

The GREENBOT dataset: Multimodal mobile robotic dataset for a typical Mediterranean greenhouse

Fernando Cañadas-Aránega, Jose Luis Blanco-Claraco,
Jose Carlos Moreno and Francisco Rodriguez

January 31st, 2024

Abstract

This paper introduces an innovative dataset specifically crafted for challenging agricultural settings (a greenhouse), where achieving precise localization is of paramount importance. The dataset was gathered using a mobile platform equipped with a set of sensors typically used in mobile robots, as it was moved through all the corridors of a typical Mediterranean greenhouse featuring tomato crop. This dataset presents a unique opportunity for constructing detailed 3D models of plants in such indoor-like space, with potential applications such as robotized spraying. For the first time to the best knowledge of authors, a dataset suitable to put at test Simultaneous Localization and Mapping (SLAM) methods is presented in a greenhouse environment, which poses unique challenges. The suitability of the dataset for such goal is assessed by presenting SLAM results with state-of-the-art algorithms. The dataset is available online in <https://arm.ual.es/arm-group/dataset-greenhouse-2024/>.

1 Introduction

In recent years, the use of technology in agriculture has experienced significant growth, driven by improved productivity and optimization of resources [1]. Technology is crucial in reducing the use of fertilizers, pesticides, and efficient water management [2]. Agricultural activities benefit considerably from technology, ranging from the application of micronutrients to the estimation of fruit quality and quantity, as well as the mechanical or electromagnetic removal of weeds [3]. Robots are presented as autonomous platforms with the ability to carry out these tasks and can also be used for example to collect a large amount of crop data continuously [4].

Central to carrying out these tasks is providing a robot with autonomous navigation capabilities in an agricultural environment. In the particular case

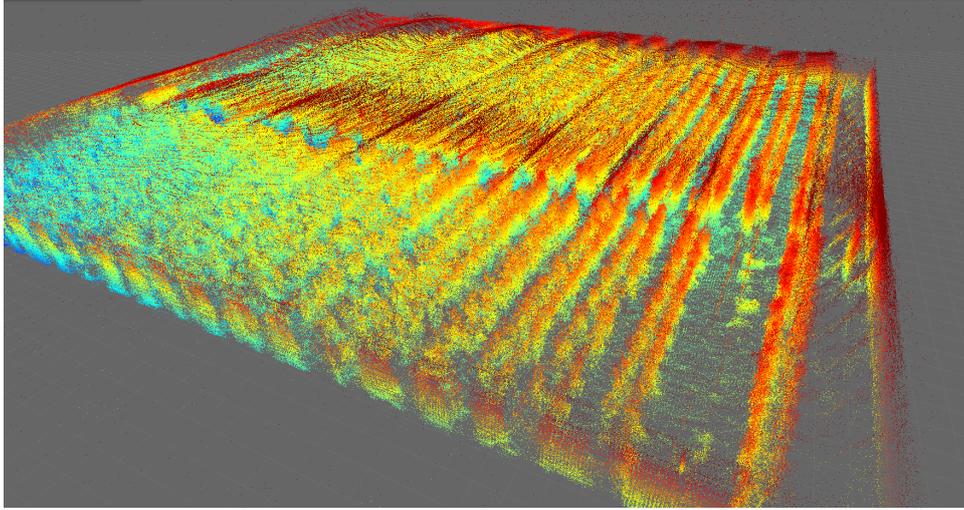


Figure 1: Mapping of the complete greenhouse, orthogonal view with Velodyne VLP16. See discussion on section 5.1.

of a greenhouse, localization requires excellent accuracy and estimation of odometry, which is almost impossible to install in greenhouses [5]. This environment presents particularly challenging features, such as repetitive patterns due to the lack of unique and consistent visual attributes of the vegetation, which poses unique challenges for closed-loop detection algorithms. In addition, these natural environments present unique computer vision challenges due to subtle human-caused movements, extreme light variability due to the closed environment, and seasonal crop changes due to crop management (pruning, harvesting,...) and weather variability.

Large and open datasets are crucial for developing data-driven solutions and benchmarks in a new era of deep learning-based algorithms. In the literature, there are abundant datasets for urban environments (e.g., [6, 7]) and, to the authors' knowledge, the most extensive datasets currently available for agricultural environments are the Rosario Dataset [8], the Sugar Beets Dataset [9] and the MAGRO Dataset [10], for an apple orchard field. However, these datasets are focused on open field agricultural environments where, on the one hand, localization and orientation may be less complex and, on the other hand, the permissible error is greater than in closed environments such as greenhouses, in particular Mediterranean greenhouses, which accounts for 92% of the total greenhouse area worldwide [11, 12], characterized for being structured in irregular and narrow corridors (90-100 cm) with sandy soil (80% of farms).

This paper presents an innovative and dynamic framework for data collection and a dataset collected throughout the growing season in a typical Mediterranean greenhouse using a sensor-equipped forklift. The data col-

lected include images from a stereo camera pointing forward, an IMU and two 3D LIDARs. This dataset provides valuable information to advance the development of visual odometry. Notably, the data were acquired in a typical Mediterranean greenhouse with a tomato crop, which offers the opportunity to develop 3D models for the plants that can be specially useful, for example, for robotized spraying tasks, where the amount of plant protection product is applied while coordinating the relationship between the displacement speed of the robot and the spray speed [13, 14]. It is important to remark that the tomato is one of the most important crops in South-Eastern Spain, occupying 17% of the greenhouse area in the province of Almería alone and accounting for 20% of its total crop production during the 20/21 season [15].

From a robotics perspective, recent advancements have introduced several robotic solutions for agriculture automation. Most of these developments are not technologically mature, however some important companies in the agriculture machinery sector, such as John Deere, New Holland and Case, and other emerging companies with applications in this sector, such as Naïo and DJI, have commercial products or one step away from commercialization. However all of these developments have been conceived to work in open field. For greenhouses, robot prototypes performing different tasks such as weeding, harvesting, pruning, spraying, pest and disease monitoring, etc. have been also developed [16, 17]. Since 1987 many researchers have worked on the development of robots for greenhouses, studying the different problems involved [18]. The AURORA project [19], for instance, introduced a resilient and cost-effective robot designed for greenhouse operations. This robot demonstrates the ability to navigate autonomously through diverse greenhouse environments. Another notable greenhouse automation initiative is [20], featuring a mobile robot equipped with a stereo vision system and a six-degree-of-freedom arm specifically designed for tomato cultivation in greenhouses. In addition, a service robot for sanitary control and localized dispensing of pesticides and fertilizers to plants in greenhouses using RGB-D cameras is presented in [21]. Other projects in Japan [22] have also received funding, contributing to the ongoing exploration and development of robotic solutions for greenhouse applications. It's important to notice that the authors of this paper have worked in three projects directly related to the use of mobile robots in greenhouses. In particular, [23] presents the implementation and testing of navigation algorithms on the FITOROBOT mobile robot [13]. In INVERSOS project [14], a multi-functional mobile robot is developed taking as basis the experience provided by the FITOROBOT project, and in the AGRICOBOT [24] project another mobile robot for greenhouses is developed, but with the particular characteristic of working in a collaborative way close to humans.

The need for this dataset is demonstrated by the fact that there are a multitude of robots that can improve the navigation task. The sensor-

equipped platform used in this paper travels the same route in different conditions and on different days, capturing data with variations in lighting, weather, and crop state. Along with the sensors data, details on the calibration of intrinsic and extrinsic sensors parameters and tools for data processing are provided.

The main contributions of this work are summarised as follows: (i) a novel dataset on an unexplored environment with rich and challenging features is presented; to the best knowledge of the authors, this is the first study on a dataset for mapping and navigation by mobile robots in greenhouses; (ii) an open-source data collection pipeline is developed and published, which will be used to augment and extend the dataset periodically, and (iii) the suitability of the proposed dataset is validated by running a state-of-the-art SLAM algorithm on it. The dataset, its download guide, and the tools to process it are publicly available in the dataset repository ¹.

The rest of the paper is organized as follows. Section 2 presents a literature review, discussing the most relevant articles and focusing on works with application to agriculture. Section 3 presents the sensor-equipped platform and the sensors calibration. Section 4 presents how the data were acquired and the data structure. In Section 5, the dataset is evaluated with a novel algorithm to validate its quality, and finally, Section 6 ends with conclusions and future work.

2 Related work

Datasets tailored to mobile robotics are of great importance for tasks such as image-based localization, LIDAR odometry, and SLAM, among others. Datasets can be classified according to whether they are synthetic (simulation of textures, objects, lighting, etc., from a virtual environment) or real (images or videos, point clouds, etc from a real environment), and according to the type of scene. Table 1 summarizes the related datasets discussed in the rest of this section.

2.1 Synthetic datasets

Starting from the works related to rural environments, [25] proposes a dataset based on a photorealistic environment where they evaluate many SLAM techniques, even going into interiors. In the same line, [26] develops a dataset based on urban photos, building an urban 3D environment and, following the same technique, [27] focuses on mountainous environments. However, for works related to agricultural environments, [28] presents a synthetic dataset that evaluates different SLAM techniques with different

¹University of Almería dataset "GREENBOT": <https://arm.ual.es/arm-group/dataset-greenhouse-2024/>

Table 1: Summary of the surveyed datasets

Dataset	Real/ synthetic	Indoor/ outdoor	Environment	Main task	RGB	Depth	GPS	IMU	LiDAR	Ground truth	Publicly available
[25]	Synthetic	Indoor	Room	SLAM	Yes	Yes	No	No	No	No	No
[26]	Synthetic	Outdoor	Urban	Semantic segmentation for navigation	Yes	Yes	No	No	No	Pose	Yes
[27]	Synthetic	Outdoor	Mountain	SLAM	No	No	Yes	Yes	Yes	Pose	Yes
[28]	Real	Outdoor	Open field	SLAM	Yes	Yes	No	No	No	Semantic label	No
[30]	Synthetic	Outdoor	Different open field	Semantic segmentation for recognise	Yes	Yes	No	No	No	No	Yes
[7]	Real	Outdoor	Urban	SLAM	Yes	No	Yes	No	Yes	Pose	Yes
[6]	Real	Outdoor	Urban	SLAM	Yes	Yes	Yes	Yes	Yes	Pose	Yes
[31]	Real	Outdoor	Urban	SLAM	Yes	No	Yes	No	Yes	Pose	Yes
[32]	Real	Outdoor	Urban	SLAM	No	No	Yes	Yes	Yes	Pose	Yes
[33]	Real	Outdoor	Lake and rivers	Measuring	No	No	Yes	Yes	Yes	Pose	No
[34]	Real	Outdoor	Rivers	SLAM	Yes	No	Yes	Yes	No	No	Yes
[35]	Real	Outdoor	Various open field	Semantic segmentation for navigation	Yes	Yes	No	No	No	Pose	Yes
[36]	Real	Outdoor	Open field	Pest detection	Yes	Yes	No	No	No	No	No
[8]	Real	Outdoor	Soybean field	SLAM	Yes	Yes	Yes	Yes	No	Pose	Yes
[9]	Real	Outdoor	Sugar beet field	Weeding	Yes	Yes	Yes	Yes	Yes	Pose	Yes
[10]	Real	Outdoor	Apple field	SLAM	Yes	Yes	Yes	No	Yes	Pose	Yes
[37]	Real	Outdoor	Building	SLAM	No	No	Yes	Yes	Yes	Pose	Yes
[38]	Real	Indoor	House	SLAM	Yes	No	No	No	No	Pose	No
[39]	Real	Indoor	Laboratory	SLAM	Yes	Yes	No	No	No	Pose	Yes
[40]	Real	Outdoor	Laboratory scale greenhouse	Navigation techniques	Yes	No	No	No	No	No	No
[41]	Real	Indoor/ Outdoor	Mountain and building	SLAM	Yes	No	Yes	Yes	No	Pose	No
GREENBOT	Real	Indoor	Mediterranean greenhouse	SLAM	Yes	Yes	No	Yes	Yes	Pose	Yes

types of plants in different camera positions, fruit labels, etc. In [29] a 3D environment of simulated vineyard data where it evaluates different mapping techniques is created. In [30], authors make a synthetic dataset for different agricultural areas divided into different plots of land for spraying or pruning of fruit trees.

2.2 Real datasets

Within these real datasets, scenes can be classified as indoor or outdoor, with outdoor scenes classifiable according to whether they are recording urban locations or other types of environments.

2.2.1 Outdoors

In terms of environments, outdoor areas often present more difficulties due to changes in lighting conditions, weather, moving objects, etc. These can be further classified into urban or open field environments, depending on the environment where the data were collected. The latter is particularly interesting for the present work because, although it is carried out in a greenhouse, which can be classified as indoor in agriculture, it is crucial to compare it with other open field agriculture datasets, highlighting the non-existence of datasets for indoor agriculture environments such as Mediterranean greenhouses.

- **Urban environments**

In [7] authors present the dataset for a university campus where data are collected using six RGB cameras, a LiDAR, and a GPS. In [6] high-resolution color and greyscale stereo cameras, a Velodyne 3D laser scanner, and a high-precision GPS/IMU are used to collect data from a city. These studies collected data using ground vehicles. In [31], data were captured by flying at low altitudes within urban streets using an aerial micro-vehicle equipped with high-resolution cameras, GPS, and IMU. A dataset based on a high-resolution stereo camera at 20 Hz for 36.8 km in Málaga with a Citroen C4 car is presented in [32].

- **Open field environments**

With the help of human-operated vehicles, [33] collected data of the river and lake vicinity over three years equipped with a 2D LiDAR, RGB camera, GPS, and IMU. Similarly, [34] managed to gather data from the Sangamon River using a canoe equipped with an RGB camera, IMU, and GPS. Focusing on agriculture in open fields, [35] classified the data obtained through an agricultural robot in a beet plantation for three months, acquiring data from an RGB-D camera, IMU, GPS, and LiDAR. Moreover, there are also articles addressing agricultural tasks unrelated to odometry, mapping, or navigation. In [36], a dataset for pest or harmful insect detection is presented, as well as fruit and plant detection [42], among others. The related existing works with the most impact in agriculture focus on the dataset collected in Rosario [8] and in the Sugar Beets [9], where, with the help of mobile robots, they can map by blending techniques based on georeferenced images and SLAM algorithms. These works are focused on weed detection, mapping of different types of crops, etc. Finally, in [10], a dataset is presented using a robot with autonomous navigation, displaying images with RGB-D cameras and 3D LiDAR on a row of apple trees, closing the control loop through a SLAM algorithm.

2.2.2 Indoors

Although most datasets focus their work on outdoors, numerous studies on indoors datasets exist. In this work, after conducting an exhaustive search, they will be divided into buildings and greenhouses. However, as mentioned above, few works discuss datasets within greenhouses. In particular for Mediterranean greenhouses, to the authors' knowledge, there is none.

- **Building environments**

Depending on the scheme to be recorded, these datasets can adopt different types of complexity. If the data are recorded in indoor environments

such as rooms, they are usually easier to recompile and process as they do not experience changes in lighting or significant differences related to seasonal changes. In addition, these environments are fabricated by objects that the user can easily place. In this case, there are a multitude of datasets that recompile authentic interiors as, for example, for a regular house, in [38], a SLAM technique based on RGB cameras using different robotic platforms is presented; in [37], a drone equipped with a stereo camera and an IMU acquires data from the interior of an industrial building. However, the variation of light, natural light, or the need to know the whole environment can lead to a disorientation of the robot, producing an erroneous map. In this case, in [39] authors contemplate the different casuistics that a service robot can experience when it navigates in interiors.

- **Agriculture environments**

Localization by means of GNSS is fundamental for navigation with mobile robots [43], although weak signal reception inside greenhouses leads to many methods exploiting alternatives such as rail-based location [44], or beacons, such as QR codes [45]. To date, most of the work on robots in greenhouses has been focused on (i) weed inspection and spraying and (ii) vegetable harvesting tasks. Regarding datasets, in [40], a small ROS dataset is presented for data collection and mapping using different techniques with RGB cameras, with an unpublished proprietary dataset. Similarly, semantic mapping is performed in [46] to address contextual navigation in a laboratory scale greenhouse with agricultural robots, using the CamVid dataset [41]. Other recent datasets focus on detailed point clouds for phenotyping plants [47].

These works do not provide a detailed dataset that can be used to evaluate different SLAM techniques. The proposed work is entirely novel as there is no dataset using two widely used LiDAR models and cameras navigating through different crop rows inside a greenhouse.

3 Platform description

The platform used to record this dataset was designed at the University of Almería ² (Figure 2). It has a base sized 900 x 500 x 700 mm with a load capacity of 180 kg, equipped with four wheels and two handles for manual operation. The sensors are mounted on top of it, taking into account the specific constraints of each sensor:

²With the support of the Regional Ministry of Economic Transformation, Industry, Knowledge and Universities and the European Regional Development Fund (FEDER) with the projects UAL2020-TEP-A1991 and PY2020_007-A1991.

- *Stereo camera* ³: It has two lenses for capturing synchronized stereo images. It connects to the PC through a FireWire IEEE 1394 interface, providing a data transmission speed of 800 Mb/s. The stored data has a resolution of 1032 x 776 pixels. The field of view is 97° horizontally and 66° vertically within a range of 0.3 m to 20 m. Recording was done at 10 Hz, with a maximum frame rate of 20 fps. This camera is placed on the front so that the images captured are ideal for obstacle identification for a robot.
- *Velodyne VLP16 LiDAR* ⁴: Velodyne’s LiDAR VLP16 with a maximum range of 100 m, a 360° horizontal and 30° vertical field of view is used. It is placed at the top of the runway to have the whole field of view available, except for the operator. It was recorded with a frequency of 10 Hz.
- *Ouster OS0 LiDAR* ⁵: Data are also collected from the Ouster model OS0. This is a 32-channel LiDAR with 360° horizontal and 90° vertical field of view, ranging between 0.3 m and 50 m. The data rate is also left at the default value of 10 Hz, including the onboard 6-axis IMU. This sensor is placed in a structure 20 cm lower than the Velodyne, obstructing its view and hence only collecting data with an effective horizontal field of view of 275°.

A NovAtel GNSS, model IMU-IGM-A1 ⁶, is also installed with the antenna ANTCOM 42G1215 ⁷. In this dataset, GNSS was only used for timestamp synchronization.

3.1 Computing platform

Greenhouses often have hard environmental conditions, with high temperature and humidity [48]. Additionally, power for sensors and computer must be supplied from a battery that balances versatility and space. Taking all this into account, the platform is equipped with a HISTTON computer featuring an Intel i7 - 8550U processor (4 GHz), an Intel UHD 620 graphics card with 24 CUDA cores, and 32 GB of DDR4 RAM. This equipment was selected for its operating temperature range of 0 - 70 °C and humidity range between 0 and 85%. It only consumes 15 W, resulting in minimal power consumption to consider for battery capacity. This computer was equipped with two disk partitions to use both, Ubuntu 20.04 with ROS Noetic, and Ubuntu 22.04 with ROS 2 Humble, providing the opportunity to work with both versions.

³Bumblebee BB2-08S2 (website).

⁴Velodyne VLP16 (website).

⁵Ouster OS0 (website).

⁶NovAtel (website).

⁷ANTCOM (website).

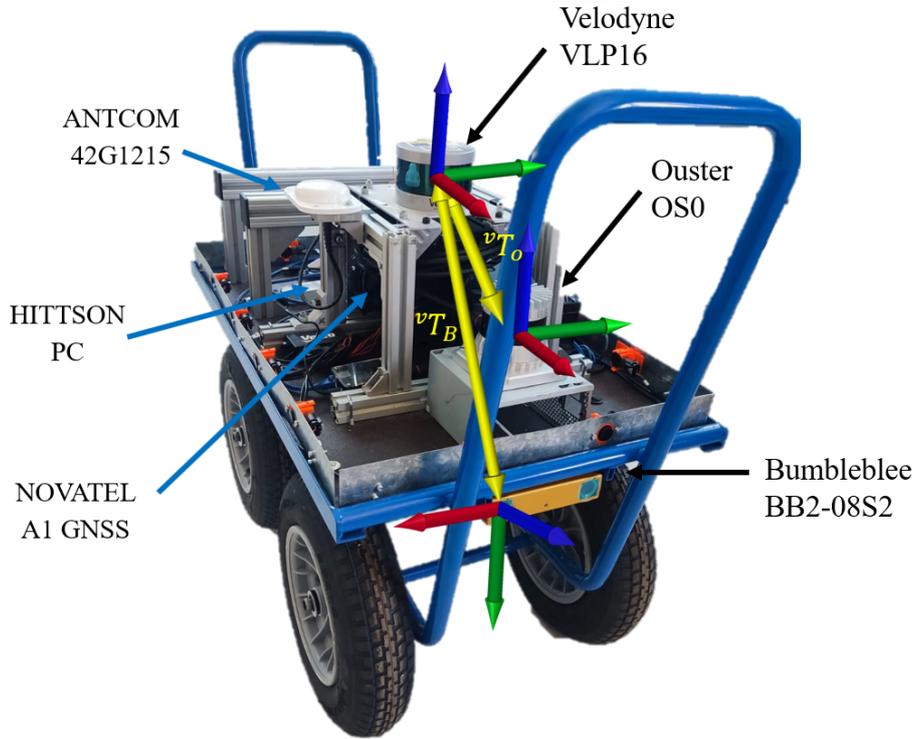


Figure 2: Sensor-equipped platform used to acquire data, equipped with a Bumblebee stereo camera, Novatel GNSS, ANTCOM 45G1215, a Velodyne VLP16 3D LiDAR, and an Ouster OS0 LiDAR with IMU. Also, map of reference local systems for platform sensors. The x-axis is represented in red, the y-axis in green, and the z-axis in blue.

3.2 Calibration

3.2.1 Intrinsic parameters

The intrinsics of the stereo camera have been determined with the `kinect-stereo-calib` application from the Mobile Robot Programming Toolkit (MRPT) [49], including the focal lengths (f_x, f_y) , the optical centers (c_x, c_y) , the pin-hole distortion parameters (k_1, k_2, p_1, p_2) , and the left-to-right relative pose. Calibration was done using a 7x10 classic checkerboard target pattern, capturing 20 pairs of images. The calibration file is available online. An example of the calibration tool, showing the result, is shown in Figure 3.

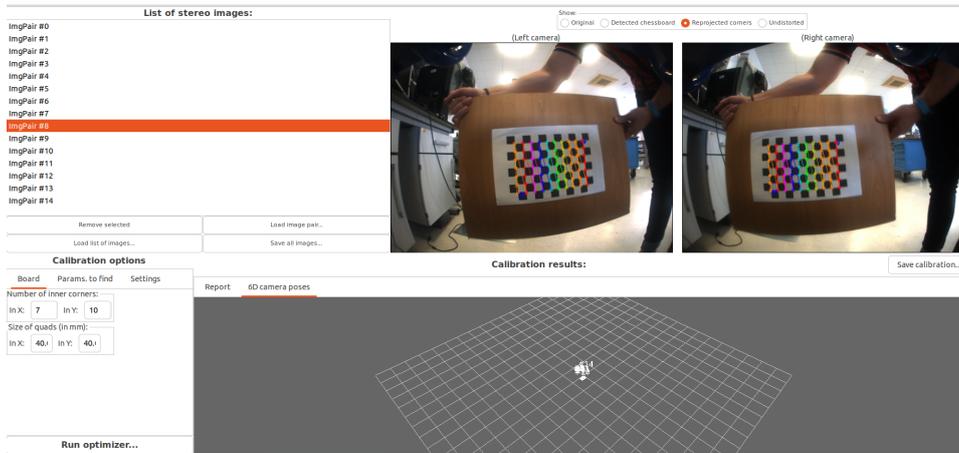


Figure 3: Screenshot of the intrinsic calibration application used for the stereo camera [49].

3.2.2 Extrinsic parameters

Extrinsic parameters define the relative poses of the different sensor frames of reference. The origin of coordinates is defined at the Velodyne, from which the other sensor poses are defined, as shown in Figure 2:

- Stereo camera: $x = \text{right}$, $y = \text{down}$, $z = \text{forward}$.
- LIDARs: $x = \text{forward}$, $y = \text{left}$, $z = \text{up}$.

The Velodyne sensor is taken as a local frame of reference for the platform.

4 GREENBOT dataset

This section describes the environment, data acquisition procedure, download method, and usage instructions.

4.1 Greenhouse and crop

The experiments took place at the Agroconnect facilities (which has received cofunding from the Ministry of Science, Innovation, and Universities in collaboration with the European Regional Development Fund (FEDER) under the grant program for acquiring cutting-edge scientific and technological equipment (2019)) in the Municipal District of La Cañada de San Urbano, Almería, located at $36^{\circ}50' \text{ N}$ and $2^{\circ}24' \text{ W}$, with an elevation of 3 meters above sea level and a slope in the terrain of 1 % in the North direction (see Figure 4).



Figure 4: IFAPA experimental greenhouse.

The greenhouse is in one of the most common styles in the region (Almería’s “raspa y amagado”), expands over 1850 square meters, and features a sturdy steel frame and a polyethylene covering. The greenhouse is arranged in an East-West ridge configuration to benefit from the natural ventilation from those two predominant wind directions in the region. A 2-meter wide central pathway in the greenhouse serves as the main road and leads to eleven aisles on each side. The aisles on the North side are 2 meters wide and 12.5 meters long, while the aisles on the South side are 2 meters wide and 22.5 meters long. Radiating from this central aisle are narrower secondary pathways, each only one meter wide, facilitating the smooth movement of mobile robotic units. The facility has two sections with various advanced systems tailored for precise crop management. These systems include natural ventilation from above (zenithal) and along the sides (lateral), an integrated air heating and cooling network, infrastructure for CO₂ enrichment, cutting-edge equipment for humidification and dehumidification, an advanced irrigation and fertilization system, as well as energy-efficient LED artificial lighting.

The greenhouse crop is tomato (*Lycopersicon esculentum*), and the plants are grown in coir bags, in rows oriented from north to south with a slope of 1%, as shown in Figure 5.



Figure 5: Example of two images acquired at the same position with different crop stages: the first one (left) after one month of plantation and the second one (right) after three months

4.2 Data acquisition

The trajectory that has been followed through the crop is shown in Figure 6. In order to distinguish the longer corridors from the shorter ones, the greenhouse is segmented into two sections, dividing it by the central corridor.

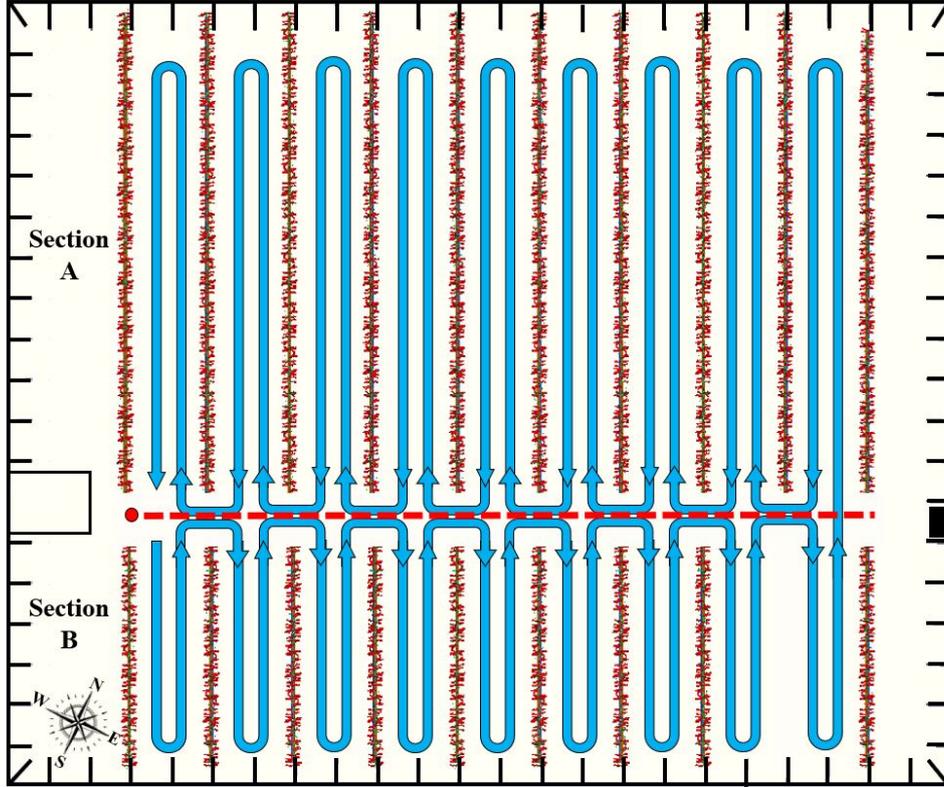


Figure 6: Trajectory taken during the data collection campaign in the greenhouse of the IFAPA centre, Almería. The red dot indicates the start and end of the route.

The platform is driven by the operator at an average velocity of 0.83 [m/s], who makes one pass until the end of each aisle and another pass on the way back, then switching to the next row. It should be noted that the ninth row in section B did not have plants at the time of recording this dataset. In order to carry out this data collection, the point cloud provided by the Velodyne and the Ouster LIDARs, the Ouster IMU, and the stereo images from the Bumblebee camera are stored in ROS 1 bag format. In the aisles located to the east and west, it has not been possible to make passes due to the impossibility of passing through the greenhouse’s supporting pillars.

In terms of data structure, there are a total of *nine sequences*, recorded between October 5th, 2022 and December, 1st 2022. Weekly recordings were made, subjected to different environmental conditions, lighting, ground changes, etc. The first two data files contain information from section B only (refere to Figure 6, as section A was temporarily inoperative). The third sequence contains section A, separated from section B. Finally, the remaining datasets contain a single *rosbag* that stores the entire greenhouse. Table 2 shows detailed and summarised information about the stored data, providing

information about the weather on the day of recording and the height of a pilot plant measured manually. The path length of each sequence is obtained from the SLAM-based reconstructions detailed later on, using the `evo_traj` tool [50]. Finally, the climatological variables of air temperature, humidity, and irradiance inside the greenhouse during the navigation, are extracted from a database of the University of Almería and also incorporated to the dataset.

Table 2: Description of each recorded segment

Sequence	Length [m]	Duration [s]	Section	Greenhouse condition	Description
2022_10_05	459.25	696	B	Temp: 27.01 [°C] Hum: 62.72 [%] Ir: 106.7 [W/m ²]	Sunny, morning plant height 0.96 [m]
2022_10_14	457.36	701	B	Temp: 23.78 [°C] Hum: 63.12 [%] Ir: 79.0 [W/m ²]	Cloudy, morning plant height 1.05 [m]
2022_10_19	1321.21	1432	A & B	Temp: 25.21 [°C] Hum: 75.83 [%] Ir: 87.9 [W/m ²]	Sunny, morning plant height 1.12 [m]
2022_10_26	1432.08	1463	A & B	Temp: 23.22 [°C] Hum: 60.09 [%] Ir: 66.4 [W/m ²]	Cloudy, morning plant height 1.35 [m]
2022_11_02	1233.87	1486	A & B	Temp: 15.84 [°C] Hum: 72.35 [%] Ir: 70.7 [W/m ²]	Cloudy, morning plant height 1.41 [m]
2022_11_09	1293.29	1532	A & B	Temp: 16.23 [°C] Hum: 62.7 [%] Ir: 82.7 [W/m ²]	Cloudy, morning plant height 1.53 [m]
2022_11_19	1332.58	1752	A & B	Temp: 17.45 [°C] Hum: 62.7 [%] Ir: 64.36 [W/m ²]	Sunny, morning plant height 1.60 [m]
2022_11_23	1428.45	1692	A & B	Temp: 15.28 [°C] Hum: 62.7 [%] Ir: 69.4 [W/m ²]	Cloudy, morning plant height 1.73 [m]
2022_11_30	1440.32	1730	A & B	Temp: 16.05 [°C] Hum: 62.7 [%] Ir: 74.4 [W/m ²]	Cloudy, morning plant height 1.85 [m]

4.3 Data structure

As already mentioned, the data are stored in *rosvbag* format as it allows storing the desired topics in the original format. This *rosvbag* keeps the ranging data from both LIDARs, the stereo images, and the IMU from the Ouster LIDAR.

Alternatively, LIDAR data are also provided in the *rawlog* format⁸.

⁸Description of the *rawlog* format available online.

This format can be easily parsed and published by ROS 2 packages such as `mrpt_rawlog` or the SLAM framework described later on.

5 SLAM suitability assesment

This section will assess the applicability of the dataset to build maps using a LIDAR SLAM framework. In particular, we used `mola_lidar_odometry` from the MOLA framework [51]. The approach belongs to the family of LIDAR odometry methods (i.e. without loop closure) with voxel-based raw point cloud representation. Experimental results from both LiDARs are presented separately in the following subsections.

5.1 Mapping - Velodyne VLP16

Figure 7 shows the result of the greenhouse mapped with the Velodyne VLP16 LiDAR for four different days.

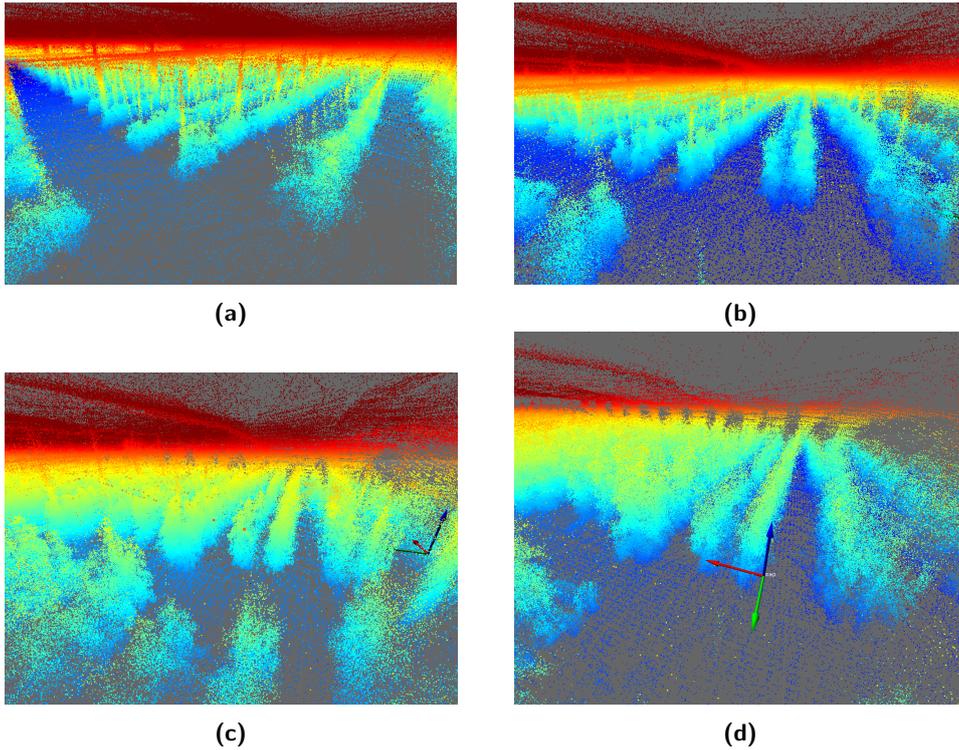


Figure 7: 3D mapping of IFAPA greenhouse with Velodyne VLP16 for sequeces (a) 2022-10-14, (b) 2022-10-26, (c) 2022-11-09. (d) 2022-11-30.

A significant visual difference in the state of the tomato plants can be observed in Figures 7(a)–(d), guaranteeing the systematic mapping of the

greenhouse. At the top of each image, phony LiDAR points are visible above the actual greenhouse top surface. These points are quite common for the VLP-16 LiDAR, which tends to detect phony points beyond the plastic surfaces of the greenhouse. The greenhouse soil is often uneven, with slopes varying between 1 and 2 %, depending on the area. In this type of mapping, being able to obtain a terrain variation model is of vital importance for robotics [52]. As output of the SLAM module we also obtain the estimated platform trajectory (in the standard *TUM* format) which is then analyzed using *evo_traj* as shown in Figure 8.

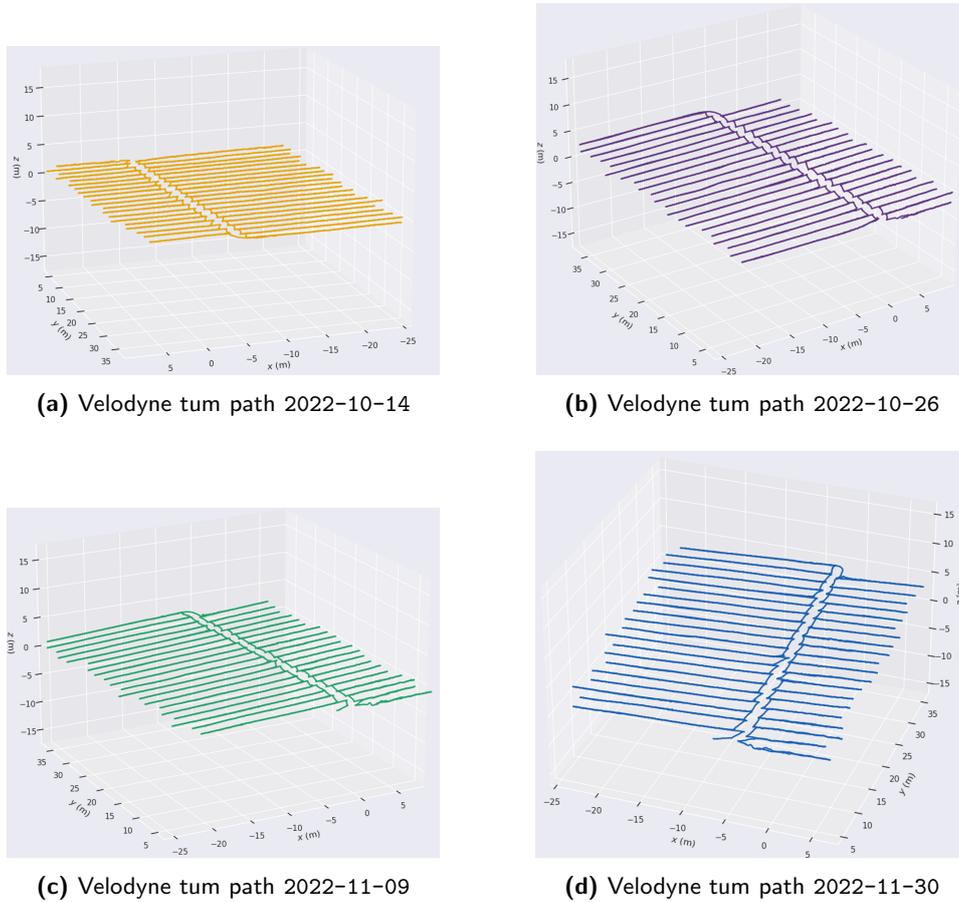


Figure 8: Path followed on different days with the Velodyne VLP16

It can be seen how the movement pattern is repeated every day. It can be seen how the pattern of movement is repeated every day. It is essential to mention that the trials start with the same orientation, as can be seen in Figures ??, ??, ?? and Figure ?? towards $+x$ and $+y$. Each axis behaves similarly irrespective of the day, although some exciting behaviors can be observed along each trajectory. To analyze this, the path followed by each

axis on November 30th, 2022 is shown in Figure 9.

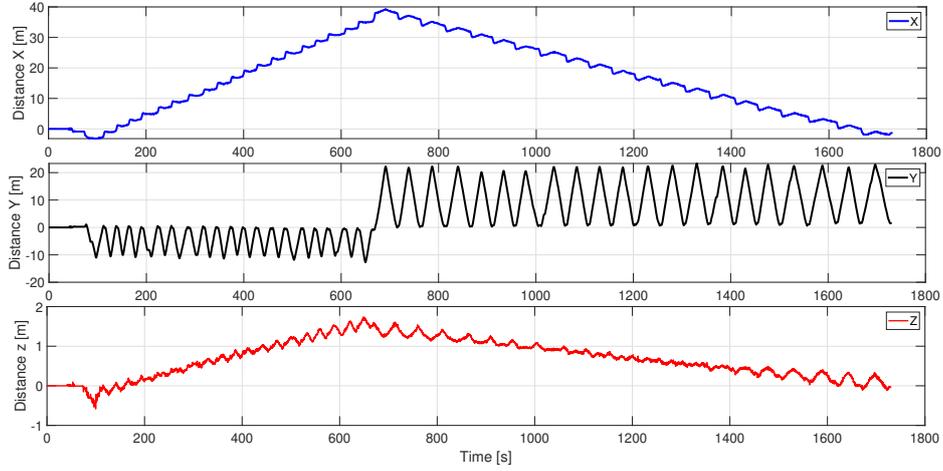


Figure 9: Trajectory estimated from the Velodyne VLP16 data for sequence 2022_11_30.

The sub-graphs corresponding to the displacement in x and y correspond to the displacement in the 2D plane, but the graph in z shows data relevant to the investigation. A noise is observed that comes from the deformed soil in the greenhouse, as well as a slope corresponding to the slope of the terrain itself, validated by the algorithm itself. Finally, Figures 1 and 10 show a final orthogonal view and plan of the result of the mapped 3D model, observing the dataset's quality.

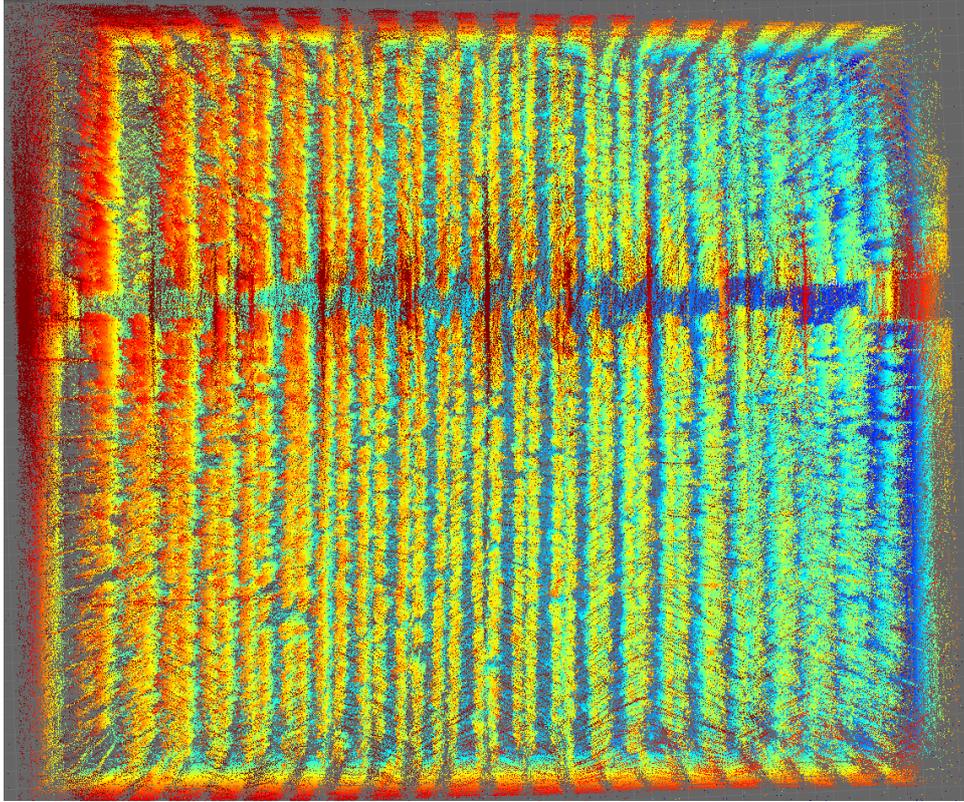
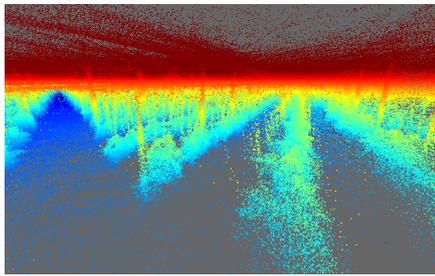


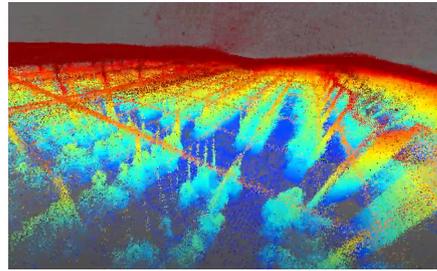
Figure 10: Mapping of the complete greenhouse, plan view with Velodyne VLP16

5.2 Mapping - Ouster OS0

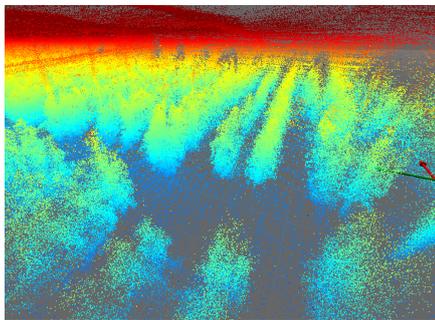
The same SLAM algorithm has been applied to the Ouster LiDAR data, obtaining the results shown in Figure 11. In this case, the mapping is shown for the same days as analysed in the previous section. Similarly, Figure 12 shows the result of the trajectories recorded by the Ouster OS0.



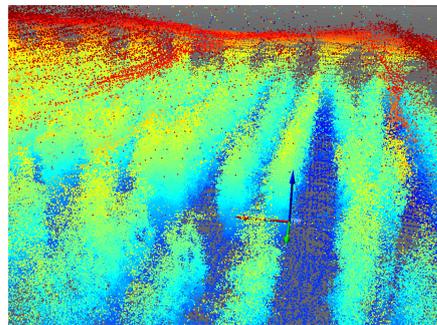
(a) Ouster MOLA Lidar odometry
2022-10-14



(b) Ouster MOLA Lidar odometry
2022-10-26

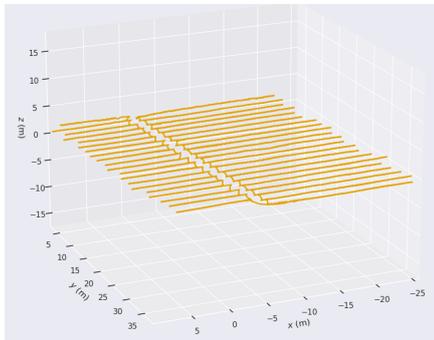


(c) Ouster MOLA Lidar odometry
2022-11-09

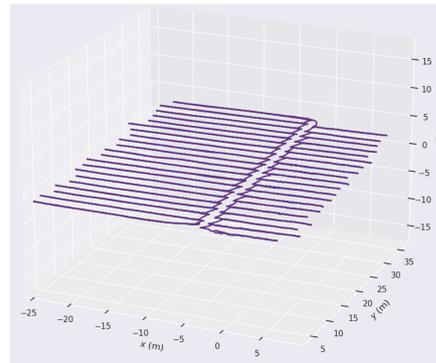


(d) Ouster MOLA Lidar odometry
2022-11-30

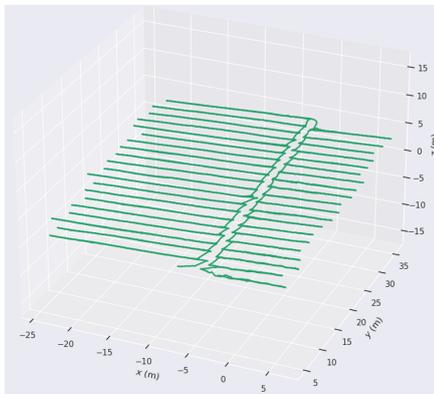
Figure 11: 3D mapping of IFAPA greenhouse with Ouster OS0



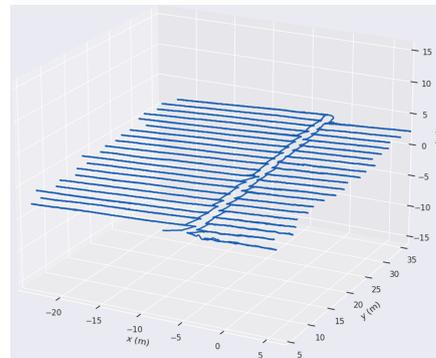
(a) Ouster path 2022-10-14



(b) Ouster path 2022-10-26



(c) Ouster path 2022-11-09



(d) Ouster path 2022-11-30

Figure 12: Path followed on different days with the Ouster OS0

A higher point cloud density can be observed as this LiDAR has twice point cloud rings than the VLP-16. In the same way, each of the segmented axes is projected, observing the same oscillatory behavior due to the uneven terrain.

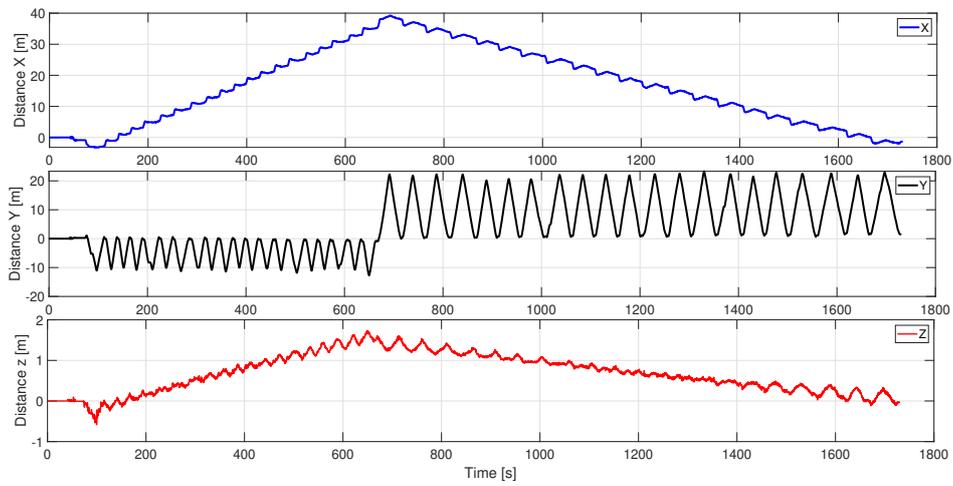


Figure 13: Trajectory of each axis for sequence 2022_11_30 with Ouster OS0.

The original and plan views are presented in Figures 14 and 15, obtaining an excellent result and exceeding the dataset expectations.

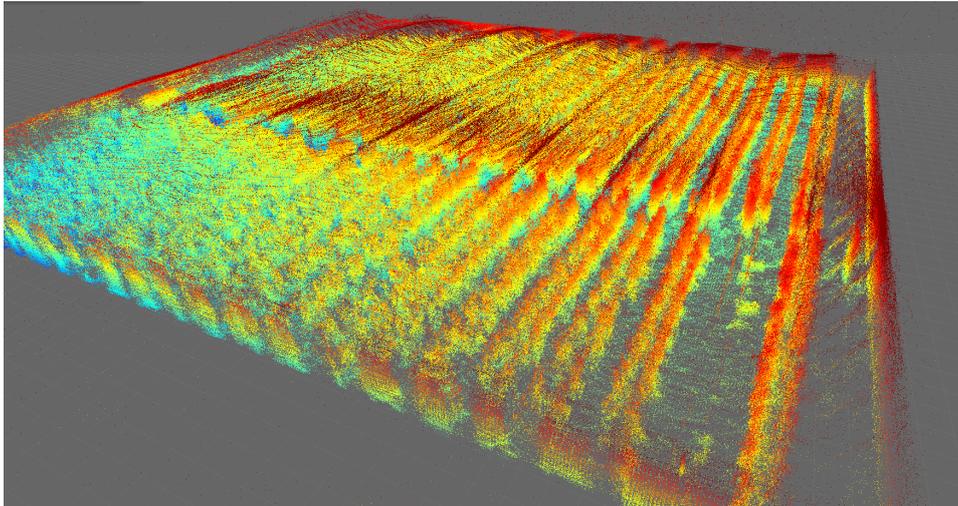


Figure 14: Mapping of the complete greenhouse, orthogonal view with Ouster OS0

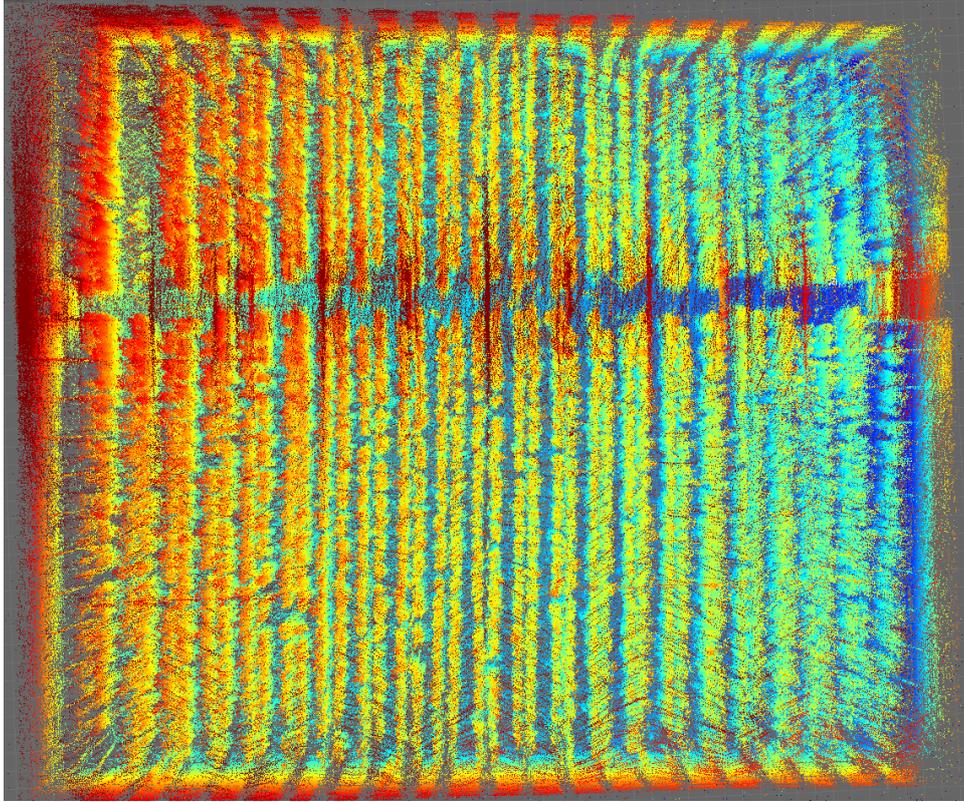


Figure 15: Mapping of the complete greenhouse, plan view with Ouster OS0

6 Conclusions and future work

A new dataset designed for challenging natural agricultural environments is presented here to drive the development of autonomous robots that can improve operations in the agricultural sector. This article details the data collection and describes the process employed. In addition, the dataset is validated using a novel SLAM technique to ensure its operability. The results indicate that the data quality is suitable for implementation in specific algorithms. Providing open datasets, such as the one presented here, is critically important to further research and implementation of autonomous robots in agricultural applications in greenhouses, given the unique characteristics of these environments, which include extreme lighting variations, changing weather conditions, and seasonal cycles.

In particular, this dataset allows the Community to develop and test its own SLAM algorithms in a novel agricultural environment, with particular emphasis on intensive greenhouse cultivation. The ability to create a map and orientate and navigate with precision is essential for all tasks a robot

can perform in a greenhouse (e.g., transport of vegetable crates, harvesting, spraying, data collection, etc).

The proposed framework is intended to be used as a database, which is unique for the time being. Data will continue to be collected from the greenhouse as all these sensors will be implemented in two robots of the University of Almería’s own. It is important to note that it is intended to collect data more frequently with the implementation of robots, seeking a more extensive dataset.

Apart from installing these sensors in real robots, the possibility of installing RGB-D cameras on both sides of the robots will be explored to obtain detailed information about the crop. The implemented cameras will also be improved to make it more robust to sudden changes in lighting, crop growth, etc. All this will be implemented in ROS 2 Humble, providing the robot with the latest technology in mobile robotics. Finally, these models can be used for other tasks, so this exciting contribution could be used to obtain the semantic information of the crop type, determining a model of plant growth in height and width.

This paper aims to present and share a unique dataset with the potential to assist researchers in developing new SLAM-based techniques and in evaluating a totally novel environment such as a greenhouse. It is outside the scope of this work to compare and analyze the performance of existing methods. The results described in Section 5 are provided as reference that researchers can use as starting point for future comparative evaluations.

Funding

This work has been partially financed by the ‘CyberGreen’ Project, PID2021-122560OB-I00, and the ‘AgroConnect.es’ infrastructure used to carry out this research, grant EQC2019-006658-P, both funded by MCIN/AEI/ 10.13039/501100011033 and by ERDF A way to make Europe. Also, author Fernando Cañadas-Aránega is supported by an FPI grant from Spanish Ministry of Science, Innovation and Universities.

Abbreviations

The following abbreviations are used in this manuscript:

CUDA	Compute Unified Device Architecture
DDR4	Double Data Rate type four Synchronous Dynamic Random-Access Memory
GPS	Global Positioning System
Hum	Humidity
IMU	Inertial Measurement Unit
Ir	Irradiance
LiDAR	Light Detection And Ranging
MOLA	Modular Optimization framework for Localization and mApping
MRPT	Mobile Robot Programming Toolkit
RAM	Random Access Memory
RGB	Red, Green and Blue
RGB-D	Red Green Blue Depth (color + Depth channels)
ROS	Robot Operating System
SLAM	Simultaneous Localization And Mapping
Temp	temperature
TIFF	Tagged Image File Format
UAL	University of Almería

References

- [1] Fan, M.; Shen, J.; Yuan, L.; Jiang, R.; Chen, X.; Davies, W.J.; Zhang, F. Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China. *Journal of experimental botany* **2012**, *63*, 13–24.
- [2] Evans, R.G.; Sadler, E.J. Methods and technologies to improve efficiency of water use. *Water resources research* **2008**, *44*.
- [3] Bindraban, P.S.; Dimkpa, C.; Nagarajan, L.; Roy, A.; Rabbinge, R. Revisiting fertilisers and fertilisation strategies for improved nutrient uptake by plants. *Biology and Fertility of Soils* **2015**, *51*, 897–911.
- [4] Bechar, A.; Vigneault, C. Agricultural robots for field operations: Concepts and components. *Biosystems Engineering* **2016**, *149*, 94–111.
- [5] Yan, Y.; Zhang, B.; Zhou, J.; Zhang, Y.; Liu, X. Real-Time Localization and Mapping Utilizing Multi-Sensor Fusion and Visual-IMU-Wheel Odometry for Agricultural Robots in Unstructured, Dynamic and GPS-Denied Greenhouse Environments. *Agronomy* **2022**, *12*, 1740.
- [6] Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets robotics: The kitti dataset. *The International Journal of Robotics Research* **2013**, *32*, 1231–1237.

- [7] Maddern, W.; Pascoe, G.; Linegar, C.; Newman, P. 1 year, 1000 km: The oxford robotcar dataset. *The International Journal of Robotics Research* **2017**, *36*, 3–15.
- [8] Pire, T.; Mujica, M.; Civera, J.; Kofman, E. The Rosario dataset: Multisensor data for localization and mapping in agricultural environments. *The International Journal of Robotics Research* **2019**, *38*, 633–641.
- [9] Chebrolu, N.; Lottes, P.; Schaefer, A.; Winterhalter, W.; Burgard, W.; Stachniss, C. Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields. *The International Journal of Robotics Research* **2017**, *36*, 1045–1052.
- [10] Marzoa Tanco, M.; Trinidad Barnech, G.; Andrade, F.; Baliosian, J.; Llofriu, M.; Di Martino, J.; Tejera, G. MAgro dataset: A dataset for simultaneous localization and mapping in agricultural environments. *The International Journal of Robotics Research* **2023**, p. 02783649231210011.
- [11] Rijswick, V. World Vegetable Map 2018. Global Trade Still Fruitful. Technical report, RaboResearch Food & Agribusiness, 2018.
- [12] Hick, G.W. World Greenhouse Vegetable Statistics- 2019 Updates Cuesta Roble Greenhouse Vegetable Consulting. Technical report, HortiDaily, USA, 2019.
- [13] Sanchez-Hermosilla, J.; Rodriguez, F.; Gonzalez, R.; Luis, J.; Berenguel, M. A Mechatronic Description of an Autonomous Mobile Robot for Agricultural Tasks in Greenhouses. In *Mobile Robots Navigation*; InTech, 2010. <https://doi.org/10.5772/9003>.
- [14] Sánchez-Hermosilla, J.; González, R.; Rodríguez, F.; Donaire, J. Mechatronic Description of a Laser Autoguided Vehicle for Greenhouse Operations. *Sensors* **2013**, *13*, 769–784. <https://doi.org/10.3390/s130100769>.
- [15] CAGPDR. Síntesis de la Campaña de Hortícolas protegidos de Almería - Campaña 2020/21. *Observatorio de la Consejería de Agricultura, Ganadería, Pesca y Desarrollo Sostenible de la Junta de Andalucía de Precios y Mercados del Sector hortícolas protegidos* **2021**.
- [16] Kondo, N.; Monta, M.; Noguchi, N. *Agricultural robots : mechanisms and practice*; Kyoto University Press, 2011; p. 348.
- [17] Rodríguez, F.; Moreno, J.C.; Sánchez, J.A.; Berenguel, M. Grasping in agriculture: State-of-the-art and main characteristics. In *Grasping in Robotics. Mechanisms and Machine Science*; Carbone, G., Ed.;

Springer London, 2013; Vol. 10, pp. 385–409. https://doi.org/10.1007/978-1-4471-4664-3_15.

- [18] Sánchez-Molina, J.A.; Rodríguez, F.; Moreno, J.C.; Sánchez-Hermosilla, J.; Giménez, A. Robotics in greenhouses. Scoping review. *Computers and Electronics in Agriculture (Under review)*.
- [19] Mandow, A.; Gomez-de Gabriel, J.; Martinez, J.L.; Munoz, V.F.; Ollero, A.; Garcia-Cerezo, A. The autonomous mobile robot AU-RORA for greenhouse operation. *IEEE Robotics & Automation Magazine* **1996**, *3*, 18–28.
- [20] Feng, Q.; Zou, W.; Fan, P.; Zhang, C.; Wang, X. Design and test of robotic harvesting system for cherry tomato. *International Journal of Agricultural and Biological Engineering* **2018**, *11*, 96–100. <https://doi.org/10.25165/j.ijabe.20181101.2853>.
- [21] Acaccia, G.; Michelini, R.; Molfino, R.; Razzoli, R. Mobile robots in greenhouse cultivation: inspection and treatment of plants. In Proceedings of the Memories. 1st International Workshop on Advances in Services Robotics. Bardolino, Italia. Citeseer, 2003.
- [22] Kurata, K. Cultivation of grafted vegetables II. Development of grafting robots in Japan. *HortScience* **1994**, *29*, 240–244.
- [23] González Sánchez, R.; Rodríguez Díaz, F.; Sánchez-Hermosilla López, J.; García Donaire, J. Navigation techniques for mobile robots in greenhouses. *Applied Engineering in Agriculture* **2009**, *25*, 153–165.
- [24] López-Gázquez, A.; Mañas-Alvarez, F.J.; Moreno, J.C.; Cañadas-Aránega, F. Navigation of a Differential Robot for Transporting Tasks in Mediterranean Greenhouses. In Proceedings of the International Symposium on New Technologies for Sustainable Greenhouse Systems (GreenSys), Cancún (México), 2023.
- [25] Li, W.; Saeedi, S.; McCormac, J.; Clark, R.; Tzoumanikas, D.; Ye, Q.; Huang, Y.; Tang, R.; Leutenegger, S. InteriorNet: Mega-scale multi-sensor photo-realistic indoor scenes dataset. *arXiv preprint arXiv:1809.00716* **2018**.
- [26] Ros, G.; Sellart, L.; Materzynska, J.; Vazquez, D.; Lopez, A.M. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas (USA), 2016; pp. 3234–3243.

- [27] Giubilato, R.; Stürzl, W.; Wedler, A.; Triebel, R. Challenges of slam in extremely unstructured environments: The dlr planetary stereo, solid-state lidar, inertial dataset. *IEEE Robotics and Automation Letters* **2022**, *7*, 8721–8728.
- [28] Yang, Y.; Tang, D.; Wang, D.; Song, W.; Wang, J.; Fu, M. Multi-camera visual SLAM for off-road navigation. *Robotics and Autonomous Systems* **2020**, *128*, 103505.
- [29] Sinha, R.; Quirós, J.J.; Sankaran, S.; Khot, L.R. High resolution aerial photogrammetry based 3D mapping of fruit crop canopies for precision inputs management. *Information Processing in Agriculture* **2022**, *9*, 11–23.
- [30] Lu, Y.; Young, S. A survey of public datasets for computer vision tasks in precision agriculture. *Computers and Electronics in Agriculture* **2020**, *178*, 105760.
- [31] Majdik, A.L.; Till, C.; Scaramuzza, D. The Zurich urban micro aerial vehicle dataset. *The International Journal of Robotics Research* **2017**, *36*, 269–273.
- [32] Blanco-Claraco, J.L.; Moreno-Duenas, F.A.; González-Jiménez, J. The Málaga urban dataset: High-rate stereo and LiDAR in a realistic urban scenario. *The International Journal of Robotics Research* **2014**, *33*, 207–214.
- [33] Bandini, F.; Jakobsen, J.; Olesen, D.; Reyna-Gutierrez, J.A.; Bauer-Gottwein, P. Measuring water level in rivers and lakes from lightweight Unmanned Aerial Vehicles. *Journal of Hydrology* **2017**, *548*, 237–250.
- [34] Miller, M.; Chung, S.J.; Hutchinson, S. The visual-inertial canoe dataset. *The International Journal of Robotics Research* **2018**, *37*, 13–20.
- [35] de Silva, R.; Cielniak, G.; Gao, J. Towards agricultural autonomy: crop row detection under varying field conditions using deep learning. *arXiv preprint arXiv:2109.08247* **2021**.
- [36] Liu, L.; Wang, R.; Xie, C.; Yang, P.; Wang, F.; Sudirman, S.; Liu, W. PestNet: An end-to-end deep learning approach for large-scale multi-class pest detection and classification. *IEEE* **2019**, *7*, 45301–45312.
- [37] Karam, S.; Nex, F.; Chidura, B.T.; Kerle, N. Microdrone-based indoor mapping with graph slam. *Drones* **2022**, *6*, 352.

- [38] Kirsanov, P.; Gaskarov, A.; Konokhov, F.; Sofiiuk, K.; Vorontsova, A.; Slinko, I.; Zhukov, D.; Bykov, S.; Barinova, O.; Konushin, A. Discoman: Dataset of indoor scenes for odometry, mapping and navigation. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, Venetian Macao (Macau), 2019; pp. 2470–2477.
- [39] Lee, T.j.; Kim, C.h.; Cho, D.i.D. A monocular vision sensor-based efficient SLAM method for indoor service robots. *IEEE Transactions on Industrial Electronics* **2018**, *66*, 318–328.
- [40] Martin, J.; Ansuategi, A.; Maurtua, I.; Gutierrez, A.; Obregón, D.; Casquero, O.; Marcos, M. A generic ROS-based control architecture for pest inspection and treatment in greenhouses using a mobile manipulator. *IEEE access* **2021**, *9*, 94981–94995.
- [41] Brostow, G.J.; Fauqueur, J.; Cipolla, R. Semantic object classes in video: A high-definition ground truth database. *Pattern Recognition Letters* **2009**, *30*, 88–97.
- [42] Duggal, V.; Sukhwani, M.; Bipin, K.; Reddy, G.S.; Krishna, K.M. Plantation monitoring and yield estimation using autonomous quadcopter for precision agriculture. In Proceedings of the 2016 IEEE international conference on robotics and automation (ICRA). IEEE, 2016, pp. 5121–5127.
- [43] Wang, Y.; Chen, X.; Liu, P. Statistical multipath model based on experimental GNSS data in static urban canyon environment. *Sensors* **2018**, *18*, 1149.
- [44] Ko, M.H.; Ryuh, B.S.; Kim, K.C.; Suprem, A.; Mahalik, N.P. Autonomous greenhouse mobile robot driving strategies from system integration perspective: Review and application. *IEEE/ASME Transactions On Mechatronics* **2014**, *20*, 1705–1716.
- [45] Feng, Q.; Zou, W.; Fan, P.; Zhang, C.; Wang, X. Design and test of robotic harvesting system for cherry tomato. *International Journal of Agricultural and Biological Engineering* **2018**, *11*, 96–100.
- [46] Matsuzaki, S.; Masuzawa, H.; Miura, J.; Oishi, S. 3D semantic mapping in greenhouses for agricultural mobile robots with robust object recognition using robots’ trajectory. In Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2018, pp. 357–362.
- [47] Marks, E.; Bömer, J.; Magistri, F.; Sah, A.; Behley, J.; Stachniss, C. BonnBeetClouds3D: A Dataset Towards Point Cloud-based Organ-

level Phenotyping of Sugar Beet Plants under Field Conditions. *arXiv preprint arXiv:2312.14706* **2023**.

- [48] Hernández, J.; Bonachela, S.; Granados, M.R.; López, J.C.; Magán, J.J.; Montero, J.I. Microclimate and agronomical effects of internal impermeable screens in an unheated Mediterranean greenhouse. *Biosystems Engineering* **2017**, *163*, 66–77.
- [49] Blanco-Claraco, J.L.; Koukis, N.; Laux, H.; Alice, N.; Briaies, J.; Monroy, J.; Jotaraul; Tarifa, M.J.; Sahdev, R.; Fernandez-Moral, E.; et al. MRPT/mrpt: Release of v2.11.7, 2024. <https://doi.org/10.5281/zenodo.10595287>.
- [50] Grupp, M. evo: Python package for the evaluation of odometry and SLAM. <https://github.com/MichaelGrupp/evo>, 2017.
- [51] Blanco-Claraco, J.L. A Modular Optimization Framework for Localization and Mapping. In Proceedings of the Robotics: Science and Systems, Freiburg im Breisgau (Germany), 2019; pp. 1–15.
- [52] GREENHOUSES, S.E.U.I.M. 4. Greenhouse climate control and energy use. *Good Agricultural Practices for greenhouse vegetable crops* **2013**, p. 63.