduction regardless of which 3D LiDAR is present on the robot. Regarding the Xsens IMU, the dataset provides data from this sensor at a frequency of 400 Hz, enabling support for sensor fusion with inertial odometric data in sequences in which the LiDAR sensor does not have an internal IMU. In comparison, the Ouster OS1-64 and Livox Mid-360 sensors operate their internal IMU at 100 Hz and 200 Hz, respectively. However, when using the UART-to-USB board included in the development kit to communicate with the IMU, the communication driver based on FTDI chipsets experiences buffering when operating at frequencies higher than 100 Hz. Since IMU data messages are timestamped upon arrival in the ROS driver, the header timestamps are affected by this buffering, as illustrated in Fig. 4a. In order to address this issue, post-processing is applied using the internal clock timestamps available in the *time_ref* topic (see Fig. 4b) and an offset correction, considering the last elements of each

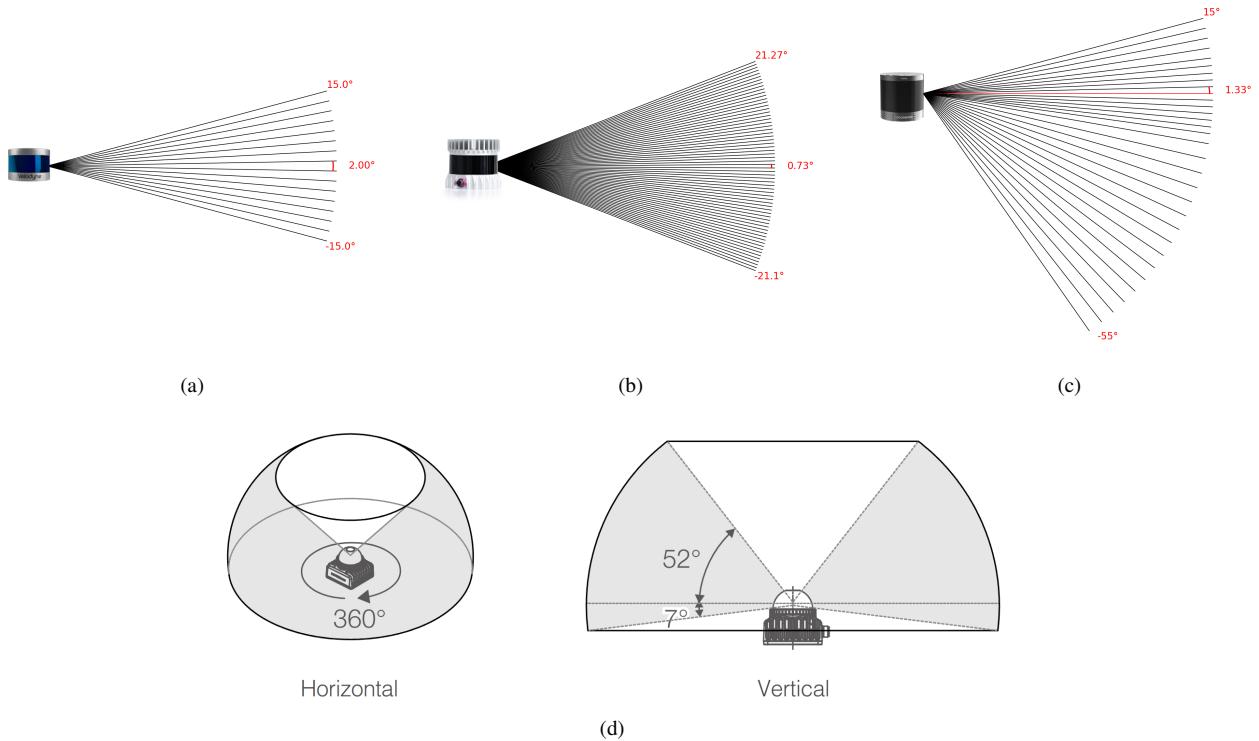


Figure 3. Field of View (FoV) of the different 3D LiDARs: (a) Velodyne VLP-16; (b) Ouster OS1-64; (c) RoboSense RS-Helios-5515; (d) Livox Mid-360.

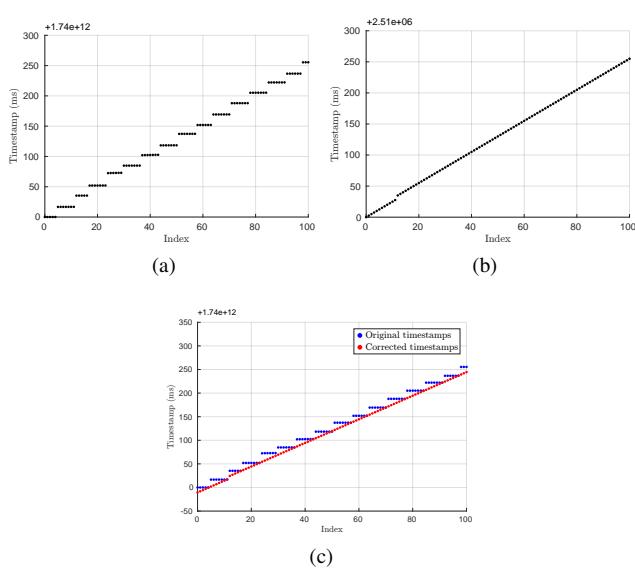


Figure 4. Timestamp buffering correction for the Xsens MTI-630 AHRS IMU: (a) original header timestamps; (b) timestamps from the *time_ref* topic; (c) corrected timestamps.

buffer segment. The effect of this correction is illustrated in Fig. 4c.

Finally, the dataset also includes wheel odometry data in two forms: raw encoder readings and pre-processed odometry. The raw data consists of encoder counts, offering low-level measurements for researchers interested in custom odometry processing. Additionally, pre-processed odometry data is provided using a ROS package developed by the 5dpo

Robotics Team⁶ and published as open-source by Sousa et al. (2024) that implements wheel odometry estimation for four-wheeled mecanum robots. This package, part of a broader robotic framework designed for the Robot@Factory 4.0 competition, is available on GitHub⁷.

3.2 Dataset sequences

The dataset is designed to include two distinct types of sequences: calibration and benchmark sequences. Calibration sequences are intended to provide the necessary data for replicating the results presented in this study and to support alternative calibration methodologies. These sequences ensure that intrinsic and extrinsic sensor calibration can be performed by users of the dataset. On the other hand, benchmark sequences are tailored to evaluate the performance of SLAM algorithms under varying conditions.

3.2.1 Calibration sequences First, the extrinsic calibration sequences are designed to determine the transformation between the IMU and 3D LiDAR reference frames. While this transformation can be derived from the distances specified in the Computer Aided Design (CAD) 3D models of the platform (Sousa et al. 2025), these sequences allow an alternative calibration method when the CAD data lacks sufficient precision. Two distinct extrinsic calibration sequences were performed. The first sequence involves lifting the robot in the air, followed by rotations around the z-axis and oscillations along the x-axis and y-axis. This setup ensures excitation in all three axes, making it compatible with various IMU–LiDAR calibration algorithms, such as LI-Init (Zhu et al. 2022). The second sequence follows a ground-based figure-eight trajectory, also known as the lemniscate of Bernoulli, as

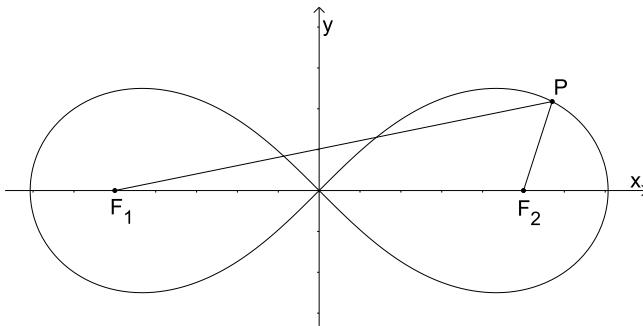


Figure 5. Extrinsic calibration using a ground-based figure-eight trajectory (lemniscate of Bernoulli).

illustrated in Fig. 5. This trajectory is particularly suitable for extrinsic calibration algorithms intended for ground robots, such as GRIL-Calib (Kim et al. 2024). The lemniscate of Bernoulli trajectory is mathematically defined as follows, where $PF_1 \cdot PF_2 = c^2$ and $a = c\sqrt{2}$:

$$\begin{cases} x = \frac{a \cdot \cos(t)}{1 + \sin^2(t)} \\ y = \frac{a \cdot \sin(t) \cdot \cos(t)}{1 + \sin^2(t)} \end{cases} \quad (1)$$

Next, the intrinsic calibration sequence for the IMU involves a three-hour stationary recording, during which the robot remains motionless. This procedure is intended to estimate the white noise and bias characteristics of IMU sensors. The dataset includes this calibration sequence for all three IMUs: the Xsens MTi-630 AHRS IMU and the internal IMUs of the Ouster OS1-64 and Livox Mid-360 3D LiDAR sensors.

As a result, the intrinsic calibration sequences can be used to generate Allan deviation plots to estimate the white noise and bias characteristics of both the gyroscope and accelerometer in the IMU. Preprocessed values for these parameters are provided in the dataset as YAML files, computed using the open-source package Allan Variance ROS (Buchanan 2021). For illustration, the calibration results of the Xsens MTi-630 AHRS IMU are presented in Fig. 6, while the same methodology applies to the internal IMUs of the other sensors. The Xsens IMU calibration yielded an accelerometer noise density of approximately $0.002 \text{ m/s}^2/\sqrt{\text{Hz}}$ and random walk of around $1.8 \times 10^{-5} \text{ m/s}^3/\sqrt{\text{Hz}}$. For the gyroscope, the noise density was approximately $1.4 \times 10^{-4} \text{ rad/s}/\sqrt{\text{Hz}}$, with a random walk of about $7.8 \times 10^{-6} \text{ rad/s}^2/\sqrt{\text{Hz}}$. To ensure a safe margin in the benchmark study, the YAML file values were amplified – bias values by a factor of 10 and white noise values by a factor of 5. These results, illustrated in Fig. 6, exemplify the noise characteristics and bias stability of the IMU sensors in the dataset.

The wheel odometry calibration sequences consist of a set of square-shaped trajectories, a widely used approach for odometry calibration in mobile robotics (Sousa et al. 2022). Four distinct trajectories were executed, including both clockwise (CW) and counterclockwise (CCW) versions (see Fig. 7), with and without rotational maneuvers at each corner. As a result, the dataset includes sixteen odometry calibration sequences – four for each 3D LiDAR

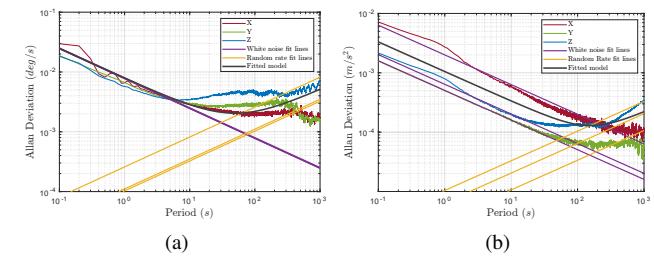


Figure 6. Allan Deviation plots of the Xsens MTi-630 AHRS IMU generated by package Allan Variance ROS (Buchanan 2021): (a) gyroscope; (b) accelerometer.

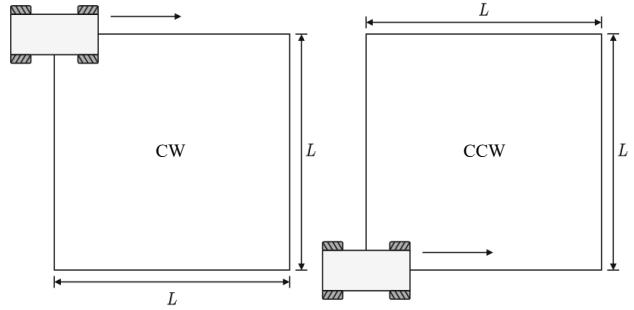


Figure 7. Square-shaped trajectories used for odometry calibration, performed in both CW and CCW directions.

–, where the robot completes two full laps of each $2 \text{ m} \times 2 \text{ m}$ square-shape trajectory. Although odometry calibration parameters are independent of other sensors, including the 3D LiDARs, providing a larger number of sequences enhances the precision and reliability of the calibration process. Furthermore, these sequences can also serve as alternative data for extrinsic calibration algorithms, complementing or substituting the figure-eight trajectory sequence when needed.

3.2.2 Benchmark sequences On the other hand, the benchmark sequences were designed to evaluate the performance of SLAM algorithms under diverse conditions, introducing challenges that range from holonomic motions to non-planar trajectories. All sequences were collected at iilab, with the majority conducted within the Nav A space, where the ground-truth system is installed. The Nav A space is characterized by its feature-rich environment, a fully planar floor, and the presence of large glass doors and windows, which can introduce challenges for LiDAR-based SLAM due to reflections and refractions.

In addition to the natural characteristics of the environment, specific challenges were deliberately introduced in some sequences to test SLAM robustness. These challenges include holonomic trajectories. In these cases, translation and rotation are decoupled by using the robot's omnidirectional steering. Another challenge is induced slippage, which causes significant wheel odometry drift. In addition, non-planar trajectories involve a ramp and an elevator, and high-speed motions introduce vibrations into the mobile robot.

Additionally, some benchmark sequences extend beyond the Nav A area, briefly exploring locations such as the first and second floor corridors or the Nav B space. Although ground-truth data cannot be directly provided in these

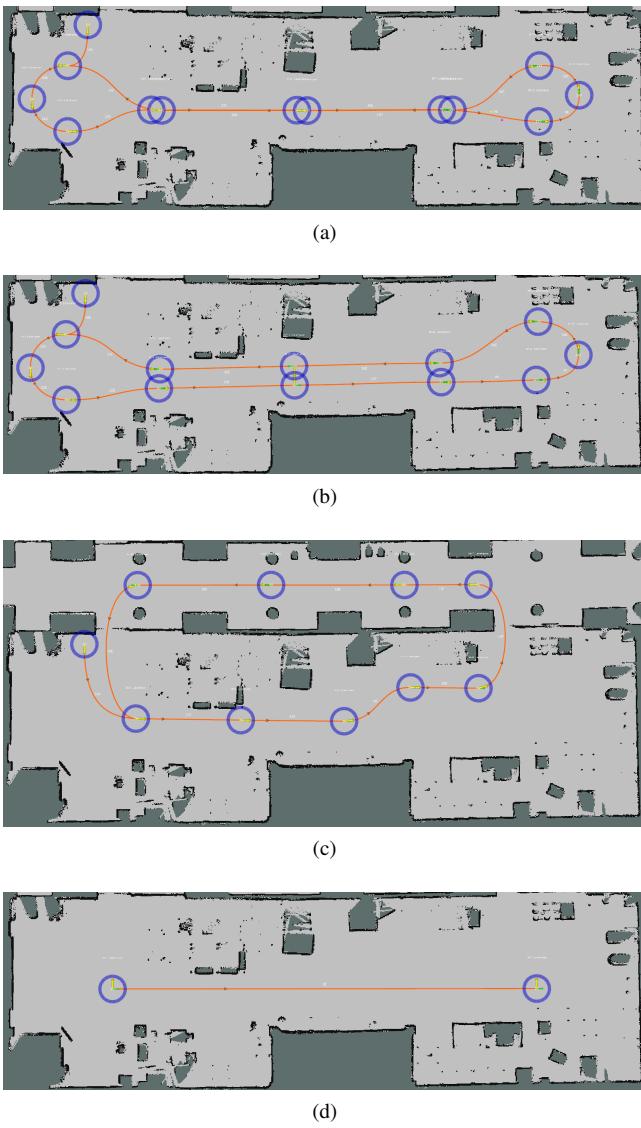


Figure 8. Benchmark trajectories analyzed in this study: (a) Nav A Diff; (b) Nav A Omni; (c) Loop; (d) Slippage.

regions, all sequences begin and end within Nav A, ensuring that accurate ground-truth information is available at those timestamps. The trajectories for each sequence are depicted in Fig. 8 and Fig. 9, while illustrative photographs of the elevator and ramp are provided in Fig. 9b and Fig. 10. Furthermore, a explanation of each sequence is presented below:

- **Nav A Diff:** A 275-meter, 755-second sequence completing 5 cycles within the Nav A space. It serves as a baseline, capturing typical environmental conditions without additional challenges;
- **Nav A Omni:** A 112-meter, 387-second sequence with 2 cycles, featuring holonomic motion. The robot executes decoupled translational and rotational movements using omnidirectional steering;
- **Loop:** A 232-meter, 620-second sequence with 5 cycles. The robot exits and re-enters the Nav A space via different doors, briefly traversing an external corridor. This sequence was designed to assess the loop closure capabilities of the SLAM algorithms;

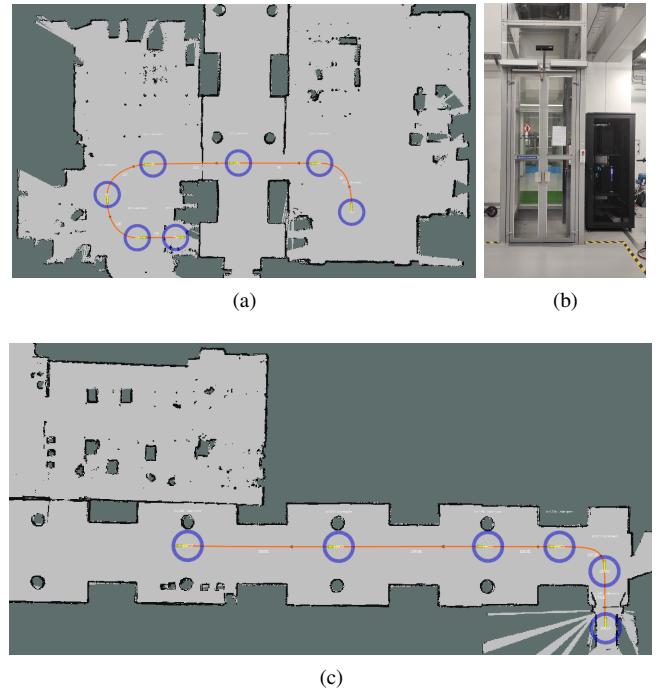


Figure 9. Elevator benchmark sequence: (a) trajectory on the first floor; (b) photograph of the elevator; (c) trajectory on the second floor.

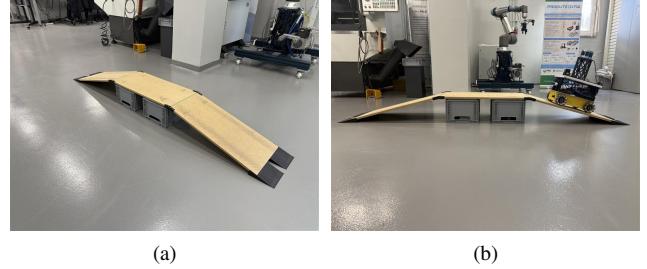


Figure 10. Ramp benchmark sequence: (a) ramp used in the experiment; (b) mobile robot navigating the ramp during the sequence.

- **Slippage:** A 39-meter, 93-second sequence with only one cycle. The robot follows a straight-line path while performing lateral motions at high speeds that intentionally induce wheel slippage. This setup challenges SLAM algorithms that rely on wheel odometry by introducing significant drift;
- **Ramp:** A 29-meter, 168-second sequence with 3 cycles, manually controlled via joystick. The primary challenge of this sequence is the non-planar motion while ascending and descending a ramp, requiring SLAM to capture vertical displacements and pitch rotations;
- **Elevator:** An 85-meter, 381-second sequence with only one cycle. The robot moves between floors via an elevator with glass windows. This sequence tests whether vertical motion can be reliably detected through IMU sensor fusion or directly via 3D LiDAR point cloud analysis.

Moreover, each benchmark sequence is repeated for all four 3D LiDAR sensors. In order to ensure consistency

across all sequences, the 2D LiDAR and the robot's 2D localization and navigation stack (Sobreira et al. 2015) were employed in every sequence, except for the Ramp sequence, where manual joystick control was used to capture non-planar motion dynamics.

Finally, in order to validate the induced odometry drift observed in the *slippage* sequence, we conducted a series of tests that evaluated both longitudinal and lateral motions at different linear velocities along the line-shaped trajectory. The Relative Translational Error (RTE) was computed on the wheel odometry data in 1-meter segments using the EVO tool (Grupp 2017), as illustrated in Fig. 11. The results confirm that odometry drift is more pronounced during longitudinal motions, likely due to the inherent limitations of the omnidirectional mobile robot's odometry model along that axis. Furthermore, when the linear velocity is increased, the mean RTE for longitudinal motions rises from 0.021 m to 0.035 m (a 66% increase), whereas for lateral motions, it increases from 0.064 m to 0.159 m (a 148% increase). These findings indicate that lateral motions are particularly susceptible to slippage-induced drift, thereby corroborating the use of this motion style in the *slippage* sequence.

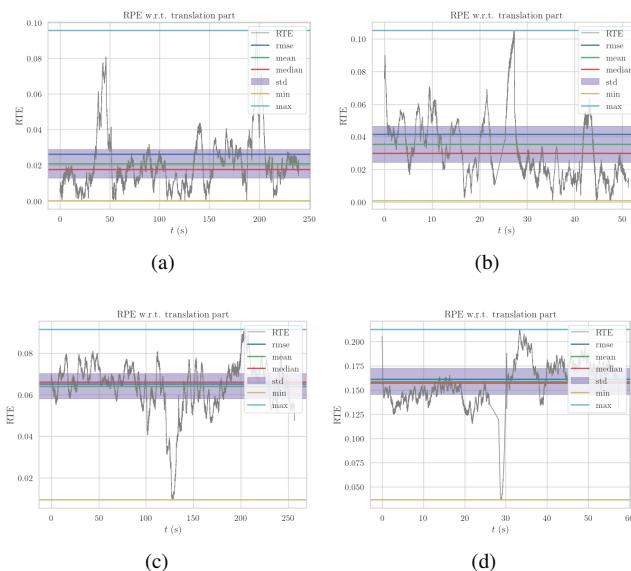


Figure 11. Odometry drift validation the Slippage trajectory: (a) with longitudinal motion at 0.1 m/s; (b) with longitudinal motion at 0.5 m/s; (c) with lateral motion at 0.1 m/s; (d) with lateral motion at 0.5 m/s.

3.3 Ground-truth system

We employed the OptiTrack MoCap system⁸ to provide the ground-truth trajectory data for the dataset. MoCap systems like OptiTrack are widely used in robotics for providing high-precision ground-truth data, particularly in indoor environments where Global Navigation Satellite System (GNSS) is unavailable (Feng et al. 2024; Yin et al. 2022). OptiTrack leverages a network of cameras to track reflective markers with sub-millimeter accuracy, offering a reliable reference for evaluating robot localization and mapping algorithms. This study utilized the OptiTrack system installed in the iilab at INESC TEC, as shown in Fig. 12. The system operates at a frequency of 240 Hz and

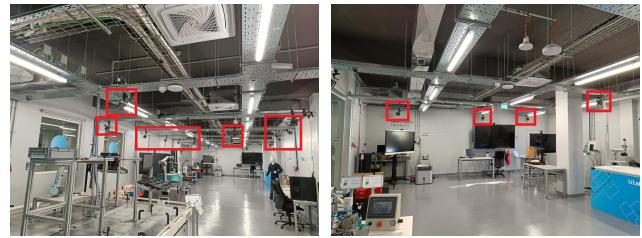


Figure 12. OptiTrack system installed in the iilab at INESC TEC.

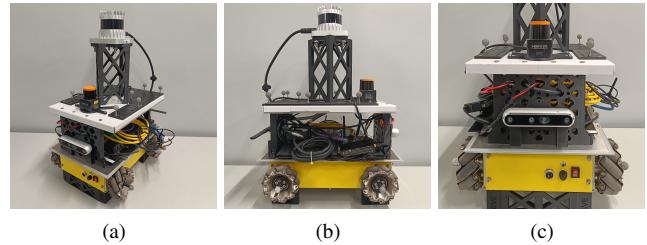


Figure 13. Reflective sphere setup on the robot: (a) perspective view; (b) side view; (c) front view.

comprises 24 high-resolution cameras – model Prime^X 22 – strategically positioned throughout the lab to ensure full coverage of the robot's operating area.

The ground-truth setup on the robot involves the integration of reflective spheres mounted in specific configurations to facilitate precise tracking and alignment. As illustrated in Fig. 13, four reflective spheres are attached to the robot's wheels to define the base link, aligning the MoCap rigid body with the robot's reference frame. Additionally, nine reflective spheres are placed on the robot's top in a non-symmetric structure. This asymmetrical arrangement ensures that the rigid body is uniquely identifiable by the OptiTrack system, even under occlusion or partial visibility. The setup process is configured using the Motive software⁹, specifically designed to work with the OptiTrack system. The steps for setting up the ground-truth measurement are as follows:

1. Define the rigid body in the Motive software, aligning it with the robot's reference frame by incorporating the four reflective spheres mounted on the wheels;
2. Add the additional spheres on the top of the robot to the rigid body as constraints, enhancing its stability and accuracy during tracking;
3. Detach the wheel-mounted spheres from the rigid body definition. This ensures that their motion may occur during robot operation and does not introduce errors in the rigid body's pose estimation.

In order to record the ground-truth data, the NatNet 4 ROS driver¹⁰ was employed to capture and store the OptiTrack data in rosbags. The OptiTrack system captures the positions of reflective spheres using cameras, which transmit raw positional data via Ethernet to a main computer running the Motive software in a Windows environment. The Motive software processes this data to compute the poses of rigid bodies and individual markers, transmitting the information

via User Datagram Protocol (UDP) to a secondary system running the NatNet 4 ROS driver on Ubuntu.

As for determining whether a secondary computer was necessary for handling ground-truth data, a comparative study was conducted on the performance of Wi-Fi versus Ethernet connections for transmitting OptiTrack data. While the NatNet 4 ROS driver could theoretically run directly on the robot using a Wi-Fi connection, the robot's inability to connect to the main computer via Ethernet during dataset recording posed a challenge. This limitation made it essential to evaluate whether the use of a secondary computer connected via Ethernet could provide a significant advantage. A fast and stable connection was critical since the timestamps of the ROS messages are based on the ROS time when the NatNet 4 ROS driver processes the message, rather than the time when the Motive software generates it. Tests revealed that an Ethernet connection offered significantly reduced jitter compared to Wi-Fi, ensuring higher precision and reliability in data transmission. As illustrated in Fig. 14, the jitter values were 1.1607 ms for the Wi-Fi setup and 0.8472 ms for the Ethernet configuration. This notable reduction highlights the importance of using a secondary computer with an Ethernet connection to ensure accurate ground-truth data recording.

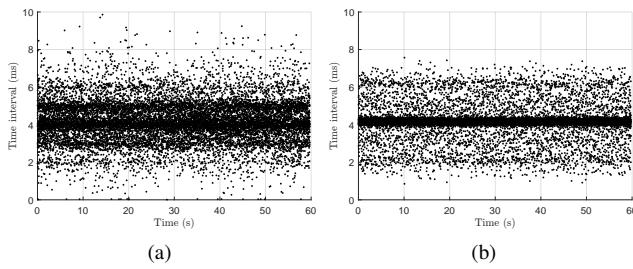


Figure 14. Jitter comparison for OptiTrack data transmission: (a) Wi-Fi connection; (b) Ethernet connection.

Given the use of a secondary computer for handling ground-truth data, clock synchronization between this computer and the robot was essential for accurate benchmarking. Initial experiments utilized the Network Time Protocol (NTP) for direct time synchronization, with the secondary computer serving as the NTP server and the robot as the client. However, this configuration resulted in offsets and jitters of approximately 100ms. The issue stemmed from variable delays in the Wi-Fi connection, where the time taken to transmit messages could fluctuate significantly. Such high latency was unacceptable for this study, as the benchmark requires precise synchronization between the odometry provided by the SLAM algorithms and ground-truth data. Since the dataset includes 10 Hz LiDAR data, even small timing discrepancies can lead to significant errors in evaluating the SLAM algorithm's performance. In order to address this issue, an alternative synchronization strategy was implemented, utilizing a public NTP server in both devices, specifically the Lisbon Astronomical Observatory (`ntp02.oal.ul.pt`). This configuration achieved significantly smaller offsets and jitter, approximately 1 ms for the Ethernet-connected secondary computer and 10 ms for the robot connected via Wi-Fi.

Consequently, the offset and jitter between robot and secondary computer were reduced to around 10 ms, ensuring adequate synchronization for the dataset. The unexpected improvement with the external server is likely related to environmental factors, such as fluctuating Wi-Fi transmission delays, affecting the local NTP setup. So, we recommend that readers replicate this benchmark setup and conduct similar synchronization tests in their own environments to determine the most effective configuration for their specific conditions.

Finally, the accuracy of our ground-truth system was validated through an OptiTrack calibration procedure. The calibration yielded a three-dimensional error of 0.646 mm, as documented in the corresponding calibration file. This error pertains to a single sphere of the OptiTrack system. Notably, the spheres used in the mobile robot's rigid body are 15.9 mm in diameter – the same size as those in the calibration tool. Consequently, since the robot's rigid body consists of nine spheres, the overall pose error is expected to be smaller. This sub-millimeter precision confirms the reliability of our ground-truth data, thereby ensuring that the dataset serves as a robust benchmark for evaluating SLAM algorithms.

4 Benchmark

This section presents a comprehensive benchmark study of nine state-of-the-art 3D LiDAR-based SLAM algorithms. The evaluation uses the dataset described in Sec. 3. By benchmarking multiple SLAM approaches on this dataset, we provide a detailed comparison that highlights their performance in terms of accuracy and robustness.

The section is structured as follows. First, in Sec. 4.1, the methodology is introduced, describing the evaluation metrics, the selected algorithms, and the overall benchmarking process. Next, in Sec. 4.2, the results are presented in a structured manner, summarizing the performance of each approach. Finally, in Sec. 4.3, the results are analyzed and discussed, considering both the differences between the algorithms and the impact of the 3D LiDAR sensors used.

4.1 Methodology

This study focuses on a selection of open-source, 3D LiDAR-based SLAM algorithms that collectively represent a diverse set of strategies, sensor integrations, and loop closure techniques. As summarized in Tab. 3, the algorithms include foundational methods such as LOAM (Zhang and Singh 2014) and LeGO-LOAM (Shan and Englot 2018), as well as more recent contributions like GLIM (Koide et al. 2024), Kinematic-ICP (Guadagnino et al. 2024), and MOLA-LO (Blanco-Claraco 2025). Some rely solely on LiDAR data, as exemplified by KISS-ICP (Vizzo et al. 2023), while others fuse data from additional sensors, such as IMUs (LIO-SAM (Shan et al. 2020), DLIO (Chen et al. 2023)) or wheel odometry (Kinematic-ICP (Guadagnino et al. 2024), VineSLAM (Aguiar et al. 2022)). They also differ in loop closure strategies: certain methods employ basic distance/time thresholds (LeGO-LOAM (Shan and Englot 2018), GLIM (Koide et al. 2024)), whereas others leverage advanced descriptors using Scan Context (Kim and Kim 2018; Kim et al. 2022). Furthermore, although most of

Table 3. Overview of the 3D LiDAR-based SLAM algorithms evaluated in this benchmark study.

| SLAM Algorithm | Open-source Code | ROS Version | 3D LiDAR Support | | | | Additional Sensor Support | | Loop Closure | Year |
|---|------------------------------|----------------|------------------|--------|------------------|---------|---------------------------|------------------|--------------|------|
| | | | VLP-16 | OS1-64 | RS-5515 | Mid-360 | IMU | Wheel Odometry | | |
| LOAM (Zhang and Singh 2014) | A-LOAM (Qin and Cao 2019) | Noetic (ROS 1) | X | X | X ⁽¹⁾ | - | - | - | - | 2014 |
| LeGO-LOAM (Shan and Englot 2018) | LeGO-LOAM-BOR ¹¹ | Noetic (ROS 1) | X | X | X ⁽¹⁾ | - | - | - | X | 2018 |
| LIO-SAM (Shan et al. 2020) | LIORF ¹² | Noetic (ROS 1) | X | X | X | - | X | - | X | 2020 |
| DLIO (Chen et al. 2023) | Official | Noetic (ROS 1) | X | X | X | X | X | - | - | 2022 |
| VineSLAM (Aguiar et al. 2022) | Official | Humble (ROS 2) | X | X | X | X | X | X | - | 2022 |
| KISS-ICP (Vizzo et al. 2023) | Official | Humble (ROS 2) | X | X | X | X | - | - | - | 2023 |
| GLIM (Koide et al. 2024) | Official | Humble (ROS 2) | X | X | X | X | X | - | X | 2024 |
| Kinematic-ICP (Guadagnino et al. 2024) | Official | Humble (ROS 2) | X | X | X | - | - | X | - | 2024 |
| MOLA-LO (Blanco-Claraco 2025) | Official | Humble (ROS 2) | X | X | X | X | - ⁽²⁾ | - ⁽²⁾ | - | 2024 |

¹ LOAM and LeGO-LOAM are optimized for mechanical spinning LiDARs that provide a symmetric and uniform vertical field of view. The RoboSense RS-Helios-5515 does not meet these criteria; however, data from sequences recorded with this sensor were still included for comparative analysis.

² Although MOLA-LO officially supports sensor fusion with both an IMU and wheel odometry, we were unable to evaluate these configurations in our experiments because the sensor fusion models were still insufficiently documented at the time of this article.

Abbreviations: Inertial Measurement Unit (IMU), Robot Operating System (ROS), Livox Mid-360 (Mid-360), Ouster OS1-64 (OS1-64), RoboSense RS-Helios-5515 (RS-5515), Velodyne VLP-16 (VLP-16)

these approaches are based on LO, VineSLAM (Aguiar et al. 2022) adopts a particle-filter framework.

Additionally, not all algorithms officially support every 3D LiDAR sensor. In particular, the oldest algorithms contain feature extraction modules that limit their support for non-repetitive scanning 3D LiDAR sensors, such as the Livox Mid-360. Moreover, LOAM (Zhang and Singh 2014) and LeGO-LOAM (Shan and Englot 2018) were initially optimized for mechanical spinning LiDARs that have a symmetric, uniform vertical FoV. Consequently, the RoboSense RS-Helios-5515 does not fully meet these criteria. However, we still evaluated the performance of these algorithms with this sensor for comparative analysis. Furthermore, although MOLA-LO (Blanco-Claraco 2025) officially supports both IMU and wheel-odometry fusion, we were unable to evaluate these configurations in our experiments, as the sensor fusion modules are, at the moment, still under development.

Given that some original implementations were built on legacy code bases, this study employs refined forks or variants of the older algorithms. Specifically, A-LOAM (Qin and Cao 2019) replaces LOAM to enhance code clarity and numerical stability. Likewise, LeGO-LOAM-BOR¹¹ refines LeGO-LOAM by improving its software structure. However, both of these implementations have removed the loosely coupled IMU-based deskewing step present in the original algorithms – a drawback not critical for our low-speed indoor dataset. In addition, LIORF¹² extends LIO-SAM by supporting a broader range of sensors (e.g., RoboSense LiDARs and 6-axis IMUs) and incorporating a Scan Context-based loop closure module. These updated forks retain the original algorithms' core principles while ensuring compatibility with modern hardware and software frameworks.

Moreover, this benchmark study employs the open-source tool EVO (Grupp 2017) to compute and analyze

the following key metrics for evaluating SLAM algorithm performance:

- **Absolute Trajectory Error (m):** This metric evaluates the overall trajectory accuracy by computing the Root Mean Square Error (RMSE) between time-aligned ground truth and estimated trajectories. Prior to error calculation, trajectories are spatially aligned using Umeyama's method (Umeyama 1991);
- **Relative Translational Error (%) and Relative Rotational Error (/m):** Initially introduced in the KITTI Vision Benchmark Suite (Geiger et al. 2013), these metrics quantify the drift in odometry by comparing relative pose estimates over successive trajectory segments. RTE expresses translational drift as a percentage of segment length, while RRE measures rotational drift in degrees per meter. These metrics provide a systematic assessment of how errors accumulate over time.

Finally, the benchmark study results were obtained by evaluating the selected SLAM algorithms across all benchmark sequences from the IILABS 3D dataset. In total, the selected SLAM algorithms span a broad range of release years, reflecting both long-standing foundational approaches and the latest research advances. In order to address compatibility challenges (e.g., some algorithms require ROS 1 on Ubuntu 20.04, whereas others need ROS 2 on Ubuntu 22.04), we used Docker containers¹³. This containerized setup ensured that each SLAM system operated under consistent computational conditions, allowing for fair and controlled evaluations across the full set of benchmark sequences. In order to promote reproducibility, the Dockerfiles and ROS packages used for launching the SLAM algorithms are available as open-source at the GitHub². Furthermore, a dedicated toolkit for downloading the dataset and computing the evaluation metrics is also available on GitHub³.

Table 4. Benchmark results of 3D LiDAR-based SLAM algorithms using sequences from Velodyne VLP-16 and Ouster OS1-64 sensors. Metrics include Absolute Trajectory Error (ATE) in meters (m), Relative Translational Error (RTE) in percentage (%), and Relative Rotational Error (RRE) in degrees per meter (°/m).

| SLAM Algorithm | Velodyne VLP-16 Sequences | | | | | Ouster OS1-64 Sequences | | | | |
|---|---------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| | Nav A Diff | Nav A Omni | Loop | Slippage | Ramp | Nav A Diff | Nav A Omni | Loop | Slippage | Ramp |
| A-LOAM (Qin and Cao 2019) | 0.031 m <u>1.28%</u> | 0.049 m 1.58% | 0.039 m 1.68% | 0.040 m 1.53% | 0.031 m 0.38% | 0.039 m 1.11% | 0.043 m 1.34% | 0.029 m 1.15% | 0.041 m 1.18% | 0.038 m 0.39% |
| | 0.066 °/m | 0.084 °/m | 0.104 °/m | 0.083 °/m | 0.072 °/m | 0.072 °/m | 0.089 °/m | 0.080 °/m | 0.088 °/m | 0.052 °/m |
| LeGO-LOAM-BOR ¹¹ | 0.054 m 1.33% | 0.047 m 1.43% | 0.042 m 1.67% | 0.053 m 1.87% | 0.050 m 0.75% | 0.042 m 1.14% | 0.036 m 1.23% | 0.033 m 1.43% | 0.045 m 1.44% | 0.036 m 0.52% |
| | 0.114 °/m | 0.110 °/m | 0.119 °/m | 0.151 °/m | 0.189 °/m | 0.097 °/m | 0.106 °/m | 0.083 °/m | 0.104 °/m | 0.150 °/m |
| LIORF ¹² | <u>0.030 m</u> 1.33% | <u>0.027 m</u> 1.43% | <u>0.025 m</u> 1.63% | <u>0.035 m</u> 1.40% | <u>0.030 m</u> 0.36% | 0.032 m 1.06% | 0.031 m 1.24% | 0.021 m 1.31% | <u>0.028 m</u> 1.19% | 0.017 m 0.21% |
| | 0.057 °/m | 0.064 °/m | 0.086 °/m | 0.084 °/m | 0.078 °/m | <u>0.060 °/m</u> | <u>0.070 °/m</u> | <u>0.055 °/m</u> | <u>0.057 °/m</u> | <u>0.026 °/m</u> |
| DLIO (Chen et al. 2023) | 0.064 m 1.77% | 0.049 m 1.81% | 0.036 m 1.69% | 0.065 m 1.42% | 0.031 m <u>0.30%</u> | 0.042 m 1.40% | 0.032 m 1.49% | 0.029 m 1.41% | 0.044 m 1.23% | 0.020 m <u>0.19%</u> |
| | 0.095 °/m | 0.121 °/m | 0.093 °/m | 0.075 °/m | 0.112 °/m | 0.086 °/m | 0.110 °/m | 0.083 °/m | 0.062 °/m | 0.080 °/m |
| VineSLAM (Aguilar et al. 2022) | 0.083 m 2.03% | 0.089 m 2.00% | 0.128 m 2.24% | 0.077 m 2.18% | 0.048 m 0.62% | 0.144 m 2.61% | 0.139 m 2.86% | 0.143 m 2.83% | 0.147 m 2.57% | 0.086 m 1.08% |
| | 0.147 °/m | 0.166 °/m | 0.182 °/m | 0.148 °/m | 0.283 °/m | 0.201 °/m | 0.242 °/m | 0.251 °/m | 0.153 °/m | 0.369 °/m |
| KISS-ICP (Vizzo et al. 2023) | 0.052 m 1.41% | 0.047 m <u>1.38%</u> | 0.045 m 1.50% | 0.045 m 1.56% | 0.041 m 0.91% | 0.036 m 1.07% | 0.038 m 1.10% | 0.032 m 1.00% | 0.035 m <u>1.12%</u> | 0.030 m 0.64% |
| | 0.093 °/m | 0.091 °/m | 0.096 °/m | 0.088 °/m | 0.149 °/m | 0.073 °/m | 0.079 °/m | 0.063 °/m | 0.069 °/m | 0.087 °/m |
| GLIM (Koide et al. 2024) | 0.525 m 2.99% | 0.286 m 2.73% | 1.563 m 18.69% | 0.082 m 2.38% | 0.068 m 0.82% | 0.029 m 1.15% | <u>0.021 m</u> 1.18% | 0.034 m 1.17% | 0.036 m 1.23% | 0.053 m 0.34% |
| | 0.048 °/m | 0.046 °/m | 0.083 °/m | 0.036 °/m | <u>0.040 °/m</u> | 0.066 °/m | 0.075 °/m | 0.058 °/m | 0.074 °/m | 0.050 °/m |
| Kinematic-ICP (Guadagnino et al. 2024) | 0.183 m 2.43% | 0.180 m 2.45% | 0.147 m 1.96% | 0.676 m 8.63% | — | 0.510 m 3.18% | 0.191 m 2.59% | 0.152 m 1.95% | 0.788 m 9.94% | — |
| | 0.140 °/m | 0.152 °/m | 0.137 °/m | 0.093 °/m | | 0.224 °/m | 0.172 °/m | 0.137 °/m | 0.135 °/m | |
| MOLA-LO (Blanco-Claraco 2025) | 0.045 m 1.41% | 0.040 m 1.48% | 0.037 m 1.69% | 0.039 m 1.54% | 0.338 m 3.53% | <u>0.028 m</u> 0.97% | 0.026 m 1.07% | <u>0.020 m</u> 1.06% | <u>0.028 m</u> 1.12% | <u>0.015 m</u> 0.24% |
| | 0.088 °/m | 0.088 °/m | 0.106 °/m | 0.092 °/m | 0.822 °/m | 0.070 °/m | 0.080 °/m | 0.057 °/m | 0.064 °/m | 0.054 °/m |
| LeGO-LOAM-BOR + LC ¹¹ | 0.059 m 1.33% | 0.047 m <u>1.37%</u> | 0.051 m 1.67% | 0.053 m 1.81% | 0.046 m 0.73% | 0.046 m 1.17% | 0.042 m 1.24% | 0.042 m 1.71% | 0.045 m 1.43% | 0.031 m 0.46% |
| | 0.104 °/m | 0.108 °/m | 0.115 °/m | 0.134 °/m | 0.183 °/m | 0.098 °/m | 0.108 °/m | 0.088 °/m | 0.105 °/m | 0.148 °/m |
| LIORF + LC ¹² | <u>0.028 m</u> 1.33% | <u>0.027 m</u> 1.45% | <u>0.019 m</u> 1.50% | <u>0.040 m</u> 1.42% | <u>0.027 m</u> 0.34% | 0.031 m 1.07% | 0.031 m 1.25% | <u>0.020 m</u> 1.33% | <u>0.028 m</u> 1.21% | <u>0.016 m</u> 0.20% |
| | 0.055 °/m | 0.068 °/m | 0.076 °/m | 0.097 °/m | 0.073 °/m | <u>0.061 °/m</u> | <u>0.069 °/m</u> | <u>0.054 °/m</u> | <u>0.056 °/m</u> | <u>0.026 °/m</u> |
| GLIM + LC (Koide et al. 2024) | 0.461 m 2.91% | 0.242 m 2.69% | 1.262 m 15.76% | 0.109 m 2.74% | 0.069 m 0.86% | <u>0.030 m</u> 1.13% | <u>0.021 m</u> 1.20% | 0.030 m 1.20% | 0.035 m 1.23% | 0.053 m 0.34% |
| | 0.046 °/m | 0.048 °/m | 0.077 °/m | 0.035 °/m | 0.041 °/m | 0.066 °/m | 0.075 °/m | 0.058 °/m | 0.073 °/m | 0.050 °/m |

Abbreviations: Loop Closure (LC)

Additionally, it is important to note that the odometry estimates produced by the SLAM algorithms are not always provided in the same reference frame as the ground-truth data (i.e., the *base_link* frame). For example, the output odometry may be provided in different reference frames: some algorithms report it in the LiDAR frame (e.g., A-LOAM, and LIORF), others in the IMU frame (e.g., GLIM), some in the robot footprint frame (e.g., Kinematic-ICP), and yet others directly in the robot frame (e.g., DLIO, KISS-ICP, and MOLA-LO). In order to ensure a fair comparison against the ground-truth data, the odometry trajectories were transformed to the robot frame by applying the appropriate CAD-specified transformations. These are the same transformations provided as static transforms in the ROS tf topic of the dataset and as configuration parameters for the algorithms when required, thereby reinforcing consistency in the analysis.

4.2 Results

The benchmark results of this study are presented in Tab. 5, Tab. 4 and Fig. 15, considering all six sequences, four 3D LiDAR sensors, and nine SLAM algorithms. The benchmark analysis is approached differently since ground-truth data is unavailable for the entire trajectory in the *elevator* sequence.

A quantitative analysis is conducted based on accuracy metrics for five of the sequences. In contrast, for the *elevator* sequence, a qualitative analysis is performed by examining the odometry trajectories produced by the SLAM algorithms.

The quantitative analysis is organized into two tables that report the ATE, RTE, and RRE for all four 3D LiDAR sensors. Tab. 5 presents the performance metrics obtained from sequences recorded with the RoboSense RS-Helios-5515 and Livox Mid-360 sensors, whereas Tab. 4 details similar metrics for data collected with the Velodyne VLP-16 and Ouster OS1-64 sensors. The RTE and RRE values are computed in both tables over 10-meter segments. Due to constraints inherent to some algorithms, not all combinations of sequences, sensors, and methods are available. For example, results for sequences acquired with the Livox Mid-360 are provided only for those algorithms that support this sensor, and quantitative outcomes for the *ramp* sequence are omitted for Kinematic-ICP, which is tailored for planar motion estimation. Nevertheless, qualitative results for this algorithm in the *elevator* sequence are included for illustrative purposes. Complementing the numerical data, Fig. 15 displays the odometry trajectories for the *elevator* sequence across all four sensor types. Due to the limited availability of ground-truth data provided only at the

Table 5. Benchmark results of 3D LiDAR-based SLAM algorithms using sequences from Livox Mid-360 and RoboSense RS-Helios-5515 sensors. Metrics include Absolute Trajectory Error (ATE) in meters (m), Relative Translational Error (RTE) in percentage (%), and Relative Rotational Error (RRE) in degrees per meter (/m).

| SLAM Algorithm | RoboSense RS-Helios-5515 Sequences | | | | | Livox Mid-360 Sequences | | | | |
|--|------------------------------------|------------------|------------------|------------------|------------------|-------------------------|------------------|------------------|------------------|------------------|
| | Nav A Diff | Nav A Omni | Loop | Slippage | Ramp | Nav A Diff | Nav A Omni | Loop | Slippage | Ramp |
| A-LOAM <i>Qin and Cao (2019)</i> | 0.032 m 0.92% | 0.032 m 1.05% | 0.045 m 1.03% | 0.042 m 1.32% | 0.057 m 0.59% | — | — | — | — | — |
| | 0.097 °/m | 0.084 °/m | 0.100 °/m | 0.084 °/m | 0.073 °/m | | | | | |
| LeGO-LOAM-BOR ¹¹ | 0.040 m 1.45% | 0.038 m 1.29% | 0.039 m 1.37% | 0.043 m 1.71% | 0.041 m 0.59% | — | — | — | — | — |
| | 0.114 °/m | 0.104 °/m | 0.132 °/m | 0.114 °/m | 0.168 °/m | | | | | |
| LIORF ¹² | 0.029 m 1.23% | 0.024 m 1.24% | 0.023 m 1.28% | 0.029 m 0.97% | 0.034 m 0.47% | — | — | — | — | — |
| | 0.090 °/m | 0.109 °/m | 0.093 °/m | 0.077 °/m | 0.119 °/m | | | | | |
| DLIO <i>Chen et al. (2023)</i> | 0.068 m 1.55% | 0.050 m 1.64% | 0.040 m 1.36% | 0.074 m 1.31% | 0.030 m 0.30% | 0.027 m 1.01% | 0.021 m 0.99% | 0.017 m 1.14% | 0.030 m 0.92% | 0.012 m 0.17% |
| | 0.107 °/m | 0.122 °/m | 0.102 °/m | 0.073 °/m | 0.125 °/m | 0.056 °/m | 0.069 °/m | 0.050 °/m | 0.042 °/m | 0.073 °/m |
| VineSLAM <i>Aguilar et al. (2022)</i> | 0.108 m 2.00% | 0.078 m 1.97% | 0.123 m 2.55% | 0.095 m 1.67% | 0.056 m 0.80% | 0.188 m 2.70% | 0.106 m 1.95% | 0.124 m 2.07% | 0.096 m 1.56% | 0.108 m 1.12% |
| | 0.148 °/m | 0.177 °/m | 0.199 °/m | 0.111 °/m | 0.315 °/m | 0.223 °/m | 0.188 °/m | 0.146 °/m | 0.072 °/m | 0.335 °/m |
| KISS-ICP <i>Vizzo et al. (2023)</i> | 0.046 m 1.10% | 0.046 m 1.10% | 0.047 m 1.00% | 0.041 m 0.96% | 0.039 m 0.58% | 0.031 m 0.76% | 0.030 m 0.75% | 0.026 m 0.83% | 0.033 m 0.95% | 0.028 m 0.53% |
| | 0.091 °/m | 0.080 °/m | 0.098 °/m | 0.081 °/m | 0.125 °/m | 0.064 °/m | 0.067 °/m | 0.059 °/m | 0.052 °/m | 0.099 °/m |
| GLIM <i>Koide et al. (2024)</i> | 0.137 m 1.60% | 0.055 m 1.16% | 0.560 m 7.36% | 0.030 m 1.43% | 0.057 m 0.65% | 0.017 m 0.78% | 0.016 m 0.81% | 0.104 m 1.76% | 0.025 m 0.82% | 0.057 m 0.64% |
| | 0.051 °/m | 0.038 °/m | 0.049 °/m | 0.038 °/m | 0.036 °/m | 0.044 °/m | 0.052 °/m | 0.036 °/m | 0.039 °/m | 0.055 °/m |
| Kinematic-ICP <i>Guadagnino et al. (2024)</i> | 0.337 m 2.87% | 0.178 m 2.27% | 0.103 m 1.62% | 0.676 m 8.86% | — | 0.085 m 1.62% | 0.181 m 2.31% | 0.053 m 1.38% | 0.584 m 7.45% | — |
| | 0.184 °/m | 0.143 °/m | 0.118 °/m | 0.110 °/m | | 0.066 °/m | 0.108 °/m | 0.084 °/m | 0.106 °/m | |
| MOLA-LO <i>Blanco-Claraco (2025)</i> | 0.044 m 1.26% | 0.031 m 1.15% | 0.032 m 1.25% | 0.040 m 1.21% | 0.134 m 0.94% | 0.025 m 0.77% | 0.022 m 0.77% | 0.018 m 0.98% | 0.027 m 0.93% | 0.023 m 0.26% |
| | 0.104 °/m | 0.076 °/m | 0.104 °/m | 0.099 °/m | 0.230 °/m | 0.057 °/m | 0.063 °/m | 0.058 °/m | 0.047 °/m | 0.056 °/m |
| LeGO-LOAM-BOR + LC ¹¹ | 0.045 m 1.44% | 0.043 m 1.37% | 0.050 m 1.81% | 0.043 m 1.69% | 0.041 m 0.56% | — | — | — | — | — |
| | 0.113 °/m | 0.099 °/m | 0.133 °/m | 0.112 °/m | 0.168 °/m | | | | | |
| LIORF + LC ¹² | 0.029 m 1.22% | 0.023 m 1.21% | 0.022 m 1.27% | 0.031 m 1.00% | 0.022 m 0.33% | — | — | — | — | — |
| | 0.093 °/m | 0.104 °/m | 0.091 °/m | 0.078 °/m | 0.110 °/m | | | | | |
| GLIM + LC <i>Koide et al. (2024)</i> | 0.148 m 1.76% | 0.059 m 1.27% | 0.487 m 6.66% | 0.030 m 1.41% | 0.057 m 0.64% | 0.016 m 0.78% | 0.016 m 0.81% | 0.103 m 1.74% | 0.025 m 0.82% | 0.057 m 0.64% |
| | 0.051 °/m | 0.038 °/m | 0.048 °/m | 0.037 °/m | 0.036 °/m | 0.044 °/m | 0.051 °/m | 0.036 °/m | 0.039 °/m | 0.056 °/m |

Abbreviations: Loop Closure (LC)

beginning and end of the trajectory, a quantitative evaluation is not feasible for this sequence, and a qualitative assessment is instead based on a visual analysis of the trajectories.

The results indicate that most algorithms maintain robust performance without significant divergences. A closer examination reveals that methods incorporating IMU sensor fusion generally outperform those that do not. However, algorithms relying on wheel odometry tend to exhibit more significant errors, with some exceptions. Among the evaluated approaches, LIORF consistently demonstrates the best overall performance, GLIM exhibits notably lower rotational drift, and KISS-ICP registers reduced translational drift. In the *elevator* sequence, only a few algorithms (LIORF, DLIO, GLIM, and MOLA-LO) could accurately estimate the trajectory. Moreover, the choice of sensor has influence on performance, with the Livox Mid-360 and Ouster OS1-64 sensors enabling more accurate trajectory estimation compared to the RoboSense RS-Helios-5515 and Velodyne VLP-16. Finally, the integration of loop closure modules appears to have a marginal impact, as the performance differences between the original implementations and their loop closure-enhanced counterparts are minimal. These observations provide a solid

foundation for the subsequent discussion of the underlying trends and algorithmic differences.

Due to the extensive scope of this benchmark study, which includes six sequences, four 3D LiDAR sensors, and nine SLAM algorithms, it is not practical to include illustrations of all resulting trajectories within this paper. The figures presented here, such as Fig. 15, Fig. 16, and Fig. 17, represent just a selected subset, that highlights key trends and specific cases. For a complete set of trajectory visualizations, please refer to the project webpage¹⁴.

4.3 Discussion

As outlined in the previous subsection, all algorithms generally achieved robust performance across the benchmark sequences, yielding an ATE between 0.02 m and 0.05 m, an RTE around 1%, and an RRE below 0.2°/m, which is consistent with the expectations for a feature-rich environment such as the iilab. However, a few notable exceptions merit closer inspection. For example, the GLIM algorithm exhibited significant difficulties with the Velodyne VLP-16 sequences. The sensor's narrow vertical FoV, from -15° to 15°, appears to have induced drift along the z-axis, most likely since it could not provide enough points of the ground and/or ceiling. Moreover, in the *loop* sequence,

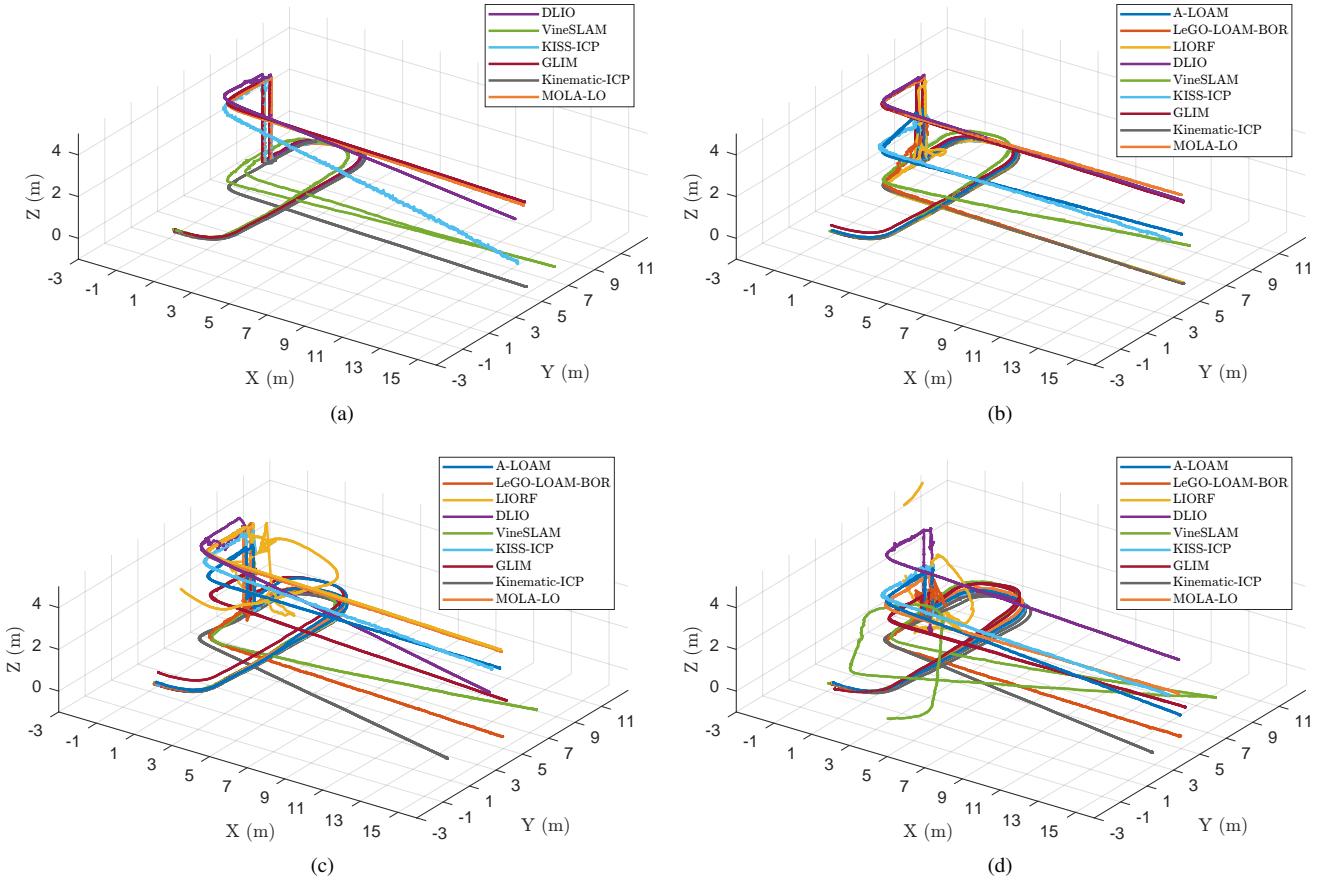


Figure 15. Odometry trajectories of 3D LiDAR-based SLAM algorithms in the *elevator* sequence for different 3D LiDAR sensors: (a) Livox Mid-360; (b) Ouster OS1-64; (c) RoboSense RS-Helios-5515; (d) Velodyne VLP-16.

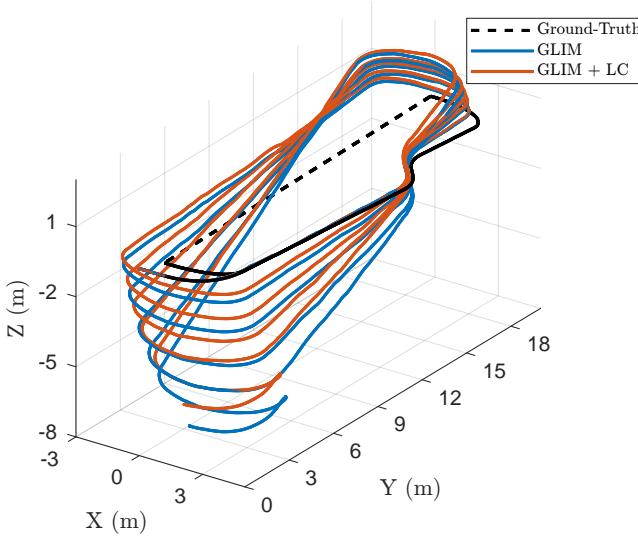


Figure 16. Odometry trajectories of the GLIM algorithm in the *loop* sequence using the Velodyne VLP-16 LiDAR sensor, with and without loop-closure module

where the robot transitions from the Nav A space to the corridor through a glass door, GLIM struggled with discrepancies between consecutive keyframe local maps in those transitions. These issues culminated in an ATE of 1.563 m and an RTE of 18.69% for that sequence, as illustrated in Fig. 17.

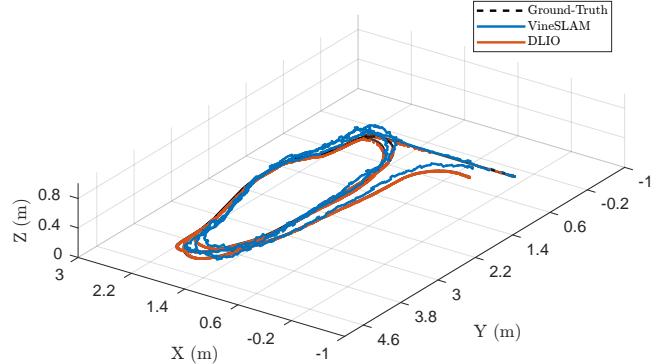


Figure 17. Odometry trajectories of the VineSLAM and DLIO algorithms in the *ramp* sequence using the Velodyne VLP-16 LiDAR sensor

Moreover, algorithms incorporating IMU sensor fusion generally delivered superior accuracy. Among these, LIORF consistently demonstrated the best overall performance by achieving the lowest ATE, RTE, and RRE in several sequences. However, when the metrics are considered individually, the rankings vary: while LIORF excels in ATE, KISS-ICP and GLIM register the best performance in terms of RTE and RRE, respectively. One contributing factor to the enhanced performance of IMU-based methods is the use of the high-precision Xsens MTi-630 AHRS IMU, which operates at 400 Hz and provides accurate motion estimates

that improve the initialization step of the ICP-based modules within these SLAM systems.

Conversely, algorithms that rely on wheel odometry data, such as Kinematic-ICP and VineSLAM, produced less accurate results. For instance, Kinematic-ICP consistently achieved an ATE exceeding 0.1 m in most sequences and even surpassed 0.5 m in the *slippage* sequence. This underperformance is likely related to its unicycle-based correction in the ICP-based pipeline, which is unsuitable for robots with omnidirectional steering. On the other hand, although the VineSLAM achieved satisfactory results in all sequences, with an ATE below 0.15 m, its performance was outstripped by several other methods. Despite being the only approach that fuses both IMU and wheel odometry data for sensor fusion, VineSLAM relies on a particle filter that requires a lower update frequency for robot pose estimation. Furthermore, its overall trajectory exhibits more jitter than those generated by the LO-based algorithms, as illustrated in Fig. 16.

When evaluating the results on a per-sensor basis, the Livox Mid-360 consistently enabled the best outcomes across sequences. For example, in the *nav a diff* sequence, the optimal metrics achieved by GLIM and KISS-ICP were an ATE of 0.017 m, an RTE of 0.76%, and an RRE of 0.044°/m, outperforming results obtained with the other sensors. This improvement can be attributed to the Livox Mid-360's distinctive vertical FoV, which extends from -7° to 52°, thereby capturing abundant ceiling features that enhance the performance of the ICP-based local optimization modules. Notably, the VineSLAM algorithm did not exhibit a similar improvement with this sensor, reinforcing the notion that the advantages of the Livox Mid-360 are closely tied to the presence of a practical local optimization component. Additionally, specific algorithms appear to perform better with particular sensors, likely due to the extensive parameter sets of these algorithms, which can be more fine-tuned for specific sensors. For instance, GLIM excelled with the Livox Mid-360, while MOLA-LO delivered better results with the Ouster OS1-64.

Despite these particular cases, the minimal differences observed in algorithm performance among the mechanical spinning LiDARs (Velodyne VLP-16, RoboSense RS-Helios-5515, and Ouster OS1-64) suggest that the vertical resolution of these sensors, which is determined by the number of beams and vertical FoV, has limited impact in indoor environments. Similarly, the unique characteristics of the RoboSense RS-Helios-5515, such as its asymmetric and non-uniform FoV resolution, do not appear to affect performance in these settings significantly.

With respect to loop-closure modules, the results remain inconclusive. In most sequences, the algorithms equipped with a loop-closure module (LeGO-LOAM-BOR, LIORF, and GLIM) did not require loop-closure optimization to correct large trajectory drifts. The only exception was the GLIM algorithm, which relied on its loop-closure module to mitigate z-axis drift, mainly when using the Velodyne VLP-16 sensor, as illustrated in Fig. 16. As such, a direct comparison between the different loop-closure modules is not possible.

Finally, regarding the *elevator* sequence, only a small subset of the algorithms could accurately estimate vertical

motion during both ascent and descent. Among the nine SLAM methods, three are explicitly designed for ground applications (LeGO-LOAM-BOR, VineSLAM, and Kinematic-ICP), and, as expected, these approaches failed to capture vertical movements. In contrast, three algorithms that incorporate IMU data (LIORF, DLIO, and GLIM) managed to estimate vertical motion accurately with certain sensors, notably the Ouster OS1-64 and Livox Mid 360. At the same time, performance with the Velodyne VLP-16 and RoboSense RS-Helios-5515 was less reliable, likely due to the lower point-cloud density provided by these sensors. Among the IMU-based methods, LIORF successfully tracked vertical motion during ascent for three sensors (Ouster OS1-64, Livox Mid-360, and RoboSense RS-Helios-5515) but produced inaccurate estimates during descent for two of them (Ouster OS1-64 and RoboSense RS-Helios-5515), and it completely diverged for the Velodyne VLP-16. Additionally, of the three algorithms that rely solely on LiDAR data (A-LOAM, KISS-ICP, and MOLA-LO), both A-LOAM and KISS-ICP failed to capture the full vertical displacement by exhibiting an offset in the vertical distance to the second floor. In contrast, MOLA-LO surprisingly managed to estimate vertical motion accurately for three sensors (Ouster OS1-64, Livox Mid-360, and RoboSense RS-Helios-5515), albeit with an offset error for the Velodyne VLP-16, likely due to insufficient point-cloud information from that sensor.

5 Conclusion

In this study, we presented a comprehensive benchmark of state-of-the-art 3D LiDAR-based SLAM algorithms evaluated on the IILABS 3D dataset. This dataset was developed to address key limitations of existing resources by providing multiple sensor sources along with calibration sequences and challenging benchmark trajectories, all acquired using a mobile robot at the iilab. The sensor sources include four 3D LiDARs (Velodyne VLP-16, Ouster OS1-64, RoboSense RS-Helios-5515) with different FoVs and scanning patterns, an IMU (Xsens MTi-630 AHRS), and wheel odometry.

By testing nine algorithms over six sequences and four distinct LiDAR sensors, our study offers a comparison of their performance in terms of accuracy and robustness under diverse indoor conditions. Overall, the results indicate that most SLAM methods perform robustly, with algorithms incorporating IMU sensor fusion consistently delivering superior accuracy, whereas approaches that rely on wheel odometry tend to yield larger errors. As such, given that wheel odometry is more accessible than IMU sensors in many industrial environments, there exists a clear research gap in developing robust SLAM methods that effectively leverage wheel odometry. Moreover, the analysis revealed that sensor characteristics play a critical role: the Livox Mid-360 and Ouster OS1-64 sensors, with their dense point clouds and favorable FoV, enabled more precise trajectory estimation compared to the RoboSense RS-Helios-5515 and Velodyne VLP-16. Additionally, the challenging task of accurately capturing vertical motion in the *elevator* sequence further highlighted the limitations of current SLAM approaches.

Additionally, it is important to note that among the nine evaluated algorithms, only VineSLAM and MOLA-LO provide both mapping and dedicated localization modules. While all the algorithms are capable of generating maps, only these two offer integrated localization functionalities that use the previously built map to accurately determine the robot's pose. This gap in the state-of-the-art highlights a promising direction for future research aimed at developing SLAM solutions that seamlessly combine robust mapping with efficient localization in challenging indoor environments. Furthermore, we provide open-source benchmark scripts used to evaluate all nine SLAM algorithms presented in this work, along with a dedicated toolkit for downloading the IILABS 3D dataset and computing evaluation metrics. These resources not only facilitate the replication of this benchmark study but also support further research in indoor LiDAR-based SLAM.

Acknowledgements

This work is co-financed by Component 5 – Capitalization and Business Innovation, integrated in the Resilience Dimension of the Recovery and Resilience Plan within the scope of the Recovery and Resilience Mechanism (MRR) of the European Union (EU), framed in the Next Generation EU, for the period 2021–2026, within project GreenAuto, with reference 54.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

- 1 IILABS 3D Dataset: <https://doi.org/10.25747/vhnj-wm80>
- 2 Project GitHub: https://github.com/jorgedfr/3d_lidar_slam_benchmark_at_iilab
- 3 IILABS 3D Toolkit: <https://github.com/jorgedfr/iilabs3d-toolkit>
- 4 https://sousarbarb.github.io/inesctec_mrdt_hangfa_discovery_q2
- 5 <https://www.hangfa-europe.com>
- 6 <https://5dpo.github.io>
- 7 https://github.com/5dpo/5dpo_ratf_2023
- 8 <https://www.optitrack.com>
- 9 <https://optitrack.com/software/motive>
- 10 https://github.com/L2S-lab/natnet_ros_cpp
- 11 <https://github.com/facontidavide/LeGO-LOAM-BOR>
- 12 <https://github.com/YJZLuckyBoy/liorf>
- 13 <https://www.docker.com>
- 14 https://jorgedfr.github.io/3d_lidar_slam_benchmark_at_iilab

References

- Aguiar A, Santos F, Sobreira H, Boaventura-Cunha J and Sousa A (2022) Localization and mapping on agriculture based on point-feature extraction and semiplanes segmentation from 3D LiDAR data. *Frontiers in Robotics and AI* 9. DOI:10.3389/frobt.2022.832165.
- Bailey T and Durrant-Whyte H (2006) Simultaneous localization and mapping (SLAM): part II. *IEEE Robotics & Automation Magazine* 13(3): 108–117. DOI:10.1109/MRA.2006.1678144.
- Besl P and McKay N (1992) A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2): 239–256. DOI:10.1109/34.121791.
- Blanco-Claraco J (2025) A flexible framework for accurate LiDAR odometry, map manipulation, and localization. *The International Journal of Robotics Research* DOI:10.1177/02783649251316881.
- Bodin B, Wagstaff H, Saecdi S, Nardi L, Vespa E, Mawer J, Nisbet A, Lujan M, Furber S, Davison A, Kelly P and O'Boyle M (2018) SLAMBench2: multi-objective head-to-head benchmarking for visual SLAM. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 3637–3644. DOI:10.1109/ICRA.2018.8460558.
- Bresson G, Alsayed Z, Yu L and Glaser S (2017) Simultaneous localization and mapping: a survey of current trends in autonomous driving. *IEEE Transactions on Intelligent Vehicles* 2(3): 194–220. DOI:10.1109/TIV.2017.2749181.
- Brizi L, Giacomini E, Giamarino L, Ferrari S, Salem O, Rebotti L and Grisetti G (2024) VBR: a vision benchmark in Rome. In: *2024 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 15868–15874. DOI:10.1109/ICRA57147.2024.10611395.
- Buchanan R (2021) Allan Variance ROS. URL https://github.com/ori-drs/allan_variance_ros.
- Bujanca M, Gafton P, Saeedi S, Nisbet A, Bodin B, O'Boyle M, Davison A, Kelly P, Riley G, Lennox B, Luján M and Furber S (2019) SLAMBench 3.0: systematic automated reproducible evaluation of SLAM systems for robot vision challenges and scene understanding. In: *2019 International Conference on Robotics and Automation (ICRA)*. pp. 6351–6358. DOI:10.1109/ICRA.2019.8794369.
- Burri M, Nikolic J, Gohl P, Schneider T, Rehder J, Omari S, Achtelik M and Siegwart R (2016) The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research* 35(10): 1157–1163. DOI:10.1177/0278364915620033.
- Cadena C, Carlone L, Carrillo H, Latif Y, Scaramuzza D, Neira J, Reid I and Leonard J (2016) Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. *IEEE Transactions on Robotics* 32(6): 1309–1332. DOI:10.1109/TRO.2016.2624754.
- Carlevaris-Bianco N, Ushani A and Eustice R (2015) University of Michigan North Campus long-term vision and LiDAR dataset. *International Journal of Robotics Research* 35(9): 1023–1035.
- Catalano I, Yu X and Queraltà J (2023) Towards robust UAV tracking in GNSS-denied environments: a multi-LiDAR multi-UAV dataset. In: *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. pp. 1–7.
- Chen K, Nemiroff R and Lopez B (2023) Direct LiDAR-inertial odometry: lightweight LIO with continuous-time motion correction. *2023 IEEE International Conference on Robotics and Automation (ICRA)* : 3983–3989 DOI:10.1109/ICRA48891.2023.10160508.
- Chen Z, Qi Y, Feng D, Zhuang X, Chen H, Hu X, Wu J, Peng K and Lu P (2024) Heterogeneous LiDAR dataset for benchmarking

- robust localization in diverse degenerate scenarios. *arXiv Preprint arXiv:2409.04961*. URL <https://arxiv.org/abs/2409.04961>.
- Durrant-Whyte H and Bailey T (2006) Simultaneous localization and mapping: part I. *IEEE Robotics & Automation Magazine* 13(2): 99–110. DOI:10.1109/MRA.2006.1638022.
- Feng D, Qi Y, Zhong S, Chen Z, Chen Q, Chen H, Wu J and Ma J (2024) S3E: a multi-robot multimodal dataset for collaborative SLAM. *IEEE Robotics and Automation Letters* : 1–8DOI: 10.1109/LRA.2024.3490402.
- Ferrari S, Giammarino L, Brizi L and Grisetti G (2024) MAD-ICP: it is all about matching data – robust and informed LiDAR odometry. *IEEE Robotics and Automation Letters* 9(11): 9175–9182. DOI:10.1109/LRA.2024.3456509.
- Geiger A, Lenz P, Stiller C and Urtasun R (2013) Vision meets robotics: the KITTI dataset. *The International Journal of Robotics Research* 32(11): 1231–1237. DOI:10.1177/0278364913491297.
- Geiger A, Lenz P and Urtasun R (2012) Are we ready for autonomous driving? The KITTI vision benchmark suite. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3354–3361. DOI:10.1109/CVPR.2012.6248074.
- Grisetti G, Kümmerle R, Stachniss C and Burgard W (2010) A tutorial on graph-based SLAM. *IEEE Intelligent Transportation Systems Magazine* 2(4): 31–43. DOI:10.1109/MITS.2010.939925.
- Grisetti G, Stachniss C and Burgard W (2007) Improved techniques for grid mapping with Rao-Blackwellized particle filters. *IEEE Transactions on Robotics* 23(1): 34–46. DOI:10.1109/TRO.2006.889486.
- Grupp M (2017) evo: Python package for the evaluation of odometry and SLAM. URL <https://github.com/MichaelGrupp/evo>.
- Guadagnino T, Mersch B, Vizzo I, Gupta S, Malladi M, Lobefaro L, Doisy G and Stachniss C (2024) Kinematic-ICP: enhancing LiDAR odometry with kinematic constraints for wheeled mobile robots moving on planar surfaces. *arXiv Preprint arXiv:2410.10277*. URL <https://arxiv.org/pdf/2410.10277.pdf>.
- Helmberger M, Morin K, Berner B, Kumar N, Cioffi G and Scaramuzza D (2022) The Hilti SLAM challenge dataset. *IEEE Robotics and Automation Letters* 7(3): 7518–7525. DOI:10.1109/LRA.2022.3183759.
- Hess W, Kohler D, Rapp H and Andor D (2016) Real-time loop closure in 2D LiDAR SLAM. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 1271–1278. DOI:10.1109/ICRA.2016.7487258.
- Jiao J, Wei H, Hu T, Hu X, Zhu Y, He Z, Wu J, Yu J, Xie X, Huang H et al. (2022) FusionPortable: A multi-sensor campus-scene dataset for evaluation of localization and mapping accuracy on diverse platforms. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 3851–3856. DOI:10.1109/IROS47612.2022.9982119.
- Kim G, Choi S and Kim A (2022) Scan context++: structural place recognition robust to rotation and lateral variations in urban environments. *IEEE Transactions on Robotics* 38(3): 1856–1874. DOI:10.1109/TRO.2021.3116424.
- Kim G and Kim A (2018) Scan context: egocentric spatial descriptor for place recognition within 3D point cloud map. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 4802–4809. DOI:10.1109/IROS.2018.8593953.
- Kim G, Park Y, Cho Y, Jeong J and Kim A (2020) MulRan: multimodal range dataset for urban place recognition. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 6246–6253. DOI:10.1109/ICRA40945.2020.9197298.
- Kim T, Pak G and Kim E (2024) GRIL-Calib: targetless ground robot IMU-LiDAR extrinsic calibration method using ground plane motion constraints. *IEEE Robotics and Automation Letters* 9(6): 5409–5416. DOI:10.1109/LRA.2024.3392081.
- Kohlbrecher S, von Stryk O, Meyer J and Klingauf U (2011) A flexible and scalable SLAM system with full 3D motion estimation. In: *2011 IEEE International Symposium on Safety, Security, and Rescue Robotics*. pp. 155–160. DOI:10.1109/SSRR.2011.6106777.
- Koide K, Miura J and Menegatti E (2019) A portable three-dimensional LiDAR-based system for long-term and wide-area people behavior measurement. *International Journal of Advanced Robotic Systems* 16(2). DOI:10.1177/1729881419841532.
- Koide K, Yokozuka M, Oishi S and Banno A (2024) GLIM: 3D range-inertial localization and mapping with GPU-accelerated scan matching factors. *Robotics and Autonomous Systems* 179. DOI:10.1016/j.robot.2024.104750.
- Korayem M, Nakhai A and Rostam T (2006) Design, modelling and errors measurement of wheeled mobile robots. *The International Journal of Advanced Manufacturing Technology* 28: 403–416. DOI:10.1007/s00170-004-2357-3.
- Liao Y, Xie J and Geiger A (2023) KITTI-360: a novel dataset and benchmarks for urban scene understanding in 2D and 3D. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45(3): 3292–3310. DOI:10.1109/TPAMI.2022.3179507.
- Lin J and Zhang F (2020) LOAM Livox: a fast, robust, high-precision LiDAR odometry and mapping package for LiDARs of small FoV. In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 3126–3131. DOI:10.1109/ICRA40945.2020.9197440.
- Liu X, Yang Y, Xu B, Feng D and Schwertfeger S (2024) Benchmarking SLAM algorithms in the cloud: the SLAM Hive benchmarking suite. *arXiv Preprint arXiv:2406.17586*. URL <https://arxiv.org/abs/2406.17586>.
- Macenski S and Jambrecic I (2021) SLAM toolbox: SLAM for the dynamic world. *Journal of Open Source Software* 6(61). DOI: 10.21105/joss.02783.
- Maddern W, Pascoe G, Linegar C and Newman P (2017) 1 year, 1000km: the Oxford RobotCar dataset. *The International Journal of Robotics Research (IJRR)* 36(1): 3–15. DOI:10.1177/0278364916679498.
- Nardi L, Bodin B, Zia M, Mawer J, Nisbet A, Kelly P, Davison A, Luján M, O’Boyle M, Riley G, Topham N and Furber S (2015) Introducing SLAMBench, a performance and accuracy benchmarking methodology for SLAM. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 5783–5790. DOI:10.1109/ICRA.2015.7140009.
- Nguyen TM, Yuan S, Cao M, Lyu Y, Nguyen T and Xie L (2022) NTU VIRAL: a visual-inertial-ranging-LiDAR dataset, from an aerial vehicle viewpoint. *The International Journal of Robotics Research* 41(3): 270–280. DOI:10.1177/0278364922109321.

- 02783649211052312.
- Pan Y, Zhong X, Wiesmann L, Posewsky T, Behley J and Stachniss C (2024) PIN-SLAM: LiDAR SLAM using a point-based implicit neural representation for achieving global map consistency. *IEEE Transactions on Robotics* 40: 4045–4064. DOI:10.1109/TRO.2024.3422055.
- Qin T and Cao S (2019) A-LOAM: advanced implementation of LOAM. URL <https://github.com/HKUST-Aerial-Robotics/A-LOAM>.
- Qingqing L, Xianjia Y, Queralta J and Westerlund T (2022) Multi-modal LiDAR dataset for benchmarking general-purpose localization and mapping algorithms. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 3837–3844. DOI:10.1109/IROS47612.2022.9981078.
- Ribeiro J, Sousa R, Martins J, Aguiar A, Santos F and Sobreira H (2025) IILABS 3D: iilab indoor LiDAR-based SLAM dataset. DOI:10.25747/VHNJ-WM80. [Dataset].
- Roriz R, Cabral J and Gomes T (2022) Automotive LiDAR technology: a survey. *IEEE Transactions on Intelligent Transportation Systems* 23(7): 6282–6297. DOI:10.1109/TITS.2021.3086804.
- Schöps T, Schönberger J, Galliani S, Sattler T, Schindler K, Pollefeys M and Geiger A (2017) A multi-view stereo benchmark with high-resolution images and multi-camera videos. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 2538–2547. DOI:10.1109/CVPR.2017.272.
- Shan T and Englot B (2018) LeGO-LOAM: lightweight and ground-optimized Lidar odometry and mapping on variable terrain. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 4758–4765. DOI:10.1109/IROS.2018.8594299.
- Shan T, Englot B, Meyers D, Wang W, Ratti C and Rus D (2020) LIO-SAM: tightly-coupled LiDAR inertial odometry via smoothing and mapping. In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 5135–5142. DOI:10.1109/IROS45743.2020.9341176.
- Siegwart R, Nourbakhsh I and Scaramuzza D (2011) *Introduction to autonomous mobile robots*, volume 23. MIT press.
- Sobreira H, Pinto M, Moreira A, Costa P and Lima J (2015) Robust robot localization based on the perfect match algorithm. In: *CONTROLO'2014 – Proceedings of the 11th Portuguese Conference on Automatic Control*. pp. 607–616. DOI:10.1007/978-3-319-10380-8_58.
- Sousa R, Petry M, Costa P and Moreira A (2022) OptiOdom: a generic approach for odometry calibration of wheeled mobile robots. *Journal of Intelligent & Robotic Systems* 105(2). DOI:10.1007/s10846-022-01630-3.
- Sousa R, Rocha C, Martins J, Costa J, Padrão J, Sarmento J, Carvalho J, Lopes M, Costa P and Moreira A (2024) A robotic framework for the Robot@Factory 4.0 competition. In: *2024 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. pp. 66–73. DOI:10.1109/ICARSC61747.2024.10535935.
- Sousa R, Sobreira H, Martins J, Costa P, Silva M and Moreira A (2025) Integrating multimodal perception into ground mobile robots. In: *2025 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. pp. 104–111. DOI:10.1109/ICARSC65809.2025.10970176.
- Sousa R, Sobreira H and Moreira A (2023) A systematic literature review on long-term localization and mapping for mobile robots. *Journal of Field Robotics* 40(5): 1245–1322. DOI:10.1002/rob.22170.
- Sturm J, Engelhard N, Endres F, Burgard W and Cremers D (2012) A benchmark for the evaluation of RGB-D SLAM systems. In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. pp. 573–580. DOI:10.1109/IROS.2012.6385773.
- Umeyama S (1991) Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(4): 376–380. DOI:10.1109/34.88573.
- Vizzo I, Guadagnino T, Mersch B, Wiesmann L, Behley J and Stachniss C (2023) KISS-ICP: in defense of point-to-point ICP – simple, accurate, and robust registration if done the right way. *IEEE Robotics and Automation Letters* 8(2): 1029–1036. DOI:10.1109/LRA.2023.3236571.
- Wang H, Wang C, Chen CL and Xie L (2021) F-LOAM: fast LiDAR odometry and mapping. In: *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 4390–4396. DOI:10.1109/IROS51168.2021.9636655.
- Wei H, Jiao J, Hu X, Yu J, Xie X, Wu J, Zhu Y, Liu Y, Wang L and Liu M (2024) FusionPortableV2: a unified multi-sensor dataset for generalized SLAM across diverse platforms and scalable environments. *The International Journal of Robotics Research* DOI:10.1177/02783649241303525.
- Xie A, Qian Y, Yan W, Wang C and Yang M (2024) Non-repetitive: a promising LiDAR scanning pattern. In: *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 2653–2659. DOI:10.1109/IROS58592.2024.10802663.
- Xu W and Zhang F (2021) FAST-LIO: a fast, robust LiDAR-inertial odometry package by tightly-coupled iterated Kalman filter. *IEEE Robotics and Automation Letters* 6(2): 3317–3324. DOI:10.1109/LRA.2021.3064227.
- Xu X, Zhang L, Yang J, Cao C, Wang W, Ran Y, Tan Z and Luo M (2022) A review of multi-sensor fusion SLAM systems based on 3D LiDAR. *Remote Sensing* 14(12). DOI:10.3390/rs14122835.
- Yang Y, Xu B, Li Y and Schwertfeger S (2023) The SLAM Hive benchmarking suite. In: *2023 IEEE International Conference on Robotics and Automation (ICRA)*. pp. 11257–11263. DOI:10.1109/ICRA48891.2023.10160302.
- Yin J, Li A, Li T, Yu W and Zou D (2022) M2DGR: a multi-sensor and multi-scenario SLAM dataset for ground robots. *IEEE Robotics and Automation Letters* 7(2): 2266–2273. DOI:10.1109/LRA.2021.3138527.
- Yin J, Yin H, Liang C, Jiang H and Zhang Z (2023) Ground-challenge: a multi-sensor SLAM dataset focusing on corner cases for ground robots. In: *2023 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. pp. 1–5. DOI:10.1109/ROBIO58561.2023.10354969.
- Zhang J and Singh S (2014) LOAM: LiDAR odometry and mapping in real-time. In: *Robotics: Science and systems*. pp. 1–9.
- Zhang L, Helmberger M, Fu L, Wisth D, Camurri M, Scaramuzza D and Fallon M (2023) Hilti-Oxford dataset: a millimeter-accurate benchmark for simultaneous localization and mapping. *IEEE Robotics and Automation Letters* 8(1): 408–415. DOI:10.1109/LRA.2022.3226077.

- Zhu F, Ren Y and Zhang F (2022) Robust real-time LiDAR-inertial initialization. In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. pp. 3948–3955. DOI: 10.1109/IROS47612.2022.9982225.