2026 IEEE International Conference on Robotics and Automation (ICRA)
June 1-5, 2026. Vienna, Austria

2026 IEEE International Conference on Robotics and Automation (ICRA) | 979-8-3315-8160-2/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICRA57385.2026.11696237

TEMPO-VINE: A Multi-Temporal Sensor Fusion Dataset for Localization and Mapping in Vineyards

Mauro Martini[1], Marco Ambrosio[1], Judith Vilella-Cantos[2], Alessandro Navone[1], and Marcello Chiaberge[1]

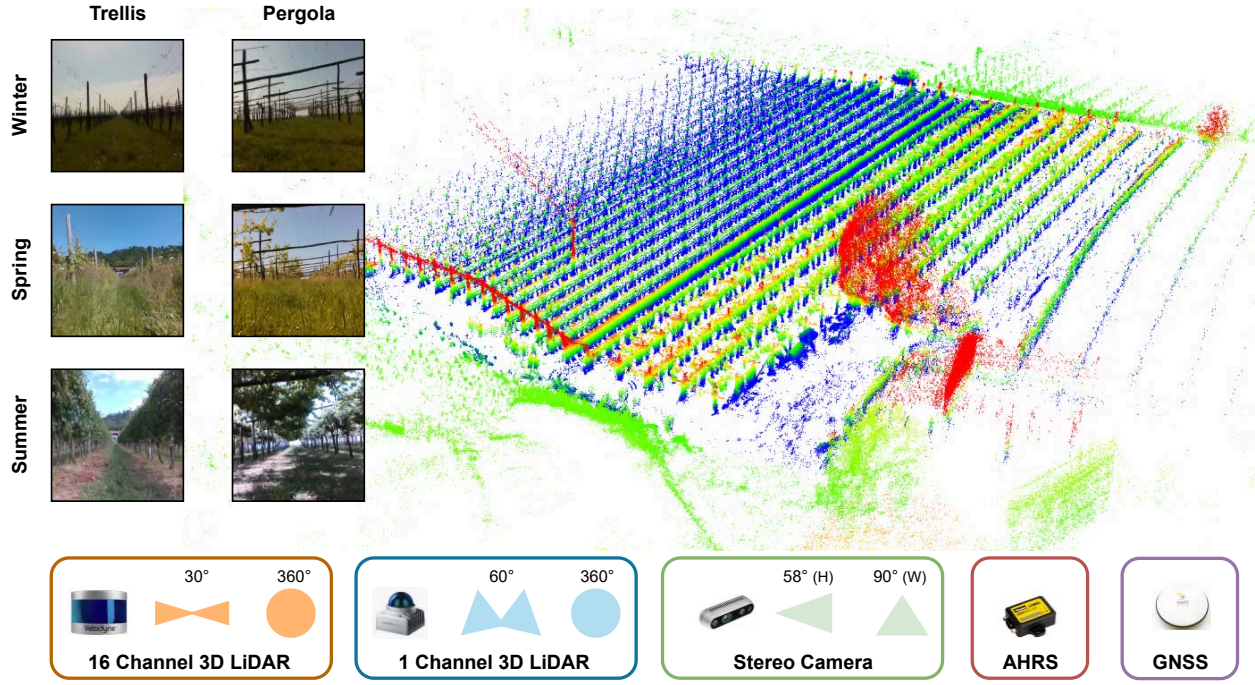

Fig. 1: TEMPO-VINE is a multi-seasonal dataset collected in trellis and pergola vineyards with different LiDARs, an RGB-D camera, an AHRS and a GNSS-RTK used for ground truth computation.

***Abstract*— In recent years, precision agriculture has been introducing groundbreaking innovations in the field, with a strong focus on automation. However, research studies in robotics and autonomous navigation often rely on controlled simulations or isolated field trials. The absence of a realistic common benchmark represents a significant limitation for the diffusion of robust autonomous systems under real complex agricultural conditions. Vineyards pose significant challenges due to their dynamic nature, and they are increasingly drawing attention from both academic and industrial stakeholders interested in automation. In this context, we introduce the TEMPO-VINE dataset, a large-scale multi-temporal dataset specifically designed for evaluating sensor fusion, simultaneous localization and mapping (SLAM), and place recognition techniques within operational vineyard environments. TEMPO-VINE is the first multi-modal public dataset that brings together data from heterogeneous LiDARs of different price levels, AHRS, RTK-GPS, and cameras in real trellis and pergola vineyards, with multiple rows exceeding 100 m in length. In this work, we address a critical gap in the landscape of agricultural datasets by providing researchers with a comprehensive data collection and ground truth trajectories in different seasons, vegetation growth stages, terrain and weather conditions. The sequence paths with multiple runs and revisits will foster the development of sensor fusion, localization, mapping and place recognition solutions for agricultural fields. The dataset, the processing tools and the benchmarking results are available on the webpage[1].**

I. INTRODUCTION

Precision agriculture is undergoing a technological transformation, with increasing integration of robotic systems and automation aimed at improving productivity, reducing labor costs, and enabling sustainable farming practices [7]. Recent progress in autonomous navigation [8], [9] and environmental perception [10] has demonstrated the potential of robotics in agricultural domain. Much of this research has focused on row-based crops, which represent roughly 75% of planted fields in the United States [11], addressing problems

[1] Department of Electronics and Communications, Politecnico di Torino, 10129, Torino, Italy. {name.surname}@polito.it
[2] University Institute for Engineering Research, Miguel Hernández University, Avda. de la Universidad s/n, Edificio Innova, Elche, 03202, Alicante, Spain. jvilella@umh.es
This work has been developed within the PoliTO Interdepartmental Centre for Service Robotics PIC4SeR.

[1] https://sites.google.com/view/tempo-vine-dataset/dataset-overview

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.

TABLE I
COMPARISON WITH EXISTING AGRICULTURAL DATASETS.

Dataset	Campaigns	Crop Type	Total length	3D LiDAR	Camera	GPS	AHRS/IMU	Odom
GREENBOT [1]	9 (2 months)	Tomato	10.4 km	2	1 stereo RGB	-	✓	-
MAgro [2]	9 (5 months)	Apple; Pear	3.06 km	1	2 stereo RGB	RTK	✓	✓
ARD-VO [3]	11 (3 months)	Grapevine, olive.	28.8 km	1	2 stereo RGB, 1 multispectral	RTK	✓	✓
ROSARIO [4]	6 (1 month)	Soybean	7.3 km	-	stereo RGB-D	RTK	✓	✓
HORTO-3DLM [5]	6 (3 months)	Apple; Strawberry; Cherry; Tomato	3.09 km	1	-	RTK	-	-
BLT [6]	11 (6 months)	Grapevine (trellis)	0.5 km	1	2 RGB-D	RTK	✓	✓
TEMPO-VINE (ours)	13 (10 months)	Grapevine (trellis and pergola)	38 km	2	1 RGB-D	RTK	✓	✓

such as localization [12], path planning[13], navigation[14], harvesting [15], pruning [16], and vegetation assessment [17]. However, the development of reliable and generalizable solutions critically depends on the availability of high-quality datasets and standardized benchmarks to evaluate algorithmic performance in realistic settings.

The lack of datasets collected over entire seasons in operational agricultural environments that capture the complexities of the field represents a key limitation to research advancements. Robotic systems are often developed and validated in simulation [18] or isolated experiments, which fail to represent the variability in terrain, vegetation, and seasonal conditions. Among the most attractive crop fields worldwide, vineyards pose significant challenges for localization and mapping due to their dynamic appearance, structural variation, and changing environmental factors throughout the year.

To address this gap, we introduce TEMPO-VINE, a novel multi-temporal and multi-modal dataset designed for evaluating odometry, sensor fusion, simultaneous localization and mapping (SLAM), and place recognition methods in large-scale vineyard environments. TEMPO-VINE includes data from heterogeneous sensors, including two LiDARs (Velodyne VLP-16 and Livox Mid-360), a RGB-D camera, Attitude and Heading Reference System (AHRS), and RTK-GPS. Data have been collected through multiple campaigns from winter to autumn 2025, capturing diverse terrain, vegetation and weather conditions. The data are provided in a Robot Operating System (ROS) compatible format, with bags of data that can be easily played to test the desired algorithm and compute benchmarking metrics using the ground truth trajectories. Multiple runs for each campaign have been performed to augment the dataset with rich trajectory samples and to enable place recognition with revisited sequences. TEMPO-VINE aims to foster the development of robust autonomous systems for agricultural robotics by providing a realistic, challenging, and diverse dataset for the robotics community.

The main contributions of this work are as follows:

- The TEMPO-VINE dataset is a novel multi-seasonal and multi-modal benchmark for odometry, SLAM, sensor fusion, and place recognition in real-world vineyards covering both standard trellis and pergola vine architectures with row lengths exceeding 100 meters.
- We present the first publicly available vineyards dataset that includes heterogeneous LiDAR data, supporting research in both high-performance and cost-effective robotic platforms.
- We provide ROS-compatible data format and ground-truth trajectories for repeated and loop-closure sequences, enabling reproducible evaluations.
- Through benchmarking experiments, we highlight the challenges posed by seasonal and structural variability in vineyards, emphasizing the need for more robust algorithms for perception and sensor fusion in agricultural fields.

II. RELATED WORKS

A. Heterogeneous LiDAR datasets

Light Detection and Ranging (LiDAR) sensors have proven highly suitable for place recognition tasks due to their invariance to illumination changes. Popular datasets, such as the Oxford [19], USyd [20] and NCLT [21] datasets, collect data with this type of sensor during long-term campaigns to provide reliable data for developing and evaluating place recognition and simultaneous localization and mapping (SLAM) algorithms. These datasets capture LiDAR and vision information, along with seasonal changes, as they were recorded over the course of more than one year of consistent data collection sessions.

Recent studies have also focused on improving place recognition through deep learning, aiming at representations that remain robust under seasonal and appearance changes [22]. Among the most common approaches is LiDAR–vision fusion, which combines the geometric invariance of LiDAR with the rich texture information provided by cameras [23]. More recently, a growing trend is the fusion of different LiDAR sensors, since their varying fields of view (FoV) and coverage can produce higher-quality environmental maps [24]. Datasets such as HeliPR [25] and GEODE [26] exemplify this direction, as they integrate data collected with multiple LiDAR sensors featuring different characteristics. These datasets also provide loop-closure trajectories for SLAM evaluation. Nevertheless, they lack seasonal variability, as they were not acquired over long-term campaigns, and

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.

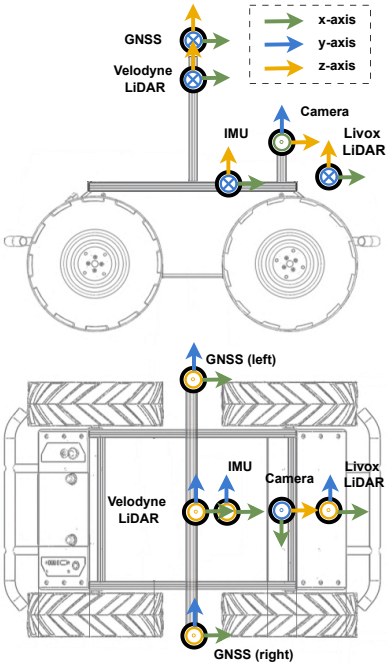

Fig. 2: Schematic of the sensors disposition on the rover used for the data collection activity in the vineyards.

the captured runs are relatively short and limited to urban environments.

B. Agricultural datasets

While the previously mentioned urban datasets have become standard benchmarks for LiDAR-based perception and localization, agricultural environments pose unique challenges that require dedicated datasets. Unlike urban scenes with structured roads, buildings, and traffic rules, agricultural fields are highly dynamic and unstructured, with seasonal variations, repetitive patterns (e.g., vineyard rows), and heavy occlusions caused by vegetation.

Many agricultural datasets to date such as ARD-VO [3], MAgro [2] or Rosario [4] primarily rely on RGB or multispectral imagery, which constrains their applicability to tasks beyond crop monitoring or visual detection. However, the lack of three-dimensional information limits progress on spatial perception problems such as mapping, navigation, and long-term place recognition. To address this gap, recent efforts like the Bacchus Long-Term (BLT) [6] or the HORTO-3DLM [5] datasets have started to incorporate 3D LiDAR recordings. Such developments open new opportunities for LiDAR-based place recognition (LPR) in agricultural environments. Nonetheless, those datasets only cover small fields with incomplete or costly sensor suites.

Table I compares agricultural datasets and shows that, to date, TEMPO-VINE offers the most complete sensor suite, making it a challenging and realistic benchmark for robot localization and mapping. Our dataset is the only one in the comparison that includes heterogeneous LiDAR; it has both a Velodyne VLP-16 and a Livox Mid-360. The Velodyne VLP-16 is a rotating LiDAR that generates a regular scan pattern with 16 simultaneous laser channels. In contrast, the Livox Mid-360 is a single-beam sensor that uses a non-repetitive scanning pattern to gradually cover its field of view. This results in different point distributions and sensing characteristics.

The TEMPO-VINE dataset was collected to address the need for agricultural datasets capturing long-term data throughout the crop's life cycle during different seasons. Vineyards are extremely dynamic environments due to vegetation growth, and, nonetheless, these factors should be combined with weather conditions and with the different field of view of the available sensors. Unlike previous vineyard-focused agricultural datasets, TEMPO-VINE provides multi-environment coverage, including both trellis and pergola vineyard structures, with rows exceeding 100 m in length. The dataset also includes measurements from two complementary LiDAR sensors (one of them low-cost) together with additional modalities, enabling both intra-modal and cross-modal sensor fusion. Beyond geometry, it captures multi-seasonal variability, including different stages of crop development and ground cover conditions. Furthermore, the dataset is fully ROS-compatible, facilitating integration into robotic pipelines.

III. Experimental Settings

The recorded sensors have been selected to mix standard perception hardware adopted in robotics and low-cost sensing solutions. Indeed, beside the more advanced and costly Velodyne VLP-16 3D LiDAR, we also include a cheaper Livox Mid-360, a toroidal non-repetitive pattern 3D LiDAR that recently appeared on the market. The sensor suite also includes an RGB-D Intel Realsense D435 camera, providing aligned RGB and depth images, an accurate AHRS MicroStrain 3DM-GX5, and wheel encoders integrated in the rover's motors. Detailed sensor specifications are reported in Table II.

The entire data collection has been performed with a Husky A200 rover from ClearPath Robotics [2] (size 990 x 670 x 390 mm) equipped with multiple commercial sensors and an on-board computer with an Intel i7-6700E processor, 16 GB DDR4 RAM, and 1 TB SSD. Figure 2 depicts the disposition of the different sensors on the robot for the data collection activity.

The rover has been manually teleoperated at a maximum linear velocity of $1[m/s]$ and a maximum angular velocity of $2.0[rad/s]$ to follow the desired trajectory. It has been remotely driven keeping it always in line of sight from the operator. Nonetheless, terrain irregularity and obstacles caused sudden drifts of the rover, which are difficult to promptly correct from long distances. This makes the trajectory collected highly realistic and comparable to an autonomous driving setting, with the rover not always staying at the centre of the rows.

[2]Husky A200 rover datasheet

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.

TABLE II
SENSOR SPECIFICATIONS

Sensor	Type	Frequency	FOV			Resolution
			Range	Horizontal	Vertical	
3D LiDAR 1	Livox Mid-360	10 Hz	0.1–40 m	360°	$-7°$ to $+52°$	0.2 m @1σ; $\pm$0.15° @1σ
3D LiDAR 2	Velodyne Puck VLP-16	10 Hz	100 m	360°	$-15°$ to $+15°$	0.03 m; 0.1°–0.4°; 2.0°
RGB-D Camera	Intel Realsense D435	30 Hz	0.3–3.0 m	69°	42°	640×480 px
AHRS-IMU	MicroStrain 3DM-GX5	100 Hz	–	–	–	0.02 mg (accel); 0.003°/s (gyro)
RTK-GPS	Swift Navigation Duro	5 Hz	–	–	–	∼1–2 cm horiz.; ∼2–5 cm vert.

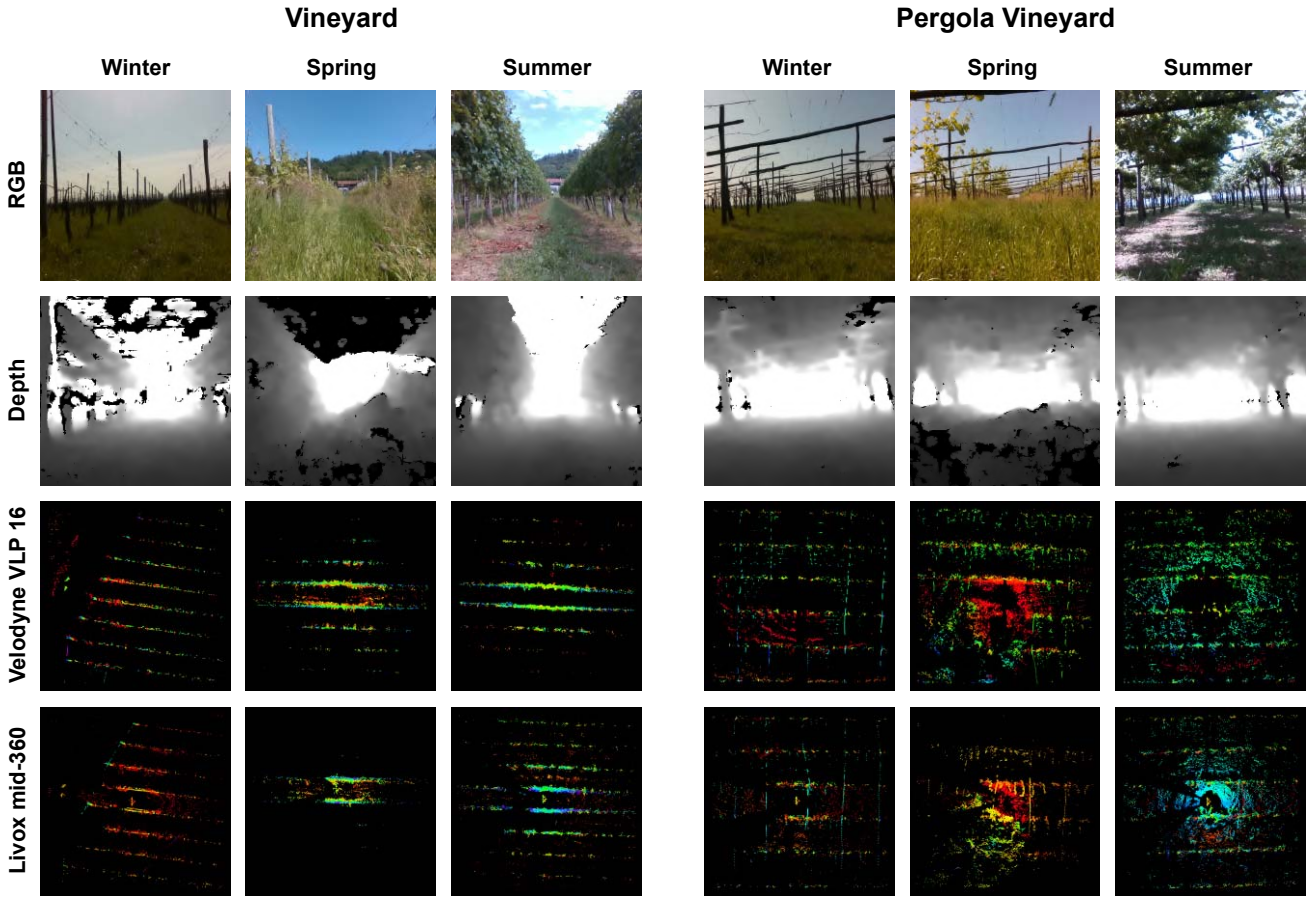

Fig. 3: Samples of RGB-D images and LiDAR point cloud collected in the trellis and pergola vineyards in the same position over different seasons from winter to summer. Different vegetation growth stages represent a key dynamic environmental aspect for robotics navigation.

IV. DESCRIPTION OF THE DATASET

In this section, the dataset composition is presented, starting from the description of the data collection campaigns and trajectories performed on the field. Then, the ground truth trajectory computation is explained, together with the file structure and organization of the dataset.

A. Vineyards Campaigns and Trajectories

The TEMPO-VINE dataset currently comprises sensor data collected over 10 experimental campaigns distributed from February to September 2025 (with two other campaigns already scheduled in October and November that will be added upon approval in the review stage).

In each campaign multiple runs have been conducted in two different vineyards:

- A trellis Guyot vineyard with size 110×25 [m]. The data have been recorded using 10 rows of the field. Each row has an average width of 2 [m].
- A pergola vineyards with 4 rows and size 110×15 [m]. Each row has an average width of 3.5 [m].

Both vineyards are located in Agliè, Torino, Italy. The peculiarity of the dataset is its multi-seasonal temporal span,

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.

TABLE III
OVERVIEW OF FIELD CAMPAIGNS IN THE VINEYARDS.

Campaign	Vineyard	Runs	Weather	Plant Growth	Grass Height
ID 01	Trellis	01, 02	Cloud	Stem, no leaves	Low (∼ 5 cm)
12/02	Pergola	01, 02	Cloud	Stem, no leaves	Low (∼ 5 cm)
ID 02	Trellis	01, 02	Sun	Stem, no leaves	Medium (∼ 20 cm)
20/03	Pergola	01, 02	Sun	Stem, no leaves	Medium (∼ 20 cm)
ID 03	Trellis	01, 02	Sun	Stem, no leaves	Tall (>40 cm)
11/04	Pergola	01, 02	Sun	Stem, no leaves	Tall (>40 cm)
ID 04	Trellis	01, 02	Sun	Stem, few leaves	Tall (>40 cm)
15/05	Pergola	01, 02	Sun	Stem, few leaves	Tall (>40 cm)
ID 05	Trellis	01, 02	Cloud	Branches, leaves	Medium (∼ 20 cm)
06/06	Pergola	01, 02	Cloud	Branches, leaves	Tall (>40 cm)
ID 06	Trellis	01, 02	Sun	Branches, leaves	Medium (∼ 20 cm)
25/06	Pergola	01, 02	Sun	Branches, leaves	Medium (∼ 20 cm)
ID 07	Trellis	01, 02, 03	Sun	Branches, leaves	Low (∼ 5 cm)
11/07	Pergola	01, 02	Sun	Branches, leaves	Medium (∼ 20 cm)
ID 08	Trellis	01, 02, 03	Sun	Branches, leaves, fruit	Low (∼ 5 cm)
28/07	Pergola	01, 02	Sun	Branches, leaves, fruit	Low (∼ 5 cm)
ID 09	Trellis	01, 02, 03	Sun	Branches, leaves, fruit	Medium (∼ 20 cm)
14/08	Pergola	01, 02	Sun	Branches, leaves, fruit	Low (∼ 5 cm)
ID 10	Trellis	01, 02, 03	Cloud	Branches, leaves, fruit	Medium (∼ 20 cm)
11/09	Pergola	01, 02	Cloud	Branches, leaves, fruit	Tall (>40 cm)
ID 11	Trellis	01, 02, 03	Sun	Branches, leaves, fruit	Low (∼ 5 cm)
02/10	Pergola	01, 02	Sun	Branches, leaves, fruit	Low (∼ 5 cm)
ID 12	Trellis	01, 02, 03	Cloudy	Branches, leaves	Low (∼ 5 cm)
30/10	Pergola	01, 02	Cloudy	Branches, leaves	Low (∼ 5 cm)
ID 13	Trellis	01, 02, 03	Cloudy	Branches, leaves	Medium (∼ 20 cm)
19/11	Pergola	01, 02	Cloudy	Branches, leaves	Medium (∼ 20 cm)

Fig. 4: Trajectories for the three runs conducted in the trellis vineyard and in the pergola vineyards, overlaid with the satellite image of the field.

offering a wide range of diverse field conditions. Table III reports the entire sequence of campaigns conducted, the number of runs and details about the state of the vineyard. The growth of plants' foliage over seasons, the presence of grass at different heights, and the weather conditions significantly shape the landscape of the vineyards over both long and short time periods, making it a highly dynamic and challenging environment for navigation. These factors influence the distribution of data taken from sensors, generating strong domain gaps in both images and point clouds and hindering the performance of machine learning, localization and place recognition algorithms. Some examples of data samples from the two LiDARs and the RGB-D camera collected in different seasons on the two vineyards are shown in Fig. 3. As can be seen in the selected samples, point clouds registered in the same position at different times drastically change, as well as the visual appearance of the row. Moreover, the two LiDARs offer different perspectives of the same row due to their specific field of view and position on the rover. This is extremely clear during Spring samples reported in Fig. 3.

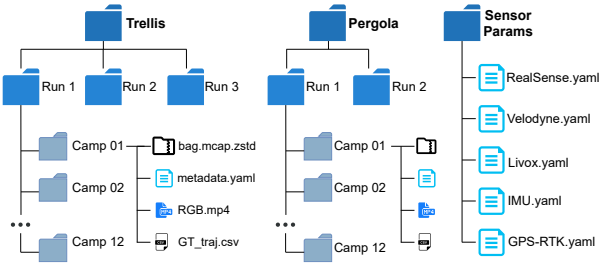

Fig. 5: File structure of the TEMPO-VINE dataset: the folders are organized to easily select the vineyard field, the run and the campaign date. For each experiment a bag file with metadata containing all the complete sensor data stream is provided, together with the RGB camera video and the ground truth trajectory file.

The runs performed follow a specific pattern among the rows of the two vineyards to guarantee consistency and reproducibility of the algorithms' performance over different seasonal conditions, and to enable place recognition with multiple revisits of the same rows. The complete visualization of the different trajectories performed in each campaign is visually shown in Fig. 4, overlaid on the satellite map of the entire vineyard. In the classic trellis vineyards (Fig. 4, top image), Run 01 is carried out traversing 10 rows continuously, while Run 02 covers the same path for the first 4 rows, then the other 4 rows are traversed in the opposite direction. Run 03 has been collected starting from the summer months when the vegetation reached a significant thickness, and it is looping over the first 3 rows to support the place recognition task in the most challenging scenario. In the pergola (Fig. 4, bottom image), we performed 2 runs with full coverage trajectories in the 4 available rows in all the months.

B. Dataset Organization and Format

The file structure of the TEMPO-VINE dataset is schematically illustrated in Figure 5. The structure is thought to easy access data of a specific experiment, choosing the vineyard field first, then the run, and finally the campaign. Each campaign folder contains 4 files: the compressed bag file `bag.mcap.zstd` with all the raw stream data from all the sensors, the `metadata.yaml` with all the bag information, the complete `RGB.mp4` video recording from the on-board camera, and the ground truth trajectory file `GT_traj.csv`.

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
RESULTS OF SLAM ALGORITHMS FOR EACH SENSOR TESTED IN WINTER (MARCH) AND SUMMER (JULY) IN BOTH THE TRELLIS (RUN 02) AND PERGOLA (RUN 01) VINEYARDS.

Sensor	RMSE	Trellis		Pergola	
		02-R02	07-R02	02-R01	07-R01
RGBD Camera (RTAB-Map [27])	ATE_t	8.32	5.97	4.58	1.72
	RPE_r	0.03	0.03	0.09	0.02
	RPE_t	0.21	0.26	2.21	0.24
Velodyne VLP-16 (Fast-LIO [28])	ATE_t	23.93	26.46	0.58	1.54
	RPE_r	0.04	0.05	0.06	0.05
	RPE_t	0.25	0.33	0.36	0.24
Velodyne VLP-16 (LIO-SAM [29])	ATE_t	1.93	3.30	3.78	1.09
	RPE_r	0.45	0.57	0.94	0.74
	RPE_t	0.03	0.03	0.04	0.03
Livox MID-360 (Fast-LIO [28])	ATE_t	1.37	23.63	2.94	2.20
	RPE_r	0.04	0.05	0.06	0.03
	RPE_t	0.17	0.38	0.32	0.22

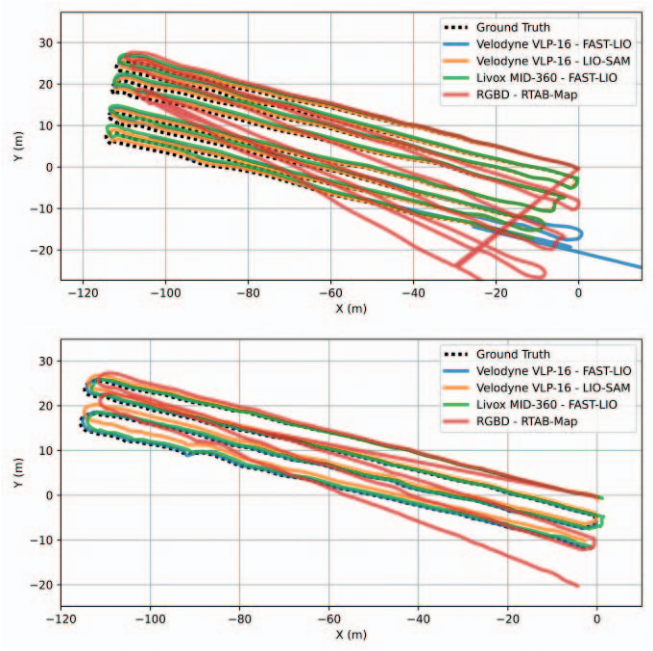

Fig. 6: SLAM results and ground truth trajectories on trellis (top) and pergola (bottom) vineyards from campaign 02 (March) with two LiDARs and the RGB-D camera.

The bag file is ROS compatible, but does not necessarily require ROS installation to be played. Sensors settings are a common feature for all the campaigns, hence the parameter files of each sensors are contained in the separate *Sensor Params* folder in a `.yaml` format.

C. Ground Truth

Ground truth is obtained using a high-precision GNSS-RTK receiver in combination with a high-end AHRS. Data from both sources are integrated to estimate the robot center's position with a temporal resolution of 0.2 seconds. For each sensor, the extrinsic transformation to the robot center is specified. The ground truth trajectory is represented in a text file, where each line contains the timestamp, translation, and rotation (expressed as a quaternion) in a local ENU frame. A geo-referenced trajectory is also provided in the same format, with translation replaced by latitude, longitude, and altitude. This format, referred to as TUM, was selected to ensure compatibility with existing evaluation tools.

TABLE V
PLACE RECOGNITION QUANTITATIVE ANALYSIS WITH SCAN CONTEXT[30]: RECALL@5 AND PRECISION@5 FOR DIFFERENT THRESHOLDS AND SEASONS IN TRELLIS VINEYARD (RUN 01).

Sensor	Camp	Recall			Precision		
		10m	15m	20m	10m	15m	20m
VLP-16	02-March	0.98	0.99	0.99	0.87	0.91	0.94
	04-May	0.85	0.90	0.93	0.58	0.655	0.71
	06-June	0.27	0.41	0.53	0.10	0.19	0.29
Mid-360	02-March	0.76	0.83	0.87	0.38	0.49	0.59
	04-May	0.41	0.55	0.68	0.17	0.26	0.36
	06-June	0.28	0.44	0.55	0.11	0.20	0.31

V. DATASET EVALUATION

In this section, we validate the data testing state-of-the-art algorithms for SLAM, both LiDAR and visual SLAM, and place recognition in the vineyard scenario. The results demonstrate the performance obtained using different sensors and the complexity generated from vegetation conditions in winter and summer.

A. SLAM results

To assess the SLAM performance of the proposed dataset, three state-of-the-art ROS2-compatible algorithms were evaluated: RTAB-Map [27] for RGB-D data, Fast-LIO [28] and LIO-SAM [29] for LiDAR data. Experiments were conducted using data from campaigns 02 (March) and 07 (July), covering different seasonal conditions and two vineyard training systems: trellis and pergola. Each sensor was tested in each condition, resulting in 12 experiments in total. Performance was evaluated against ground truth using Absolute Pose Error (APE, in meters) and Relative Pose Error (RPE), with RPE_t representing translational error (in meters) and RPE_r rotational error (unitless). APE values exceeding 4 meters were considered failures, indicating SLAM divergence. LIO-SAM obtained the most robust result overall, performing successfully in all seasons. However, the algorithm is compatible only with ring-organized data like Velodyne. Fast-LIO with Livox data achieved precise performace, only failing in the summer trellis scenario. With Velodyne, Fast-LIO performed well in the pergola but failed in the final rows of the trellis. In contrast, RTAB-Map consistently underperformed across all conditions, highlighting the challenges posed by the environment for RGB-D-based SLAM. Results are summarized in Table IV and Figure 6 shows the trajectories of the experiments from campaign 02.

B. Place recognition results

To evaluate the place recognition capabilities of our dataset, we used a handcrafted global descriptor, Scan

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.

TABLE VI
PLACE RECOGNITION QUANTITATIVE ANALYSIS OF LEARNED METHODS: RECALL@1% AND RECALL@1.

Sensor	Camp	Backbone	Recall@1%	Recall@1
VLP-16	02-March	PointNetVLAD	0.68	-
	02-March	MinkLoc3Dvs	0.42	0.32
	04-May	PointNetVLAD	0.52	-
	04-May	MinkLoc3Dv2	0.49	0.30
	06-June	PointNetVLAD	0.47	-
	06-June	MinkLoc3Dv2	0.38	0.30
Mid-360	02-March	PointNetVLAD	0.59	-
	02-March	MinkLoc3Dv2	0.39	0.32
	04-May	PointNetVLAD	0.46	-
	04-May	MinkLoc3Dv2	0.35	0.30
	06-June	PointNetVLAD	0.50	-
	06-June	MinkLoc3Dv2	0.40	0.29

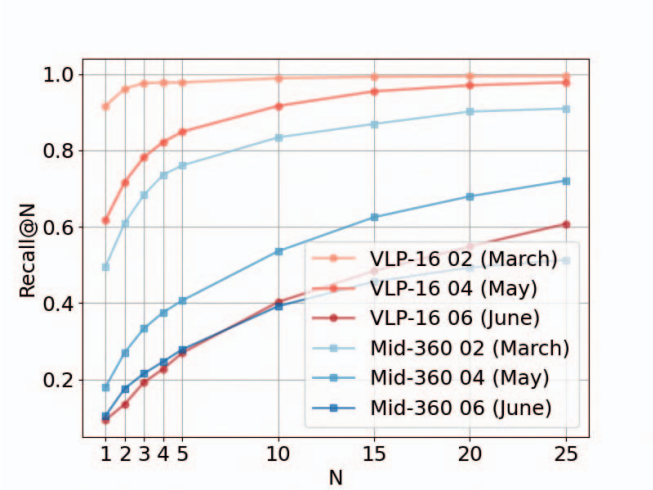

Fig. 7: Place recognition: Recall graph for different values of N. Blue curves are obtained with the Livox Mid-360, the red curves with the Velodyne VLP-16 LiDAR in different seasons.

Context [30], and two state-of-the-art deep learning global descriptors, PointNetVLAD [31] and MinkLoc3Dv2 [32]. Following an inter-session setup, we tested how well the method generalizes when revisiting the same vineyard locations at different stages of the vine growth cycle. For this evaluation, we only considered the vineyard sequences, excluding the pergola. Specifically, we focused on the Run 01 recordings, which provide the most extensive coverage of the vineyard terrain. To highlight the impact of seasonal changes on agricultural place recognition, we used the first run (recorded in February) as the database, while the subsequent monthly campaigns were used as queries in each iteration. Table V reports the Scan Context quantitative results for all campaigns, using Recall@5 and Precision@5 under different positive match thresholds. On the other hand, Table VI shows the results for the selected months with the two learned descriptors in terms of Recall@1% and Recall@1, using Run 01 from the selected campaign as the training set and Run 02 as the evaluation set. A correct match is selected in a maximum range of 5 meters.

The results show high performance when queries belong to

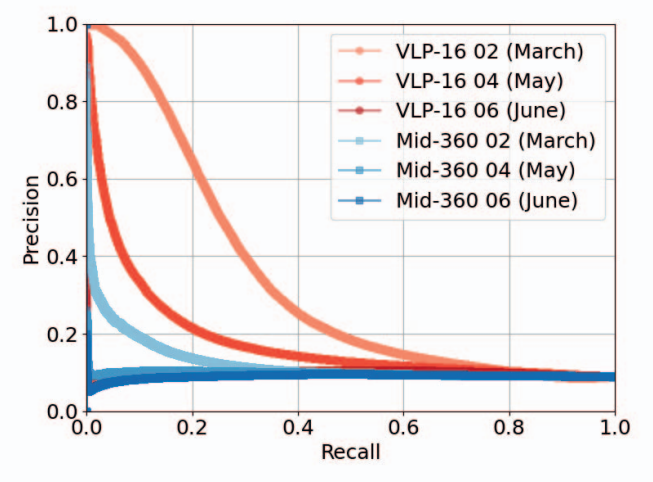

Fig. 8: Place recognition: Precision-recall curve for different season generalization scenarios. Blue curves are obtained with the Livox Mid-360, the red curves with the Velodyne VLP-16 LiDAR.

the same season database, but performance decreases as the crop undergoes significant changes in subsequent months. Furthermore, the geometry-based descriptor Scan Context outperforms the two learned methods, which were designed for urban environments. This evidences the challenges posed by agricultural environments for localization. Figures 7 and 8 present the results in graphical form, showing Recall@N values and the precision–recall curves across the different months.

VI. CONCLUSIONS

The TEMPO-VINE dataset is a multi-temporal dataset for benchmarking sensor fusion, SLAM, and place recognition research methods in a realistic vineyard environment. Robotics and autonomous driving solutions are often tested in simulated or isolated field experiments, often lacking the necessary validation process to introduce robust and continuous operations on the field. Vineyards are a strongly dynamic and challenging environment where automation solutions are raising interest of both research and industry. TEMPO-VINE is the first public dataset comprising an extensive data collection on large-scale vineyards with multiple rows longer than 100 m. The benchmark is fully ROS-compatible and it offers raw data and ground truth trajectories for multiple runs in a standard trellis and also in a pergola architecture, which was missing in the landscape of agricultural datasets. Nonetheless, TEMPO-VINE is a heterogeneous LiDAR dataset, mixing advanced 3D LiDAR such as the Velodyne VLP-16 with the more recent and low-cost sensing solution offered by the Livox Mid-360. Platform cost reduction is another fundamental research goal that must be explored to lead automation to worldwide success and diffusion, especially for small agricultural realities.

Through benchmark evaluations of both SLAM and place recognition tasks, we demonstrated the limitations in reaching robust performance from winter to summer, underscoring

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.

the need for further sensor fusion and SLAM research in these real settings.

As future work, the dataset will be extended to reach two years temporal coverage, including all the four seasons, and different trajectories. SLAM benchmarking will be enriched and open to the community, carrying out introspection analyses about strengths and failures of methods over the different field conditions. By providing multi-seasonal, multi-modal data under different vineyards and environmental conditions, TEMPO-VINE raises the bar of autonomous navigation benchmarking in agriculture, paving the way for the next generation of robotics and automation solutions.

REFERENCES

[1] F. Cañadas-Aránega, J. L. Blanco-Claraco, J. C. Moreno, and F. Rodriguez-Diaz, "Multimodal mobile robotic dataset for a typical mediterranean greenhouse: The greenbot dataset," *Sensors*, vol. 24, no. 6, p. 1874, 2024.

[2] M. Marzoa Tanco, G. Trinidad Barnech, F. Andrade, J. Baliosian, M. LLofriu, J. M. Di Martino, and G. Tejera, "Magro dataset: A dataset for simultaneous localization and mapping in agricultural environments," *The International Journal of Robotics Research*, vol. 43, no. 5, pp. 591–601, 2024.

[3] F. Crocetti, E. Bellocchio, A. Dionigi, S. Felicioni, G. Costante, M. L. Fravolini, and P. Valigi, "Ard-vo: Agricultural robot data set of vineyards and olive groves," *Journal of Field Robotics*, vol. 40, no. 6, pp. 1678–1696, 2023.

[4] N. Soncini, J. Cremona, E. Vidal, M. García, G. Castro, and T. Pire, "The rosario dataset v2: Multi-modal dataset for agricultural robotics," *The International Journal of Robotics Research*, p. 02783649251368909, 2024.

[5] T. Barros, C. Premebida, S. Aravecchia, C. Pradalier, and U. Nunes, "Spvsoap3d: A second-order average pooling approach to enhance 3d place recognition in horticultural environments," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 9–15.

[6] R. Polvara, S. Molina, I. Hroob, A. Papadimitriou, K. Tsiolis, D. Giakoumis, S. Likothanassis, D. Tzovaras, G. Cielniak, and M. Hanheide, "Bacchus long-term (blt) data set: Acquisition of the agricultural multimodal blt data set with automated robot deployment," *Journal of Field Robotics*, vol. 41, no. 7, pp. 2280–2298, 2024.

[7] Z. Zhai, J. F. Martínez, V. Beltran, and N. L. Martínez, "Decision support systems for agriculture 4.0: Survey and challenges," *Computers and Electronics in Agriculture*, vol. 170, p. 105256, 2020.

[8] S. K. Panda, Y. Lee, and M. K. Jawed, "Agronav: Autonomous navigation framework for agricultural robots and vehicles using semantic segmentation and semantic line detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 6272–6281.

[9] S. Cerrato, V. Mazzia, F. Salvetti, M. Martini, S. Angarano, A. Navone, and M. Chiaberge, "A deep learning driven algorithmic pipeline for autonomous navigation in row-based crops," *IEEE Access*, 2024.

[10] H. Ding, B. Zhang, J. Zhou, Y. Yan, G. Tian, and B. Gu, "Recent developments and applications of simultaneous localization and mapping in agriculture," *Journal of field robotics*, vol. 39, no. 6, pp. 956–983, 2022.

[11] D. Bigelow and A. Borchers, "Major uses of land in the united states, 2012," *Economic Information Bulletin Number 178*, no. 1476-2017-4340, p. 69, 2017.

[12] W. Winterhalter, F. Fleckenstein, C. Dornhege, and W. Burgard, "Localization for precision navigation in agricultural fields—beyond crop row following," *Journal of Field Robotics*, vol. 38, no. 3, pp. 429–451, 2021.

[13] F. Salvetti, S. Angarano, M. Martini, S. Cerrato, and M. Chiaberge, "Waypoint generation in row-based crops with deep learning and contrastive clustering," in *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2022, Grenoble, France, September 19–23, 2022, Proceedings, Part VI*. Springer, 2023, pp. 203–218.

[14] Z. Man, J. Yuhan, L. Shichao, C. Ruyue, X. Hongzhen, and Z. Zhenqian, "Research progress of agricultural machinery navigation technology," *Nongye Jixie Xuebao/Transactions of the Chinese Society of Agricultural Machinery*, vol. 51, no. 4, 2020.

[15] W. Hua, Z. Zhang, W. Zhang, X. Liu, C. Hu, Y. He, M. Mhamed, X. Li, H. Dong, C. K. Saha, *et al.*, "Key technologies in apple harvesting robot for standardized orchards: A comprehensive review of innovations, challenges, and future directions," *Computers and Electronics in Agriculture*, vol. 235, p. 110343, 2025.

[16] A. Navone, M. Martini, and M. Chiaberge, "Autonomous robotic pruning in orchards and vineyards: A review," *Smart Agricultural Technology*, vol. 12, p. 101283, 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2772375525005143

[17] A. Feng, J. Zhou, E. D. Vories, K. A. Sudduth, and M. Zhang, "Yield estimation in cotton using uav-based multi-sensor imagery," *Biosystems Engineering*, vol. 193, pp. 101–114, 2020.

[18] M. Martini, A. Eirale, B. Tuberga, M. Ambrosio, A. Ostuni, F. Messina, L. Mazzara, and M. Chiaberge, "Enhancing navigation benchmarking and perception data generation for row-based crops in simulation," in *Precision agriculture'23*. Wageningen Academic, 2023, pp. 451–457.

[19] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, "1 year, 1000 km: The oxford robotcar dataset," *The International Journal of Robotics Research*, vol. 36, no. 1, pp. 3–15, 2017.

[20] W. Zhou, J. S. Berrio, C. De Alvis, M. Shan, S. Worrall, J. Ward, and E. Nebot, "Developing and testing robust autonomy: The university of sydney campus data set," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 4, pp. 23–40, 2020.

[21] N. Carlevaris-Bianco, A. K. Ushani, and R. M. Eustice, "University of michigan north campus long-term vision and lidar dataset," *The International Journal of Robotics Research*, vol. 35, no. 9, pp. 1023–1035, 2016.

[22] J. Vilella-Cantos, J. J. Cabrera, L. Payá, M. Ballesta, and D. Valiente, "Minkunext-si: Improving point cloud-based place recognition including spherical coordinates and lidar intensity," *arXiv preprint arXiv:2505.17591*, 2025.

[23] J. Komorowski, M. Wysoczańska, and T. Trzcinski, "Minkloc++: lidar and monocular image fusion for place recognition," in *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2021, pp. 1–8.

[24] M. Jung, S. Jung, H. Gil, and A. Kim, "Helios: Heterogeneous lidar place recognition via overlap-based learning and local spherical transformer," *arXiv preprint arXiv:2501.18943*, 2025.

[25] M. Jung, W. Yang, D. Lee, H. Gil, G. Kim, and A. Kim, "Helipr: Heterogeneous lidar dataset for inter-lidar place recognition under spatiotemporal variations," *The International Journal of Robotics Research*, vol. 43, no. 12, pp. 1867–1883, 2024.

[26] Z. Chen, Y. Qi, D. Feng, X. Zhuang, H. Chen, X. Hu, J. Wu, K. Peng, and P. Lu, "Heterogeneous lidar dataset for benchmarking robust localization in diverse degenerate scenarios," *The International Journal of Robotics Research*, p. 02783649251344967, 2024.

[27] M. Labbé and F. Michaud, "Rtab-map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation," *Journal of Field Robotics*, vol. 36, no. 2, pp. 416–446, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21831

[28] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "FAST-LIO2: fast direct lidar-inertial odometry," *CoRR*, vol. abs/2107.06829, 2021. [Online]. Available: https://arxiv.org/abs/2107.06829

[29] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and R. Daniela, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 5135–5142.

[30] G. Kim and A. Kim, "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)*. IEEE, 2018, pp. 4802–4809.

[31] M. A. Uy and G. H. Lee, "PointNetVLAD: Deep point cloud based retrieval for large-scale place recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4470–4479.

[32] J. Komorowski, "Improving point cloud based place recognition with ranking-based loss and large batch training," in *2022 26th international conference on pattern recognition (ICPR)*. IEEE, 2022, pp. 3699–3705.

Authorized licensed use limited to: b-on: Universidade de Coimbra. Downloaded on October 08,2026 at 09:43:50 UTC from IEEE Xplore. Restrictions apply.