

Cooperative lane mapping for autonomous mobile robots in horticulture by aerial photogrammetry

Tjark Schütte^{a,b,*}, Volker Dworak^b, Redmond R. Shamshiri^b, Cornelia Weltzien^{a,b}

^a Technische Universität Berlin, Mechanical Engineering and Transport Systems, Straße des 17. Juni 135, 10623 Berlin, Germany

^b Leibniz-Institut für Agrartechnik und Bioökonomie e.V., Department of Agromechatronics, Max-Eyth-Allee 100, 14469 Potsdam, Germany

* tschuette@atb-potsdam.de

Abstract

11 Classical methods for navigation of autonomous mobile robots often fail in orchards with
12 **dense inter-row vegetation** such as **tall grass and protruding branches**. Since **robustness** and **ease**
13 of use are essential to the adoption of innovative technology, this can be a hindrance for the
14 development of new types of robotic services. **At the same time, the use of unmanned aerial**
15 **systems for monitoring applications in horticulture is increasing**. In this study we analyse whether
16 this data can be used to derive navigation maps and graphs of orchards, **even in conditions where**
17 **classical navigation methods fail**. To answer this question, we developed and analysed an
18 automated approach that derives maps with distinct levels of detail from unmanned aerial system
19 imagery for robot navigation in orchards. We test our approach on an apple tree orchard with
20 missing plants, and a mixed-berry orchard with strong inter-row vegetation and overarching
21 branches. We generate voxel maps, 2D grid maps and topological maps and introduce the concept
22 of **structure maps**. These enhance the existing row structure of the orchards. We show that our
23 method for grid map computation generalises to different crop types without retraining, while

24 providing grid maps that are on-par with, or better than current machine learning based approaches
25 with mean IoUs of 0.87 to 0.97. While the structure maps show limitations in areas without any
26 trees or plants, the derived paths show mean absolute errors equal to or less than 0.12 metres
27 compared to GNSS-based references, which is accurate enough for many agricultural applications.

28

29 **Keywords.** precision horticulture; UAV photogrammetry; AMR navigation; structure map

30

31 1. Introduction

32 The lack of skilled workers as well as the goal to reduce pesticide use (European
33 Commission, Directorate-General for Health and Food Safety, 2020) is driving the implementation
34 of ground-based robotics in horticulture for tasks that require physical contact like mechanical
35 weeding, mowing, pruning or harvesting. Even though the potential and the interest and
36 willingness to adopt these technologies are generally high, according to a review by Degieter et al.
37 (2023), practical application of robotics is still limited. In this review, lacking robustness and ease
38 of use are identified amongst the factors hindering the adoption of robotics.

39 In order for an autonomous mobile robot (AMR) to navigate successfully inside a given
40 environment, most approaches use maps for path planning and localisation. Occupancy grid maps,
41 already proposed in 1985 by Moravec and Elfes (1985), are still commonly used in modern mobile
42 robotic software (Macenski et al., 2020a). The basic idea of occupancy grid maps is to map the
43 environment on an evenly spaced two-dimensional grid. Extensions of this into the 3-dimensional
44 space are amongst others octree- or OpenVDB-based maps (Hornung et al., 2013; Macenski et al.,
45 2020b).

In agriculture, robots are often performing actions on the whole field or all crops, therefore the task of coverage planning is of special importance. Coverage planning for homogenous fields and fields with known obstacles is a well-researched problem (Santos et al., 2020b). In horticulture, however, planning a path along all crop rows can only be done if their position and that of the drivable lanes is known. The common approach of navigation for AMRs in orchard environments is therefore either to omit the mapping and leave the path planning to the human by using (i) a teach-and-repeat approach using GNSS (global navigation satellite system) receivers with real-time kinematics (RTK-GNSS) or optical sensors (Cox et al., 2023; Furgale and Barfoot, 2010). Here, the robot is driven manually through the complete orchard during first setup and repeats the exact path on consecutive runs; (ii) map the environment using a combination of simultaneous localization and mapping (SLAM) and manual driving for surveying (Pan et al., 2024; Shalal et al., 2015); or (iii) an in-depth manual surveying of the orchard using manual RTK-GNSS measurements or terrestrial laser scanners (Nørremark et al., 2008; Yagfarov et al., 2018). All three mentioned options are time consuming as they require continued human interaction, especially considering that orchards are dynamic environments and thus require repeated surveying. While extensions to SLAM, called active mapping, exist that use space exploration methods to let the robot map the environment autonomously, their applicability and robustness remain an open area of research even in structured indoor environments (Lluvia et al., 2021). Specifically, these methods will fail in situations where an online discrimination of drivable lanes, crops and other obstacles is not achieved, for example in situations where overarching branches ‘block’ the lanes. These challenges of high effort and lacking robustness can be a hindrance for the deployment of new robots, especially when offering the Robot-as-a-Service (RaaS) (Milella et al., 2024), where

68 the step of surveying, or in the example of Milella et al., building a digital field twin, has to be
69 repeated for every single customer and potentially on a single mission basis.

70 According to Degetier, the use of unmanned aerial systems (UAS), in contrast to AMRs,
71 is in many countries a lot higher. These systems are expected to play an increasingly important
72 role in precision horticulture, mainly in applications like crop monitoring and spraying, but also in
73 orchard management in general (Pflanz et al., 2018; Zhang et al., 2021). UAS provide large
74 amounts of data like RGB images, multi- or hyper-spectral images and light detection and ranging
75 (LiDAR) measurements (Amarasingam et al., 2022). With structure from motion, or in the case of
76 LiDAR through point cloud registration, large point clouds of the complete orchard can be
77 generated.

78 UAS can cover large areas in short time without interfering with potential obstacles on the
79 ground and can thus help overcome both problems of high manual effort and incompleteness or
80 failure in the mapping process for AMR. This could improve the robustness and ease of use,
81 especially during deployment and surveying of orchards. Accordingly, this type of UAS-AMR
82 collaboration is an active area of research. As an example, Sun et al. (2022) perform semantic
83 segmentation using an improved U-Net convolutional neural network (Ronneberger et al., 2015)
84 on orthoimages in order to plan straight paths for a robot inside a meadow orchard. Hobart et al.
85 (2020) use 3D point clouds from photogrammetry to derive tree parameters that can be used for
86 application maps for spraying in apple orchards. In their work on vineyard mapping, Comba et al.
87 (2018) introduce a likelihood score based on local normalised frequencies in point distances, to
88 detect and delineate vineyard parcels from other areas. In an extension by Biglia et al. (2022), a
89 similar approach, based on central point detection, is used to detect vine rows in 3D point clouds.
90 Similarly, Jurado et al. (2020) present a method for the detection of single vine trunks and missing

91 plants in high resolution point clouds of vineyards. While the works by Biglia et al. and Jurado et
92 al. are already quite comprehensive and show promising results for vineyards, they rely on specific
93 geometric features in the vines and the vine rows. Therefore the published approaches cannot be
94 expected to work with other orchards and crops. Specifically, in the hard case of a lot of
95 overarching branches, very sparse orchards, or very high inter-row grassing or vegetation, finding
96 central points that are the centre of the crop rows can be impossible. Additionally, when searching
97 for specific geometric features, these need to be adapted to every single crop type. This increases
98 the required effort in orchards with high variance or mixed crops. We want to close this gap on
99 missing scalability and generalization capability. In this work, we therefore develop and analyze a
100 **novel map generation** and **path planning method** that

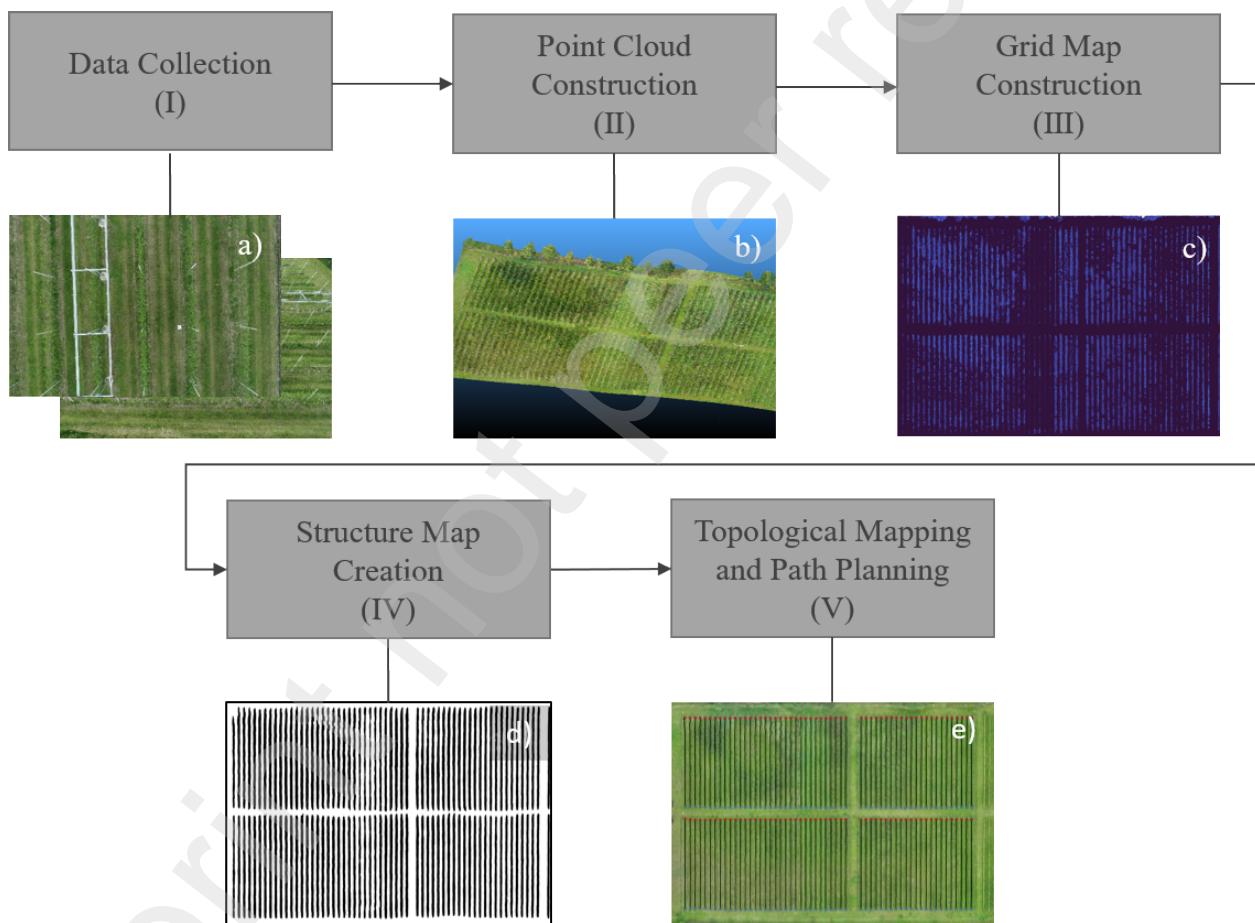
- 101 a) **does not rely on specific geometric features of the crops or a central point in the point cloud**
102 **density** and can be used for different crop types and mixed crop orchards
103 b) is **robust** to strong inter-row vegetation like overarching branches and grassing as well as
104 missing plants
105 c) **provides maps** for distinct levels of detail for the different tasks of navigation (high level
106 path planning, path following control and localization)

107

108 **2. Materials and Methods**

109 Overall, the aim of this works was to solve the problem of surveying, mapping, and
110 planning paths along all crop rows for AMRs in orchards, especially those with overgrown lanes,
111 mixed crops, and high variance. The developed method consisted of five main steps as presented
112 in Fig. 1. First, the **data collection** or **UAS surveying**, to capture **images of the orchard** (I). Secondly,
113 the **generation of a 3D point cloud of the orchard by photogrammetry** from these images (step II).

114 Then, the analysis and segmentation of this point cloud and its rasterisation to produce a 2D
 115 occupancy grid map of the orchard (step III and Fig. 1 c)). In the next step, this occupancy grid
 116 map was further processed to generate a structure map (Fig. 1 d)) of the environment with crop
 117 rows and drivable areas (IV). Lastly, based on this structure map, we derived a topological map
 118 that allowed for efficient path planning of the lanes through the crop rows in a graph structure (IV).
 119 We then evaluated the results of all these steps using GNSS measurements and manually labelled
 120 orthoimages as reference.



121
 122 Fig. 1. Workflow of the proposed method. Image a) shows exemplary UAV images b) a reconstructed 3D point
 123 map c) is the generated occupancy grid map, d) the structure map and e) the generated lane paths through
 124 the orchard displayed on the orthoimage.
 125

126 **2.1 Case Study Area and Data Collection**

127 For this study, two datasets were collected by UAS at two orchards in Brandenburg,
128 Germany. The **first dataset (berry field)** was collected at a 1.6-hectare area of a commercial farm
129 in Weggun, in October 2021. In this area, three types of berries (raspberries, red currant and black
130 berries) are grown. The **second dataset (apple orchard)** from May 2023 was collected at an
131 experimental apple orchard at the Fieldlab for Digital Agriculture (52.467349, 12.958571) of the
132 Leibniz-Institut für Agrartechnik und Bioökonomie. In this experimental orchard covering 0.76
133 hectare, the seven westernmost rows were complete, while the rest of the orchard was either
134 without trees or only partially planted with trees (see Fig. 2). A DJI Matrice 300 RTK with an L1
135 LiDAR and camera system (Da-Jiang Innovations Science and Technology Co., Ltd, Shenzhen,
136 China) was used for collecting the aerial images.



137 **Fig. 2. Orthoimages of the two surveyed orchards. Left: berry field in Weggun, right: experimental apple**
138 **orchard at the Fieldlab for Digital Agriculture.**

139
140 For reference measurements, two different RTK-GNSS system have been used. The HiPer
141
142 Pro (Topcon, Tokyo, Japan) with rover and base station was used in the berry field, and a Trimble

143 R12i (Trimble Inc., Westminster, USA) was used at the apple orchard. The positions of the slender
144 spindle- or trellis poles, as well as the ground control points (GCP) were measured. For all datasets,
145 EPSG 25833 (UTM 33N) was used as the spatial reference system (SRS) for all geo-spatial data.

146 The UAS survey of the apple orchard was divided into 3 separate flight missions: one nadir
147 flight and two oblique flights parallel to the crop rows with the sensor looking at the crops at a 45-
148 degree angle. The berry field dataset was collected in nadir view only. All flights in the case study
149 were conducted at a target altitude of 20 metres above ground and with an overlap of the images
150 of at least 60%. Before the actual flights, GCP markers were distributed in the surveyed area and
151 their position was measured with an RTK-GNSS system. The combination of the presented
152 surveying method and the selected sensor resulted in an average ground sampling distance of 5.3
153 and 7.1 millimetres for the berry field and apple orchard dataset, respectively. The difference in
154 ground sampling distance is attributed to the different flight profiles and the landscape profile in
155 the berry field that lead to a 2.8 metre lower average distance over ground based on estimated
156 camera position.

157 We used the IRUS, a mobile robot based on a remote-controlled platform (IRUS deltrak,
158 IRUS Motorgeräte GmbH, Burladingen-Salmendingen, Germany), to generate a second point
159 cloud of the apple orchard, at the same day and registered in the same SRS as the UAV-based point
160 cloud for comparison with the UAS-based point cloud. The IRUS is equipped with a dual-antenna
161 RTK-GNSS (ZED-F9P, u-blox, Thalwil, Switzerland), a 9-DOF inertial measurement unit (IMU)
162 (MTi-30-2A8G4, Xsens Technologies B.V., Netherlands) and a SICK MRS-1000 LiDAR sensor
163 (MRS1104C-011010, SICK AG, Germany) as well as wheel-tick sensors. It is localised based on
164 a GNSS, IMU and odometry fusion and registers point clouds from the LiDAR based on its current

165 position. A description of the process of point cloud collection and registration can be found in
166 Schütte et al. (2023).

167

168 **2.2 Point Cloud Construction**

169 After collection of the aerial images and the ground truth measurements of the GCPs, the
170 point cloud was **constructed** using the photogrammetry software **Metashape** (Agisoft, St.
171 Petersburg, Russia). Generally, the workflow followed the one suggested by Tinkham and Swayze
172 (2021) for tree canopy height analysis, with a few adaptations. Source preselection was chosen,
173 meaning that the images were pre-aligned based on the drones GNSS position. After image
174 alignment, GCPs were marked in all images in which they were visible. This was done to update
175 the geo-referencing of the model, as well as to optimise the image alignment. For the alignment
176 optimisation, only some of the GCPs were used, leaving out three (berry field) respectively four
177 (apple orchard) markers as check points that were later used to analyse the reconstruction results.
178 After image alignment, the key points detected in the alignment process were filtered. Key points
179 with a re-projection error higher than 0.15 metres and a reconstruction uncertainty with a ratio of
180 higher than 15 were excluded. Here, the re-projection error is the maximum error for a tie point. It
181 is computed as

$$max_i = \frac{|x'_i - x_i|}{s_i} \quad (1)$$

182 where x'_i are the coordinates of the point projection according to the estimated orientation
183 parameters on the i -th image in pixels, x_i are the measured point projection coordinates on the i -th
184 image in pixels and s_i the scale of the image that was used for measuring the coordinates of the

185 corresponding projection on the i -th image. The reconstruction uncertainty is computed as the ratio
186 of the largest to the smallest calculated error in the different dimensions. (Agisoft LLC, 2023)

187 In the next step, the dense 3D point cloud was computed using high quality settings,
188 meaning that no down- or up-sampling of the input images was performed during the
189 reconstruction process. For a detailed analysis of the influence of these settings refer to Tinkham
190 and Swayze (2021). After the process of point cloud construction, the point cloud was subsampled
191 using Octree based subsampling of the CloudCompare software (CloudCompare: 3D point cloud
192 and mesh processing software, 2024) with an octree depth of 12, giving a trade-off between detail
193 and memory requirements and resulting in point clouds with a density of 703 (berry field) and
194 1618 (apple orchard) points per square metre.

195

196 **2.3 Grid Map Construction**

197 To derive a grid map from our data, the point cloud was segmented into ground and off-
198 ground points using an algorithm that is based on the cloth simulation filter as described by Zhang
199 et al. (2016). It estimates a ground level for the complete point cloud by simulating particles of a
200 rigid cloth falling on the flipped point cloud. Points with a higher distance from the ground model
201 than a given height threshold (in our case the height of the tallest platform: 2.0 metres) were
202 removed. The segmented point cloud was rasterised by evaluating the per cell classes of a grid
203 with a given pixel width and a given height threshold of 0.2 metres. The resulting per cell ratios
204 of ground to non-ground points were then normalised and stored as a geotiff file. For the steps of
205 rasterisation and handling of the geo-information, apart from the Point Cloud Library (Rusu and
206 Cousins, 2011) the library GDAL (Rouault et al., 2024) with its C++ API was used.

207

208

2.4 Structure Map Creation

209 As can be seen in Fig. 7, not all paths crossing an area with crops are topologically wrong,
210 as some branches are overarching the complete lane in between two crop rows. In these cases, for
211 example for pruning or harvesting tasks, the information that a lane is going through is still vital.
212 To extract that information and to simplify the generated 2D map to the information required for
213 high-level or global path planning, we introduce the concept of a structure map. This map is a
214 special form of an occupancy grid, where obstacles are smoothed and small gaps are closed, while
215 maintaining and enhancing the overall structure of the orchard. To achieve this, there were three
216 steps involved; first, we rotated the image such that the main orientation, based on Hugh line
217 detection, was aligned with the image y-axis. Secondly, we applied 2D cross-correlation using a
218 series of kernels with a rectangular pattern that is based on an estimated minimum and maximum
219 lane- and crop width of the orchard. This method is based on the work done by Dvorak et al.
220 (2015), who introduced a visual crop lane tracking algorithm for automated tractors based on cross-
221 correlation. The result of the cross-correlation $g(x,y)$ is computed as the sum of the single kernel
222 operations, as shown in formula 2. Here, $f(x,y)$ is the original map at position (x, y) ; n is the
223 number of different kernels; $\omega_k(i,j)$ is the k-th kernel at position (i,j) , according to formula 3. a_k ,
224 b_k are the width and breadth of the k-th kernel in pixels, respectively.

$$g(x,y) = \sum_{k=1}^n \sum_{i=0}^{a_k-1} \sum_{j=0}^{b_k-1} \omega_k(i,j)f(x+i, y+j) \quad (2)$$

$$\omega_k(i,j) = \begin{cases} 1 & \text{if } i \% (c_k + r_k) \leq c_k \\ 0 & \text{if } i \% (c_k + r_k) > c_k \end{cases} \quad (3)$$

225 We applied n different variations of this kernel, based on the range of the estimated crop
226 and lane widths with one kernel for every 10 centimetres of range. This was done to account for

variability in crop width as well as varying crop row distances in orchards. It can also account for non-straight crop rows and lanes that increase the cross-section of lanes and crops. For the datasets in this work, the used ranges for the crop rows were [0.2...0.7] metres for both berry field and the apple orchard dataset. For the lane width, the ranges were [1.0...1.5] metres and [4.0...4.5] metres, respectively.

After correlation, a moving average filter was applied to the pixel rows of the resulting correlated image and then it is normalised globally as shown in Fig. 3b). To achieve the binary map shown in Fig. 3c), a threshold is applied using a dynamic auto-adjusting threshold, based on Otsu thresholding (Otsu, 1979) that maximises inter-class variance. After the thresholding operation, we apply image closing to close smaller gaps.

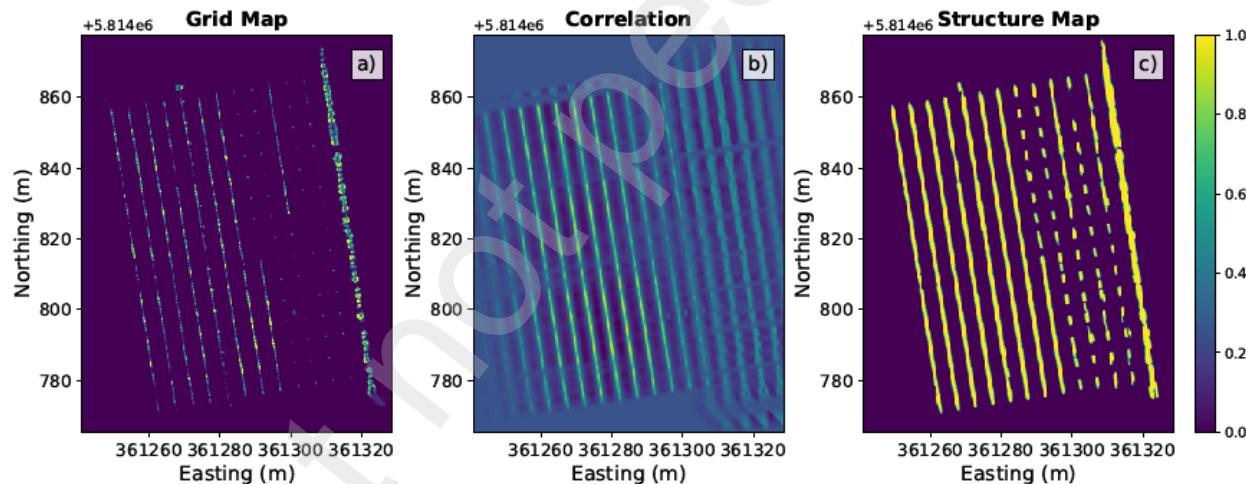


Fig. 3. The structure map process: a) grid map, values adjusted for visibility, b) normalized cross-correlation result and c) the final structure map.

2.5 Topological Mapping and Path Planning

To compute an easily searchable graph that encodes lanes and crop rows, we used the previously generated structure maps. We computed the distance field on these maps and searched for the rims of this field to generate the voronoi diagram. This structure was then converted into a

graph structure, leveraging the boost graph C++ library (Siek et al., 2001). We then applied graph pruning by removing short end segments and narrow passages as described in the work by Schwertfeger and Birk (2016). Specifically, end segments shorter than 2.0 metres, and segments with less than 0.5 metres width in the structure map, were removed. Paths in the resulting graph were then planned with Dijkstra's algorithm implementation in the boost graph library.

250

251 **2.6 Method Validation**

252 To validate our method, we evaluated each of the five main steps (data collection and point
253 cloud construction, grid map construction, structure map computation and topological mapping)
254 separately. For the evaluation of the UAS surveying and 3D reconstruction, we first compared
255 ground control point locations in the reconstructed 3D point cloud with their RTK-GNSS-
256 measured real world positions. Since this evaluation only gives an estimate of the accuracy of a
257 few distinct points, we additionally calculated the point distances to the LiDAR-based point cloud
258 of the apple orchard dataset. For evaluation of the 2D grid maps, we compared them to manually
259 labelled orthoimages. The labels used for the berry field dataset are displayed in Fig. 4, overlaid
260 on satellite imagery. Five classes were used in the labelling process: *drivable* for grass, lanes and
261 paths, *crops* for actual crops, *other vegetation* for other non-traversable vegetation like trees, and
262 *obstacles* for any other sort of obstacles including fences and poles. The fifth class, *crop row*,
263 stands for the space in the rows that is planned for, or dedicated to crops, even if there were no
264 crops present at the time of the measurement. Pixel based metrics were chosen since the topological
265 correctness of the maps is evaluated in the next step of the evaluation. Another argument for using
266 metrics that are based on direct pixel comparison, is the fact that we were interested in exact spatial

Fig. 4. Manual labels for the berry field dataset with the categories drivable in light-green, crops in purple, other vegetation in blue and obstacles in red. Labels are displayed as semi-transparent overlay on the orthoimage.



267 representations to be able to use the maps seamlessly with both environment sensor- and GNSS-
268 based localisation techniques and across different systems. We used accuracy, root mean squared
269 error (RSME), f1- and f2-score as well as mean intersection over Union (mIoU) for the evaluation
270 of the occupancy grid maps.

271 As reference for the structure maps, we used a map of crop rows derived by buffering the
272 RTK-GNSS measured pole positions, respectively their connecting lines, by the average crop
273 width. Finally, we evaluated the produced metric topological maps and paths generated by a)
274 comparing them to the centre lanes between the poles and b) by evaluating the relative area of
275 collision (rAoC) of the generated paths. The rAoC is calculated as the area of collision of a robot
276 with a given circular size with areas labelled as obstacle or crop, divided by the total area covered.
277 This is done by buffering the computed path p with half the given tool width. The buffered path is
278 the Minkowski sum of the original path with a circle c_w with the diameter w as shown in formula
279 4. The rAoC is then computed as the intersection of this buffered path with areas labelled as not
280 drivable or obstacle, divided by the total area covered (formula 5).

$$p_{buff_w} = p \oplus c_w \quad (4)$$

$$rAoC_w = \frac{p_{buff_w} \cap \text{not_drivable}}{p_{buff_w}} \quad (5)$$

281 This metric gives an indication of the quality of the path, even for lanes where no
282 completely collision-free path for the robot exists. As a reference for the rAoC values, we
283 additionally computed the metric for the reference centre lanes. This is relevant in orchards with
284 crops overlapping the drivable lanes, as the paths computed through these lanes might still be
285 considered correct. This is different from most robotic applications where generally the target is
286 to compute collision free paths.

287 We used two different metrics for the comparison of computed lanes or paths to reference
288 paths as shown in Table 1; the mean absolute error (MAE) that is computed as the mean Euclidean
289 distance between the reference points, and the root mean squared error (RMSE) for the same error.
290 The reference points are computed based on the RTK-GNSS measurements of the positions of the
291 poles in the orchards. These points were combined to one line per row with linear interpolation
292 between the poles. For every second row, the measured points were projected onto the next row
293 and the centre point of the shortest path between the point and the next row was used as a reference
294 point for the lane.

295

296 **Table 1. Metrics used for the evaluation of the computed paths.**

metric	description
rAoC	area of overlap with places labelled as non-drivable, divided by total area covered for given tool/robot width
MAE	mean absolute error (Euclidean distance) of the path to the reference path.
RMSE	root mean squared error (Euclidean distance) of the path to the reference path.

297

298 **3. Results**

299 All data processing after the Point Cloud Construction was performed on a mobile
 300 workstation with an octa core Intel i7-10875H 2.30GHz CPU with 32 GB DDR4 RAM memory
 301 running at 3200 MHz. It is equipped with an NVIDIA Quadro RTX 4000 mobile GPU that was
 302 used for the cross-correlation needed to construct the structure map. The average computing time
 303 for processing the point cloud of the berry field dataset with a density of about 700 points/m² and
 304 covering a 1.6 ha orchard was 2 minutes and 41 seconds for the complete process after point cloud
 305 construction. 39 seconds of this was used for constructing the grid map, 23 seconds for the
 306 construction of the structure map and 17 seconds for the computation of the navigation graph.
 307 Computing the 112 lanes of the berry field dataset from the graph made up the biggest part of the
 308 time with 82 seconds.

309

310 **3.1 Point Clouds**

311 As can be seen in Table 2, the RMSE of the check points (ground control points that were
 312 not used in the image alignment and georeferencing) is below 20 millimetres in all dimensions for
 313 the apple orchard. The dataset from Weggum (berry field), which was collected using a different

314 RTK-GNSS for surveying the GCPs and only contains images in nadir view shows higher errors
315 of 52 (y) to 386 (z) millimetres in all dimensions. It also covers a larger area than the apple orchard
316 dataset. When calculating the cloud-to-cloud-distances using the distance to the closest neighbour
317 in the LiDAR reference cloud for the apple orchard dataset, the average distance of the GNSS
318 referenced dataset to the UAV-based dataset is 105 ± 194 millimetres.

319 **Table 2. Number of GCPs used for alignment and georeferencing as well as check points used for validation**
320 **and the RMSE of the check points for both datasets**

Dataset	no. GCPs	no. check points	X error (mm)	Y error (mm)	Z error (mm)	XY error (mm)	Total (mm)
berry field	9	3	179	52	386	187	428
apple orchard	8	4	3	14	11	15	18

321

322 **3.2 Grid Maps**

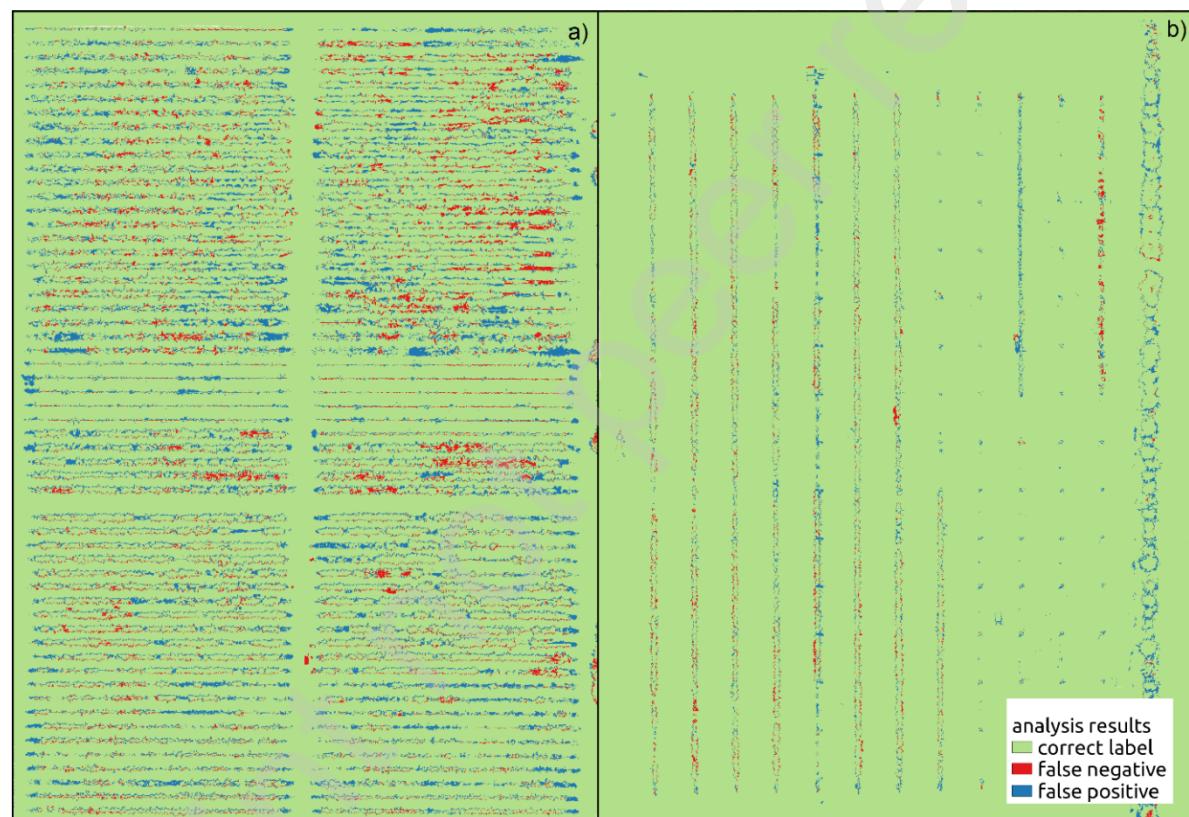
323 As a reference for the mapping process, the produced orthoimages were labelled in
324 different semantic categories as can be seen in Fig. 4. These labels were then sorted into the two
325 macro categories *drivable*: [drivable, crop row] and *non-drivable*: [other vegetation, crops,
326 obstacle]. The areas labelled as ‘crop row’ (and not at the same time as crop) were added to the
327 *drivable* category in this step, as they mark areas where no crops, other vegetation, or obstacles
328 were present at the time the dataset was collected. The evaluation of the grid maps was then carried
329 out based on the two macro labels. As can be seen in Table 3, the accuracies were 0.87 and 0.97
330 respectively for the grid maps. The f1 (f2)-scores were 0.77 (0.81) and 0.83 (0.86) for the same
331 datasets. Errors appear around the edges of the crop rows and in some missed parts of the crop
332 rows (red areas in Fig. 5).

333

334
335**Table 3.** Quantitative results for the label-based grid map and structure map analysis for the two datasets. The column ‘detected crop rows’ is the ratio of the crop rows that are clearly represented in the structure map.

Dataset	map type	mIoU	acc.	f1-score	f2-score	detected crop rows
berry field	grid	0.87	0.91	0.77	0.81	-
berry field	structure	0.87	0.90	0.86	0.93	1.0
apple orchard	grid	0.97	0.98	0.83	0.86	-
apple orchard	structure	0.89	0.94	0.73	0.77	0.54

336

**Fig. 5.** Confusion maps for both datasets a) berry field, b) apple. Overestimations (false positives) are displayed in blue, underestimations, or false negatives, in red. Correctly labelled areas in light green. Maps have been rotated and scaled for visibility.

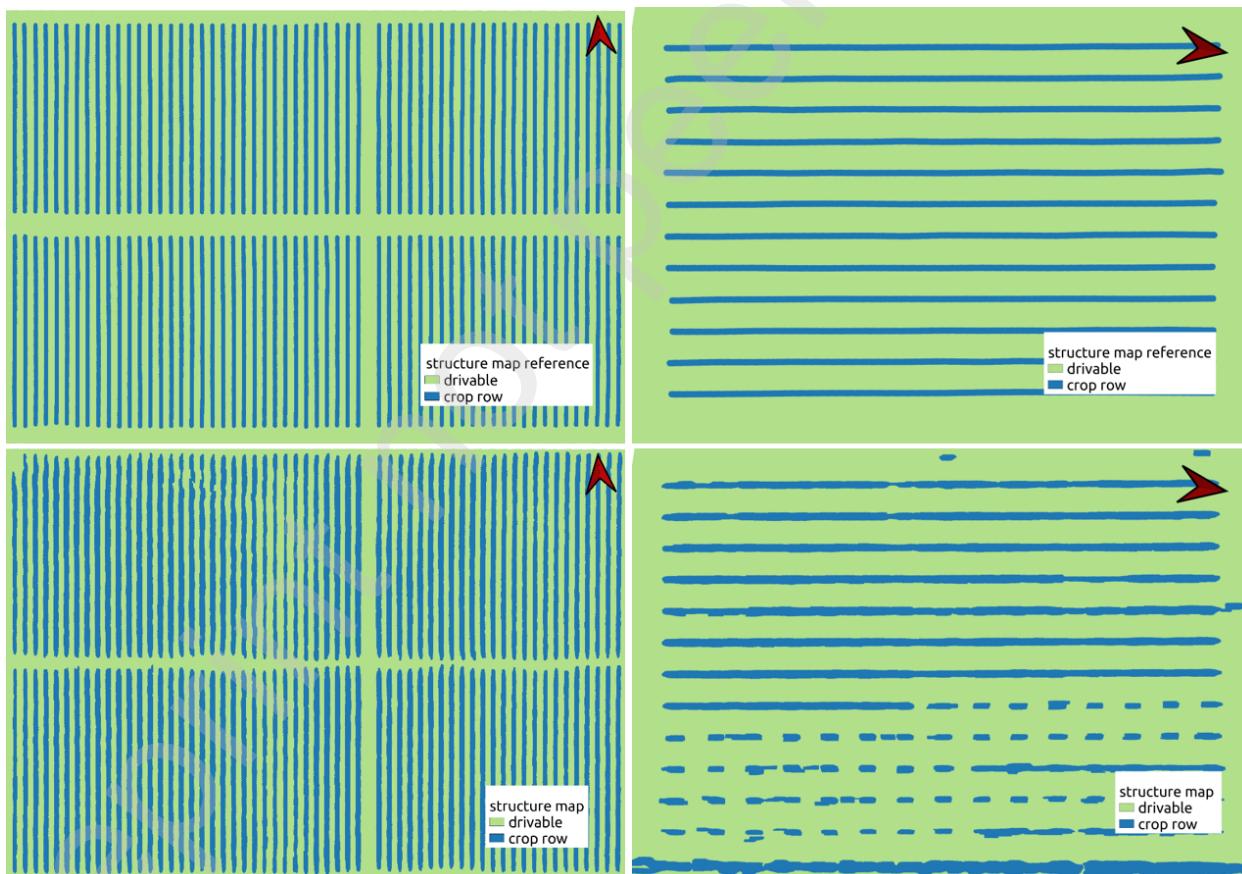
337

338

339 **3.3 Structure Maps**

340 As described in subsection 2.4, a metrically correct map is important for localisation and
341 obstacle avoidance but might not always be the most informative when it comes to global path
342 planning and navigation. We therefore introduced the concept of structure maps. As reference for
343 evaluation, we used the positions of the pole trellis, measured by RTK-GNSS, and expanded them
344 by the manually estimated mean crop row width.

345 The resulting references can be seen in Fig. 6 in the top row, with the computed structure
346 maps in the bottom row. It is visible that all crop rows were detected by the algorithm for the berry
347 field dataset (left column). When comparing the results of the structure maps with this reference,

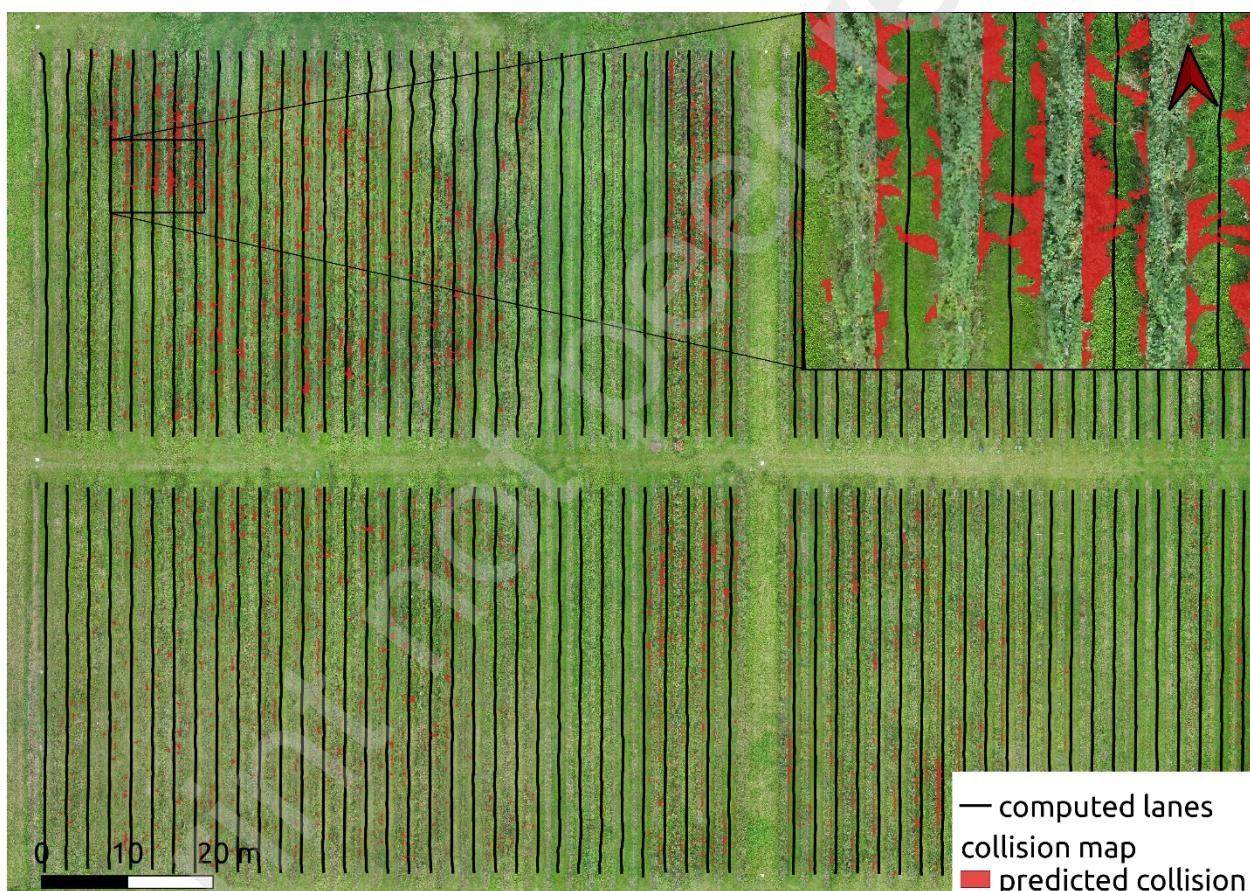


348 **Fig. 6. Reference maps derived from pole positions (top) and estimated structure maps (bottom) for both**
349 **datasets (berry field left, apple orchard right). Drivable areas in green, crop rows are marked in blue.**
350 the accuracy of the structure map of the apple orchard dataset is 8 percent lower than for the grid

351 map (0.97 vs 0.89) and similar for the berry field dataset (0.90 vs 0.91) with higher f1/f2 scores
352 for the structure map (compare Table 3). In the apple orchard dataset, the rows with plants were
353 detected correctly. For the rows with no, or only partial crop coverage, the algorithm did not close
354 the gaps between the poles, leading to the dotted rows visible in the lower half of the map in Fig.
355 6, bottom right. Also, a hedgerow next to the orchard is visible in the structure map.

356

357 **3.4 Topological Maps and Path Planning**



359 **Fig. 7.** Lanes generated for given start- and endpoints from the topological graph displayed on top of the
360 orthoimage of the mixed berry field in Weggum. The sub-image on the top right shows an enlarged area with
361 overarching branches. Areas that are labelled as crop in the reference data and would be hit by a circular
362 tool/robot with 1.5 metres diameter are highlighted in red.

363 As discussed in chapter 2.6, one primary metric for evaluating the topological maps and
364 generated lane paths is the rAoC (formula 5). We calculated this value for 500 randomly generated

365 paths in the topological map and a selected tool width of 1.5 m, 2.5 m and 4.5 meters. Additionally,
366 we also did this for all lanes of the orchards using given start- and endpoints. As can be seen in
367 Table 4, about 6 percent of the area covered by a 1.5 metre tool in the berry field dataset was
368 labelled as crop or obstacle. For the apple orchard dataset, the 4.5 metre tool width gave an overlap
369 of 1.1 percent, the 2.5 metre setting gave a 0.0 percent rAoC.

370

371 **Table 4. Mean rAoC and standard deviation for 500 randomly selected paths and all lanes with given start-**
372 **and endpoints, values in percent.**

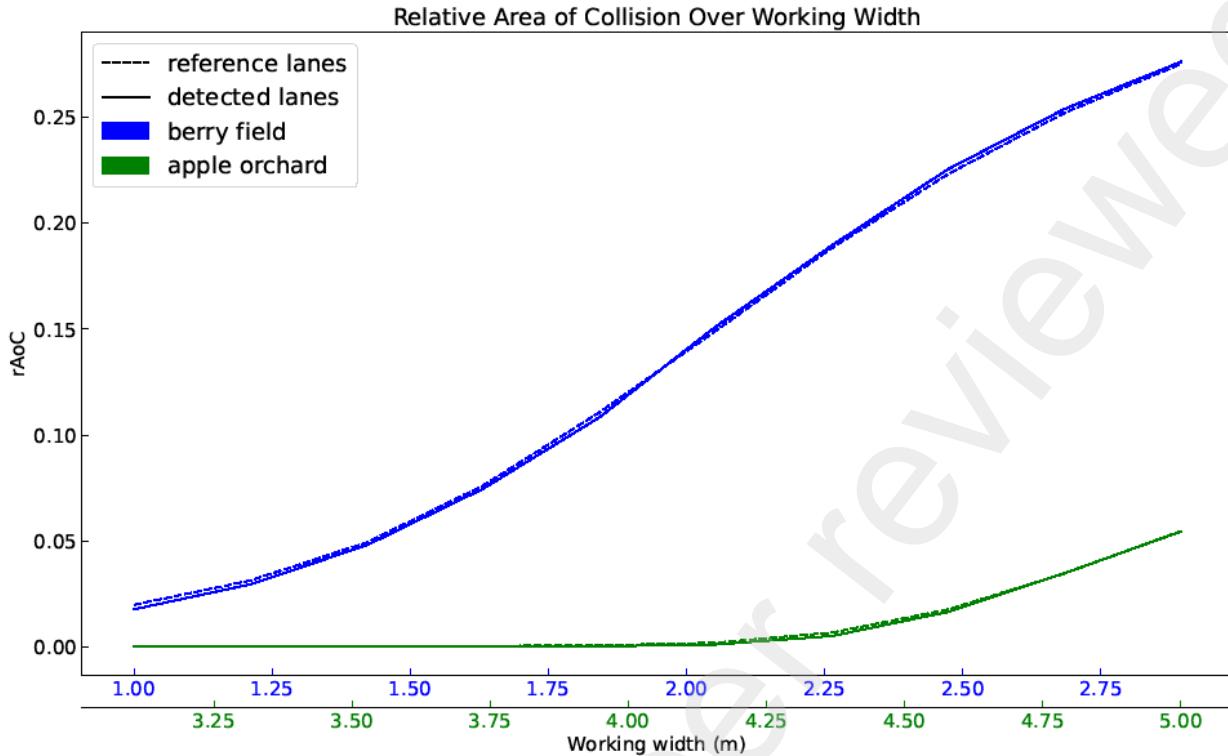
Dataset	rAoC _{rand,1.5} (%)	rAoC _{rand,2.5} (%)	rAoC _{lanes,1.5} (%)	rAoC _{lanes,2.5} (%)	rAoC _{lanes,4.5} ¹ (%)	rAoC _{rand,4.5} ¹ (%)
berry field	2.96±2.67	10.8±7.4	5.87±5.12	22.6±11.2	-	-
Apple orchard	0.1±0.3	0.1±0.2	0.1±0.2	0.0±0.1	1.1±0.7	0.5±0.5

373

374 To further analyse the errors and put the values in context, we additionally computed the rAoC
375 values for the reference centre lanes that are based on the measured pole positions. As can be seen
376 in Fig. 8, the rAoC values for the estimated paths only slightly differ from the reference paths,
377 even outperforming it for several configurations. The x-axis scales in this figure were adjusted for
378 the two datasets, to account for the unconventionally wide lanes in the apple orchard.

379

¹ 4.5 metre values were omitted for the Berry field dataset as a full overlap of the covered areas was achieved at 2.5 metre tool width



380

381 **Fig. 8. Ratio of crop area to total area covered (rAoC) over working width. Lower values mean less overlap**
 382 **with non-drivable areas. Dotted lines are manual reference centre lanes. X-axis scales adjusted to value ranges**
 383 **of the two datasets.**

384

385 As a last step we directly compared the computed lanes to the reference lanes. The MAE and
 386 RMSE were 0.09 to 0.12 and 0.12 to 0.17 metres respectively, with standard derivations of 0.02
 387 to 0.08.

388 **Table 5. Results of comparison with manual reference centre lanes based on pole positions.**

area	MAE (m)	RMSE (m)	number of points evaluated
berry field p0	0.12±0.03	0.17±0.05	189
berry field p1	0.11±0.04	0.16±0.07	192
berry field p2	0.10±0.02	0.13±0.02	219
berry field p3	0.09±0.02	0.12±0.03	214
apple orchard	0.09±0.06	0.12±0.08	185

389
390

391 **4. Discussion**

392 Using UAV images instead of satellite imagery, enables higher resolution for generating
393 the point cloud and grid maps. The overall accuracy in this case is, however, limited by the
394 accuracy of the reconstruction process. In our case, the error of check points reached a maximum
395 of 180 millimetres in the most important x and y dimensions. The xy errors on the berry field
396 dataset were more than 10 times higher than those of the apple orchard. Reasons for that might be
397 the fact that only nadir images of the berry field were taken and that it only had one GCP more
398 than the apple orchard dataset, while covering about twice the area. It might also be caused by less
399 clear features being present in the images, leading to worse image alignment. We show that mean
400 absolute errors of the generated paths to the reference are below the error margins of the point
401 cloud construction itself for this dataset. An improvement compared to the work by Biglia et al.
402 (2022) is that our results do not show the same increase in error when comparing to an absolute
403 (GNSS-based) reference, for both datasets.

404 The resulting accuracies of the grid maps and the predicted lanes are comparable to similar,
405 classical point cloud analysis results while the processes are computationally more efficient as they
406 work on the complete point cloud instead of a small subset at a time. Even though not on par with
407 deep-learning-based methods that segment 2D orthoimages in terms of speed, the low computation
408 time opens the opportunity to use the method in online or on-demand applications. The former
409 implies that the point cloud can be provided online e.g. using UAS-mounted LiDAR with RTK-
410 GNSS. An advantage of the point cloud-based approach is the option to additionally provide point
411 cloud maps like octree- or OpenVDB-based maps (Hornung et al., 2013; Macenski et al., 2020b).
412 Additionally, unlike machine learning approaches, our method does not require retraining for new

413 crop types. This was shown by segmenting diverse types of crops and orchards with the same
414 algorithms that only required changes in the lane and crop width range parameters.

415 The results in terms of grid map accuracies and mIoU of our method (acc.: 0.91-0.98, mIoU:
416 0.87-0.97, compare Table 3) are on-par with, or outperform those with Support Vector Machine
417 (accuracy 0.88-0.89) and Semantic Segmentation Suite (accuracy 0.73-0.87) by Santos et al.
418 (2020a) and are comparable to the results by Sun et al. (2022) that showed an accuracy of 0.93 and
419 mIoU of 0.88.

420 The higher f2- than f1-scores in our results indicate that the algorithm overestimates
421 obstacles and crops. This effect can also be seen in Fig. 5 were overestimation of the crop rows
422 can be observed around the edges. Reasons for that are inaccuracies in the point cloud
423 reconstruction and the manual labelling, but also the fact that tall grass next to the crops, which is
424 not labelled as crop in the reference dataset, cannot be differentiated from crops by the algorithm.

425 For the structure map, while working for both datasets, the apple orchard dataset reveals
426 the limitations of the structure extraction. In the areas without trees, the algorithm is not able to
427 close the gaps between the poles, leading to it only completely enhancing the rows in the fully
428 planted areas. Also, a hedge that is growing next to the orchard in a similar distance is included in
429 the structure map. This leads to the reduced performance metrics for the structure map of the apple
430 dataset compared to the grid map. The mean absolute errors and RMSEs of the resulting paths, are
431 still within the error margins of the 3D reconstruction, even including the parts were the structure
432 map showed errors. They are comparable to the errors given by Biglia et al. for the crop row
433 localisation and outperform those posted by Sun et al. One major difference to our results is that,
434 except for one parcel in the analysis by Biglia et al. that shows 0.2 metres higher Euclidean
435 distances, the reported accuracies in the literature do not regard errors introduced by the 3D

436 reconstruction as the references were marked in the 3D data or orthoimages and not measured
437 independently. As we compare to independently measured GNSS positions, these errors are
438 accounted for in our path planning results.

439

440 **5. Conclusions**

441 The developed method enables the automated generation of accurate 3D-, grid- and structure
442 maps from UAS data even in non-trivial orchards with missing plants and strong inter-row
443 vegetation. It does so faster than other point cloud-based approaches at equal or better performance
444 compared to similar studies. While the approach of structure mapping using cross-correlation
445 shows limitations with very sparse orchards, the generated paths are within reasonable error
446 margins for navigation tasks inside orchard rows. We could therefore confirm the hypothesis,
447 showing that it is possible to derive maps for in-lane navigation in orchards with lanes that are
448 completely overgrown and that have a high variance with the developed method. This enables fast
449 and automated mapping with minimal manual intervention for a wider range of orchards. It helps
450 reduce effort during the deployment of robots while providing full coverage. Additionally, the
451 maps can be used to aid online decision making. While some work remains to be done on the edge
452 cases of completely missing rows, the presented method achieves the objectives of improving
453 generalisation and scalability and thus improving and simplifying the surveying of orchards for
454 AMR navigation.

455 One remaining question is the suitability of the point cloud and 2D grid maps for localisation
456 that will be analysed in a following study. As future work, we will also examine the combination
457 of the presented point cloud analysis with deep learning-based segmentation to improve the

458 performance in the non-trivial cases and extract additional semantic information from the UAS

459 data.

460

461 **Funding**

462 This research was partially funded by the project “SMART4ALL - Selfsustained Cross-
463 Border Customized Cyberphysical System Experiments for Capacity Building among European
464 Stakeholders“ (Horizon 2020 – GA No 872614).

465

466 **Acknowledgements**

467 The authors would like to thank “Bauernhof Weggun GbR” and the team of the “Fieldlab for
468 Digital Agriculture” for hosting the experiments, as well as the team of the “Technical Support
469 Agromechtronics” of the Leibniz Institut für Agrartechnik und Bioökonomie for conducting the
470 drone flights.

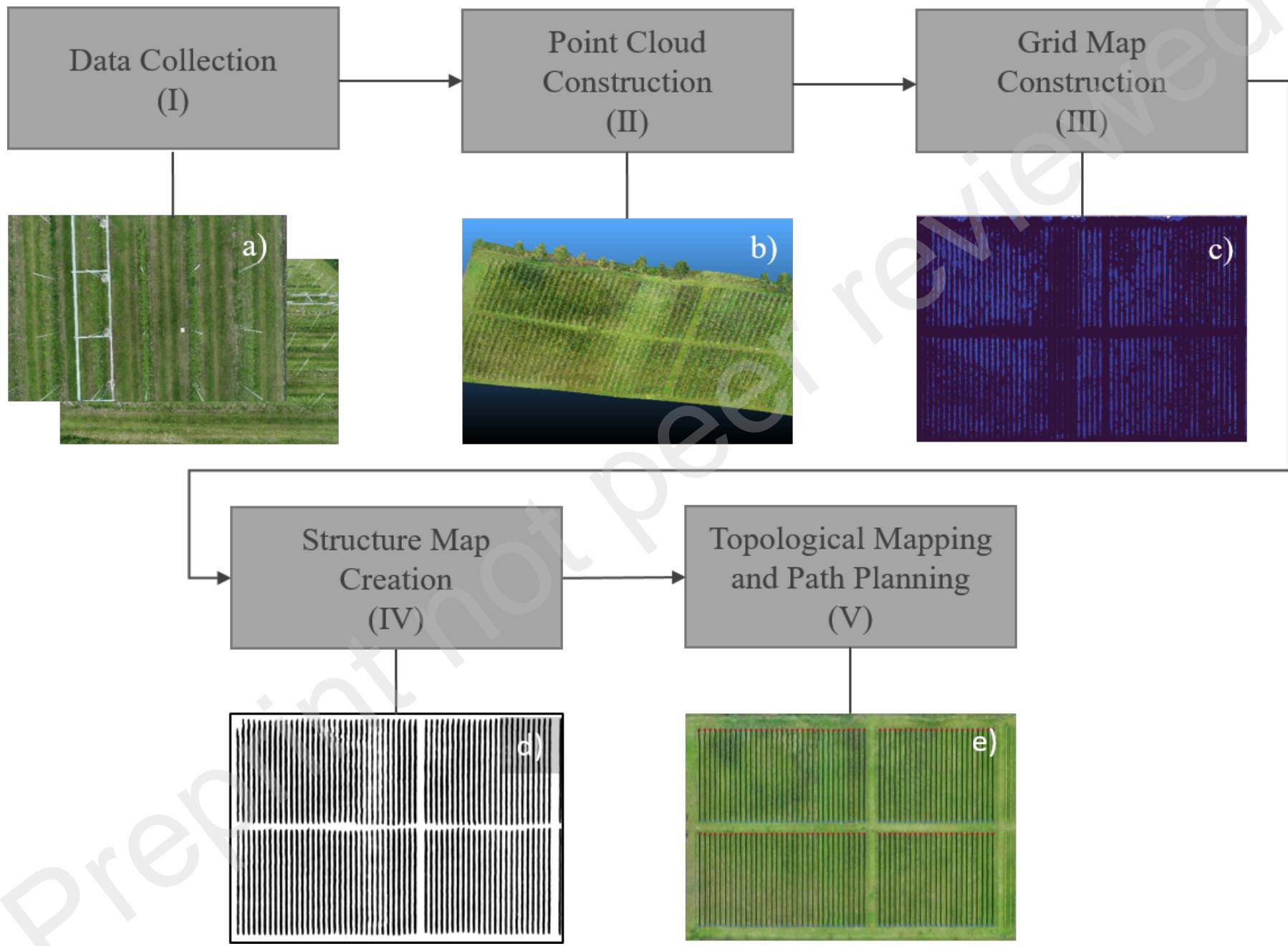
471

472 **References**

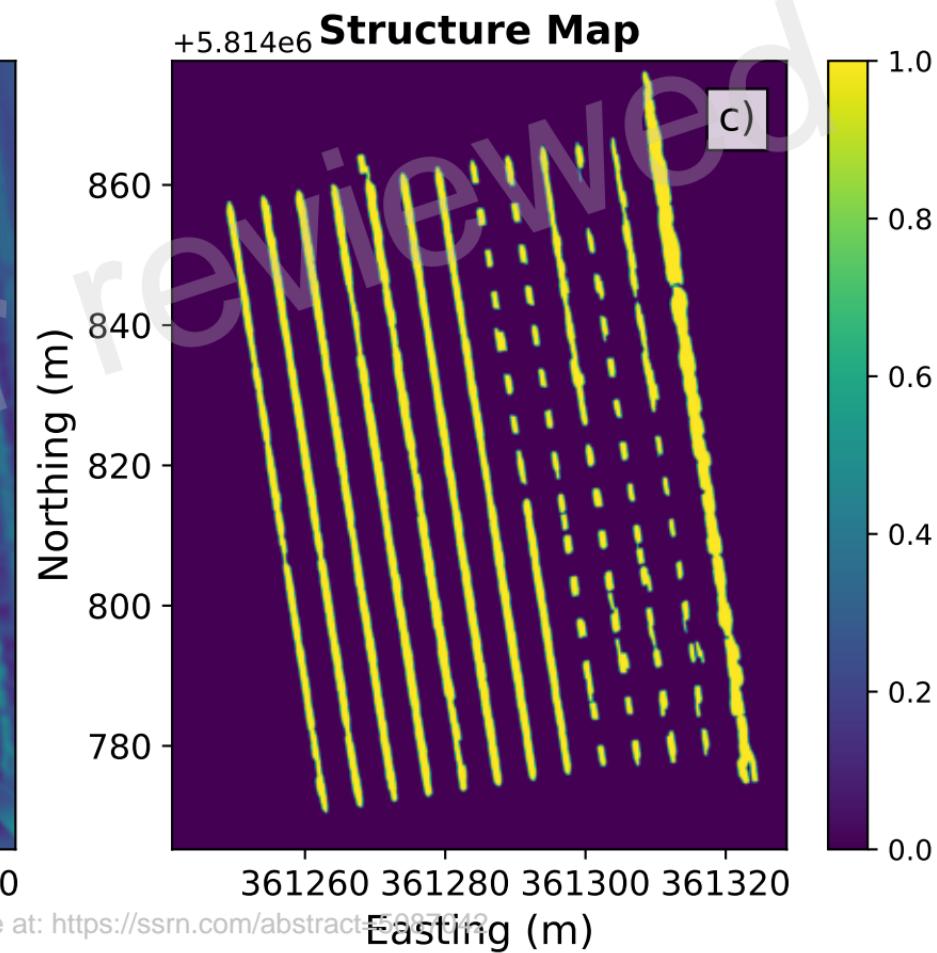
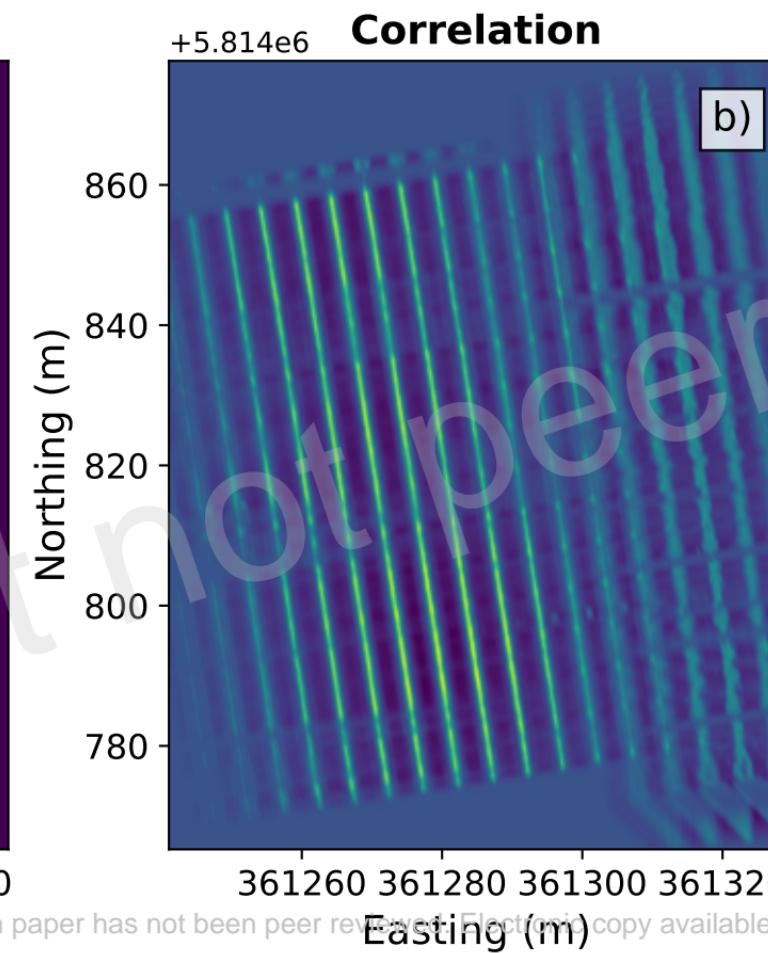
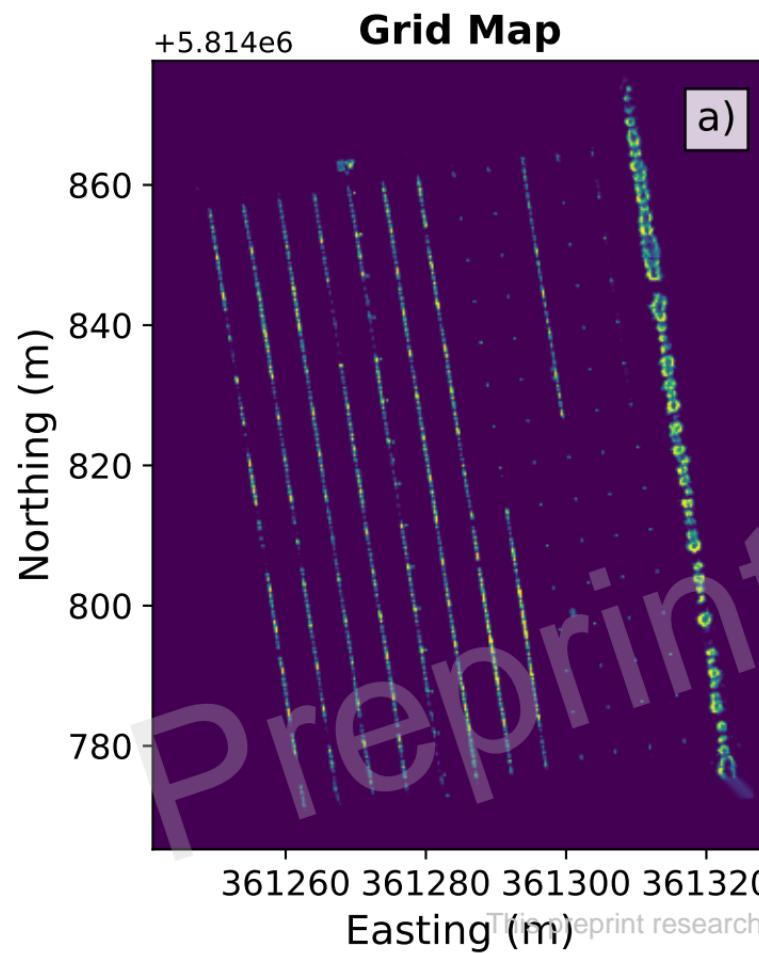
- 473 Agisoft LLC, 2023. Agisoft Metashape User Manual: Professional Edition, Version 2.1.
474 https://www.agisoft.com/pdf/metashape-pro_2_1_en.pdf (accessed 31 January 2024).
- 475 Amarasingam, N., Ashan Salgadoe, A.S., Powell, K., Gonzalez, L.F., Natarajan, S., 2022. A
476 review of UAV platforms, sensors, and applications for monitoring of sugarcane crops.
477 *Remote Sensing Applications: Society and Environment* 26, 100712.
478 <https://doi.org/10.1016/j.rsase.2022.100712>.
- 479 Biglia, A., Zaman, S., Gay, P., Ricauda Aimonino, D., Comba, L., 2022. 3D point cloud density-
480 based segmentation for vine rows detection and localisation. *Computers and Electronics in*
481 *Agriculture* 199, 107166. <https://doi.org/10.1016/j.compag.2022.107166>.
- 482 2024. CloudCompare: 3D point cloud and mesh processing software. GPL software.
- 483 Comba, L., Biglia, A., Ricauda Aimonino, D., Gay, P., 2018. Unsupervised detection of
484 vineyards by 3D point-cloud UAV photogrammetry for precision agriculture. *Computers and*
485 *Electronics in Agriculture* 155, 84–95. <https://doi.org/10.1016/j.compag.2018.10.005>.
- 486 Cox, J., Tsagkopoulos, N., Rozsypálek, Z., Krajník, T., Sklar, E., Hanheide, M., 2023. Visual
487 teach and generalise (VTAG)—Exploiting perceptual aliasing for scalable autonomous
488 robotic navigation in horticultural environments. *Computers and Electronics in Agriculture*
489 212, 108054. <https://doi.org/10.1016/j.compag.2023.108054>.
- 490 Degieter, M., Steur, H. de, Tran, D., Gellynck, X., Schouteten, J.J., 2023. Farmers' acceptance of
491 robotics and unmanned aerial vehicles: A systematic review. *Agronomy Journal* 115, 2159–
492 2173. <https://doi.org/10.1002/agj2.21427>.
- 493 Dworak, V., Huebner, M., Selbeck, J., 2015. Precise Navigation of Small Agricultural Robots in
494 Sensitive Areas with a Smart Plant Camera. *Journal of Imaging* 1, 115–133.
495 <https://doi.org/10.3390/jimaging1010115>.
- 496 European Commission, Directorate-General for Health and Food Safety, 2020. Communication
497 from the Commission to the European Parliament, the Council, the European Economic and
498 Social Committee and the Committee of the Regions: A Farm to Fork Strategy for a fair,
499 healthy and environmentally-friendly food system.
- 500 Furgale, P., Barfoot, T.D., 2010. Visual teach and repeat for long-range rover autonomy. *Journal*
501 *of Field Robotics* 27, 534–560. <https://doi.org/10.1002/rob.20342>.
- 502 Hobart, M., Pflanz, M., Weltzien, C., Schirrmann, M., 2020. Growth Height Determination of
503 Tree Walls for Precise Monitoring in Apple Fruit Production Using UAV Photogrammetry.
504 *Remote Sensing* 12, 1656. <https://doi.org/10.3390/rs12101656>.
- 505 Hornung, A., Wurm, K.M., Bennewitz, M., Stachniss, C., Burgard, W., 2013. OctoMap: An
506 Efficient Probabilistic 3D Mapping Framework Based on Octrees. *Autonomous Robots*.
507 <https://doi.org/10.1007/s10514-012-9321-0>.
- 508 Jurado, J.M., Pádua, L., Feito, F.R., Sousa, J.J., 2020. Automatic Grapevine Trunk Detection on
509 UAV-Based Point Cloud. *Remote Sensing* 12, 3043. <https://doi.org/10.3390/rs12183043>.
- 510 Lluvia, I., Lazcano, E., Ansuategi, A., 2021. Active Mapping and Robot Exploration: A Survey.
511 *Sensors* 21. <https://doi.org/10.3390/s21072445>.

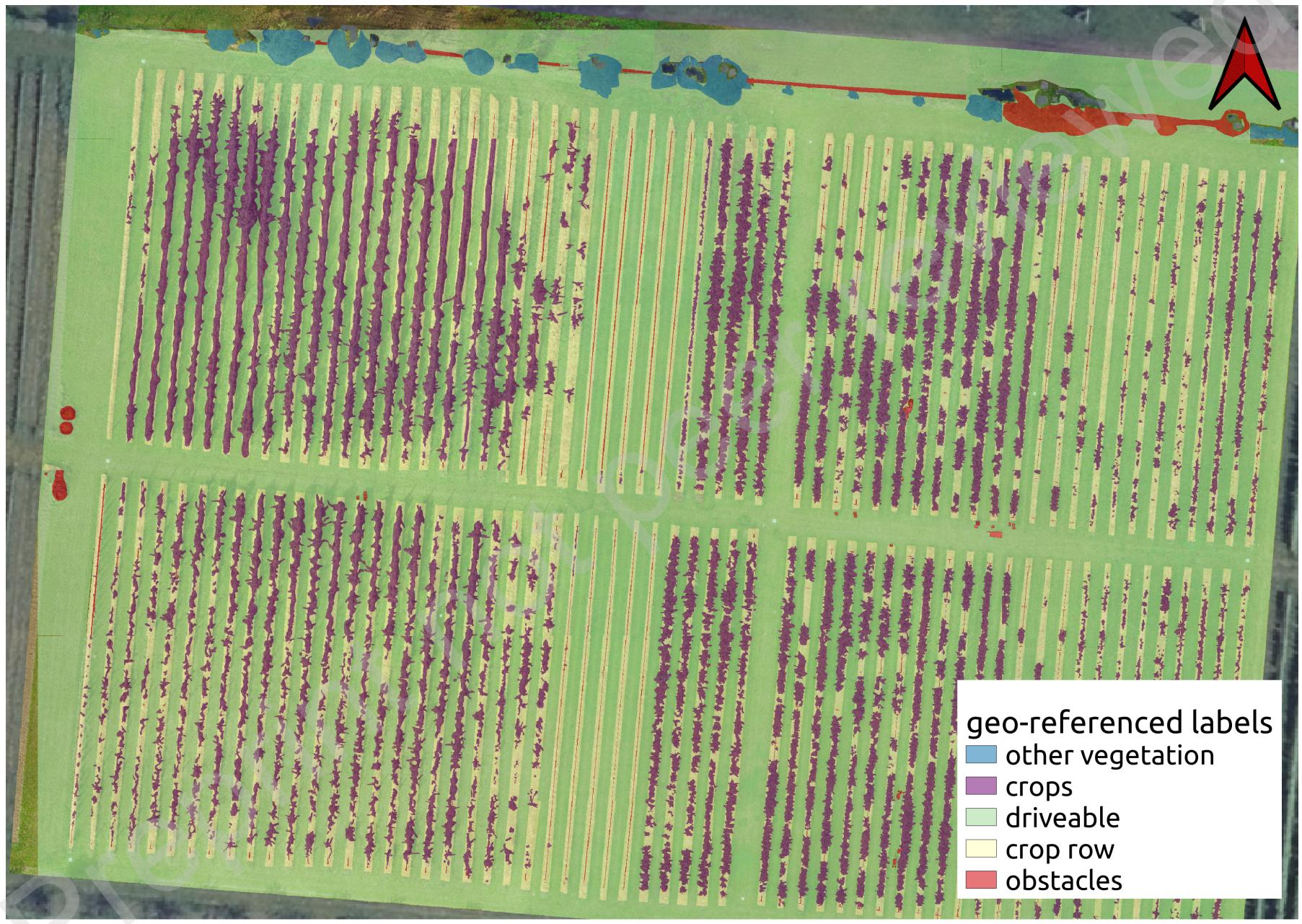
- 512 Macenski, S., Martin, F., White, R., Clavero, J.G., 2020a. The Marathon 2: A Navigation
513 System, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems
514 (IROS), pp. 2718–2725.
- 515 Macenski, S., Tsai, D., Feinberg, M., 2020b. Spatio-temporal voxel layer: A view on robot
516 perception for the dynamic world. International Journal of Advanced Robotic Systems 17,
517 172988142091053. <https://doi.org/10.1177/1729881420910530>.
- 518 Milella, A., Rilling, S., Rana, A., Galati, R., Petitti, A., Hoffmann, M., Stanly, J.L., Reina, G.,
519 2024. Robot-as-a-Service as a New Paradigm in Precision Farming. IEEE Access 12, 47942–
520 47949. <https://doi.org/10.1109/ACCESS.2024.3381511>.
- 521 Moravec, H., Elfes, A., 1985. High resolution maps from wide angle sonar, in: Proceedings.
522 1985 IEEE International Conference on Robotics and Automation, pp. 116–121.
- 523 Nørremark, M., Griepentrog, H.W., Nielsen, J., Søgaard, H.T., 2008. The development and
524 assessment of the accuracy of an autonomous GPS-based system for intra-row mechanical
525 weed control in row crops. Biosystems Engineering 101, 396–410.
526 <https://doi.org/10.1016/j.biosystemseng.2008.09.007>.
- 527 Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions
528 on Systems, Man, and Cybernetics 9, 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>.
- 529 Pan, Y., Hu, K., Cao, H., Kang, H., Wang, X., 2024. A novel perception and semantic mapping
530 method for robot autonomy in orchards. Computers and Electronics in Agriculture 219,
531 108769. <https://doi.org/10.1016/j.compag.2024.108769>.
- 532 Pflanz, M., Nordmeyer, H., Schirrmann, M., 2018. Weed Mapping with UAS Imagery and a Bag
533 of Visual Words Based Image Classifier. Remote Sensing 10, 1530.
534 <https://doi.org/10.3390/rs10101530>.
- 535 Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical
536 Image Segmentation, in: Medical Image Computing and Computer-Assisted Intervention
537 (MICCAI). Springer, pp. 234–241.
- 538 Rouault, E., Warmerdam, F., Schwehr, K., Kiselev, A., Butler, H., Łoskot, M., Szekeres, T.,
539 Tourigny, E., Landa, M., Miara, I., Elliston, B., Chaitanya, K., Plesea, L., Morissette, D.,
540 Jolma, A., Dawson, N., Baston, D., Stigter, C. de, Miura, H., 2024. GDAL. Zenodo.
541 <https://doi.org/10.5281/zenodo.5884351>. <https://gdal.org>.
- 542 Rusu, R.B., Cousins, S., 2011. 3D is here: Point Cloud Library (PCL), in: 2011 IEEE
543 International Conference on Robotics and Automation, pp. 1–4.
- 544 Santos, L.C., Aguiar, A.S., Santos, F.N., Valente, A., Petry, M., 2020a. Occupancy Grid and
545 Topological Maps Extraction from Satellite Images for Path Planning in Agricultural Robots.
546 Robotics 9, 77. <https://doi.org/10.3390/robotics9040077>.
- 547 Santos, L.C., Santos, F.N., Solteiro Pires, E.J., Valente, A., Costa, P., Magalhaes, S., 2020b. Path
548 Planning for ground robots in agriculture: a short review, in: 2020 IEEE International
549 Conference on Autonomous Robot Systems and Competitions (ICARSC), pp. 61–66.
- 550 Schütte, T., Koch, K., Dworak, V., Weltzien, C., 2023. LAND.TECHNIK AgEng 2023:
551 Improving Navigation of GNSS-based Autonomous Mobile Robots by Leveraging GNSS
552 and LiDAR Data for Structural Mapping and Path Planning., in: VDI Wissensforum GmbH
553 (Ed.), LAND.TECHNIK AgEng 2023: The Forum for Agricultural Engineering Innovations.
554 VDI Verlag, Düsseldorf, pp. 199–208.

- 555 Schwertfeger, S., Birk, A., 2016. Map evaluation using matched topology graphs. *Auton Robot*
556 40, 761–787. <https://doi.org/10.1007/s10514-015-9493-5>.
- 557 Shalal, N., Low, T., McCarthy, C., Hancock, N., 2015. Orchard mapping and mobile robot
558 localisation using on-board camera and laser scanner data fusion – Part B: Mapping and
559 localisation. *Computers and Electronics in Agriculture* 119, 267–278.
560 <https://doi.org/10.1016/j.compag.2015.09.026>.
- 561 Siek, J., Lee, L.-Q., Lumsdaine, A., 2001. The boost graph library: User guide and reference
562 manual. Addison-Wesley, Boston [etc.], XXIV, 321 s.
- 563 Sun, Q., Zhang, R., Chen, L., Zhang, L., Zhang, H., Zhao, C., 2022. Semantic segmentation and
564 path planning for orchards based on UAV images. *Computers and Electronics in Agriculture*
565 200, 107222. <https://doi.org/10.1016/j.compag.2022.107222>.
- 566 Tinkham, W.T., Swayze, N.C., 2021. Influence of Agisoft Metashape Parameters on UAS
567 Structure from Motion Individual Tree Detection from Canopy Height Models. *Forests* 12.
568 <https://doi.org/10.3390/f12020250>.
- 569 Yagfarov, R., Ivanou, M., Afanasyev, I., 2018. Map Comparison of Lidar-based 2D SLAM
570 Algorithms Using Precise Ground Truth, in: 2018 15th International Conference on Control,
571 Automation, Robotics and Vision (ICARCV). 2018 15th International Conference on
572 Control, Automation, Robotics and Vision (ICARCV), Singapore. 18.11.2018 - 21.11.2018.
573 IEEE, pp. 1979–1983.
- 574 Zhang, C., Valente, J., Kooistra, L., Guo, L., Wang, W., 2021. Orchard management with small
575 unmanned aerial vehicles: a survey of sensing and analysis approaches. *Precision Agric* 22,
576 2007–2052. <https://doi.org/10.1007/s11119-021-09813-y>.
- 577 Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X., Yan, G., 2016. An Easy-to-Use
578 Airborne LiDAR Data Filtering Method Based on Cloth Simulation. *Remote Sensing* 8, 501.
579 <https://doi.org/10.3390/rs8060501>.









geo-referenced labels

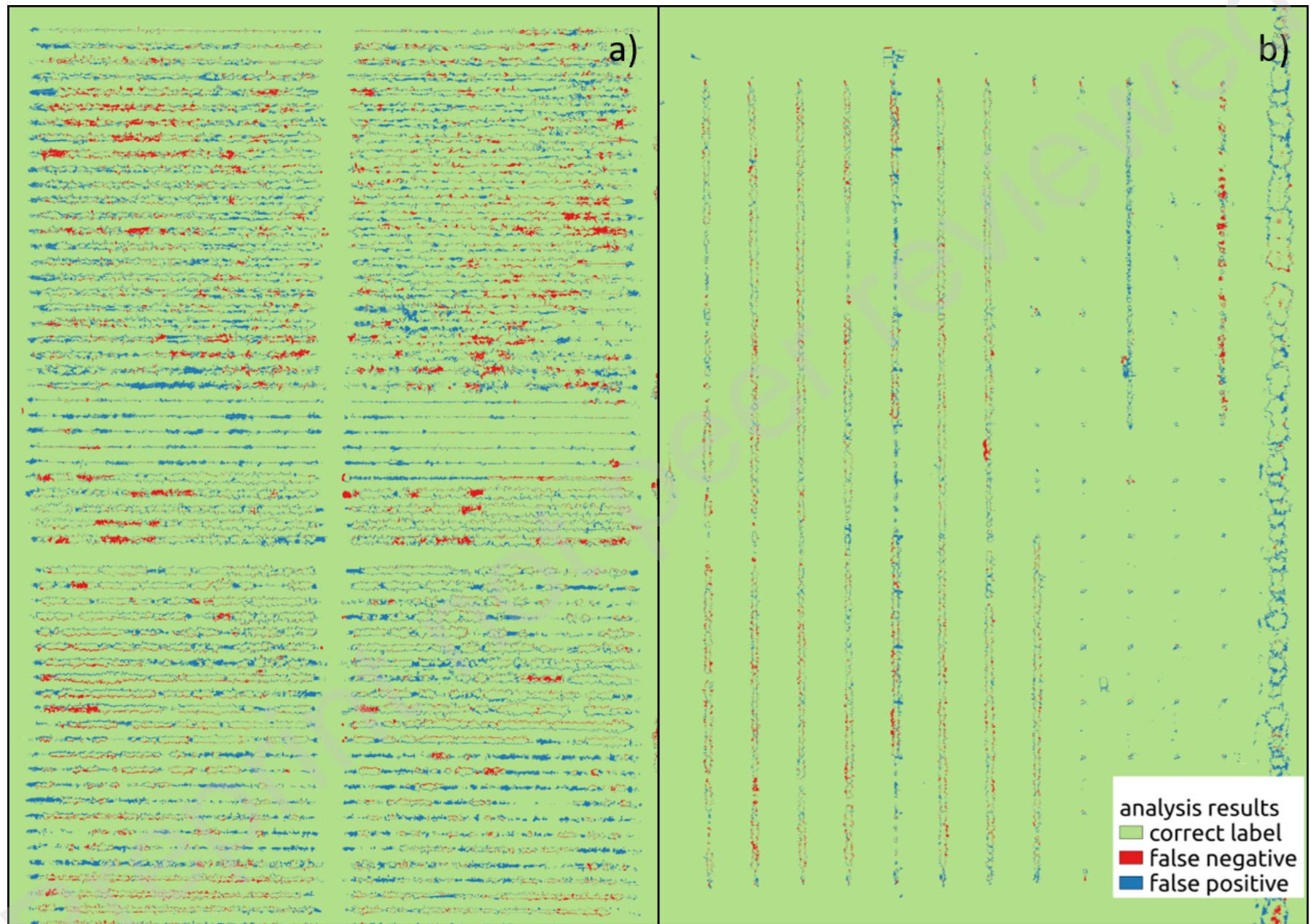
other vegetation

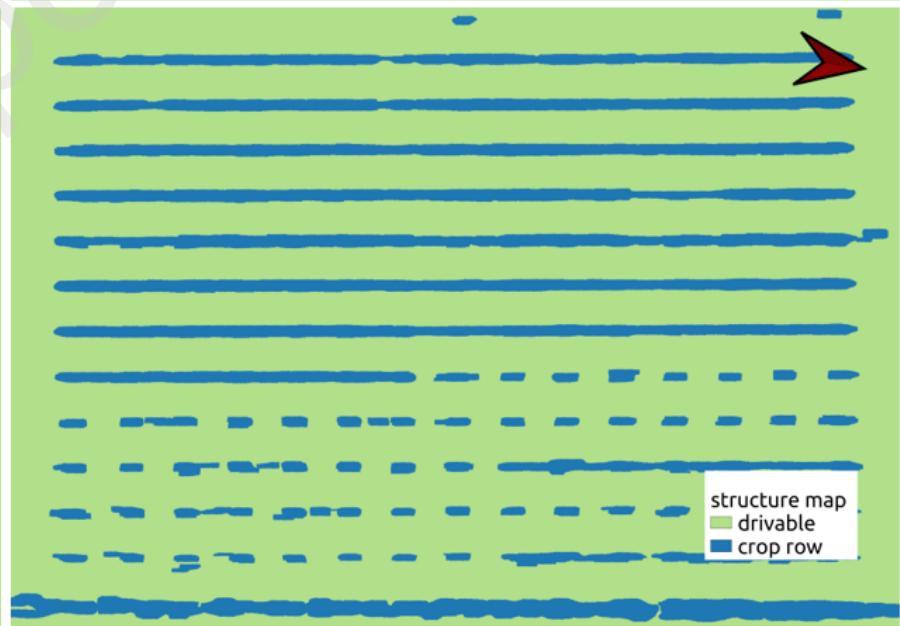
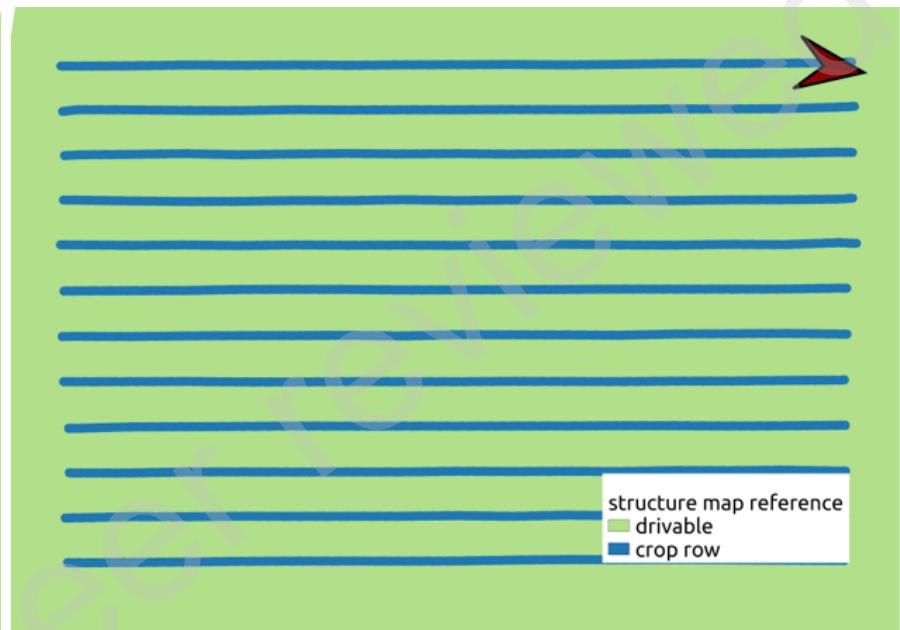
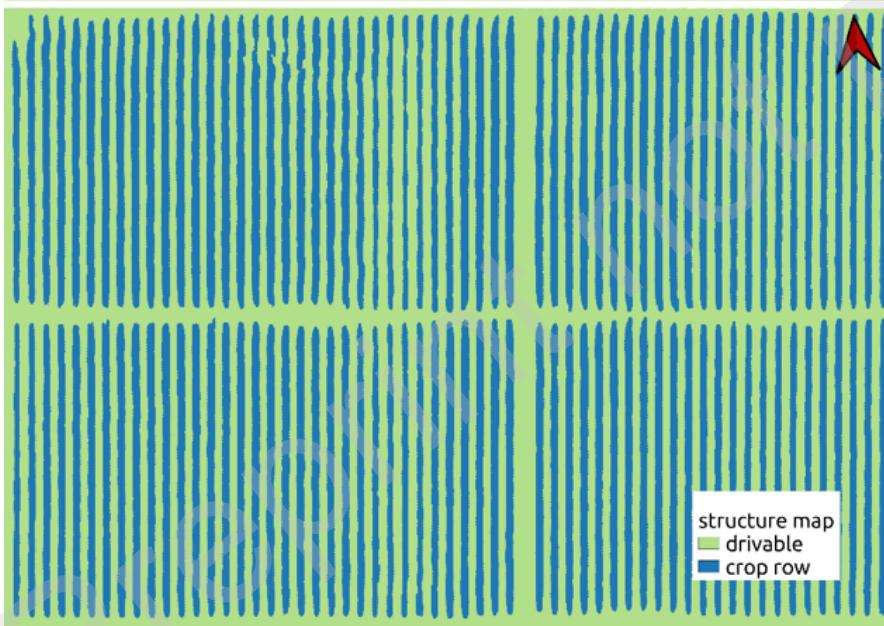
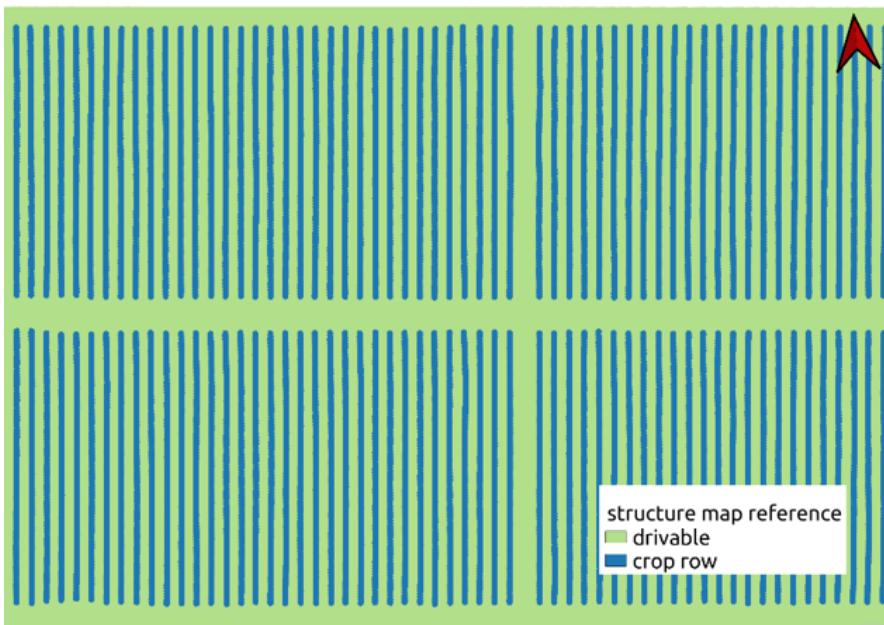
crops

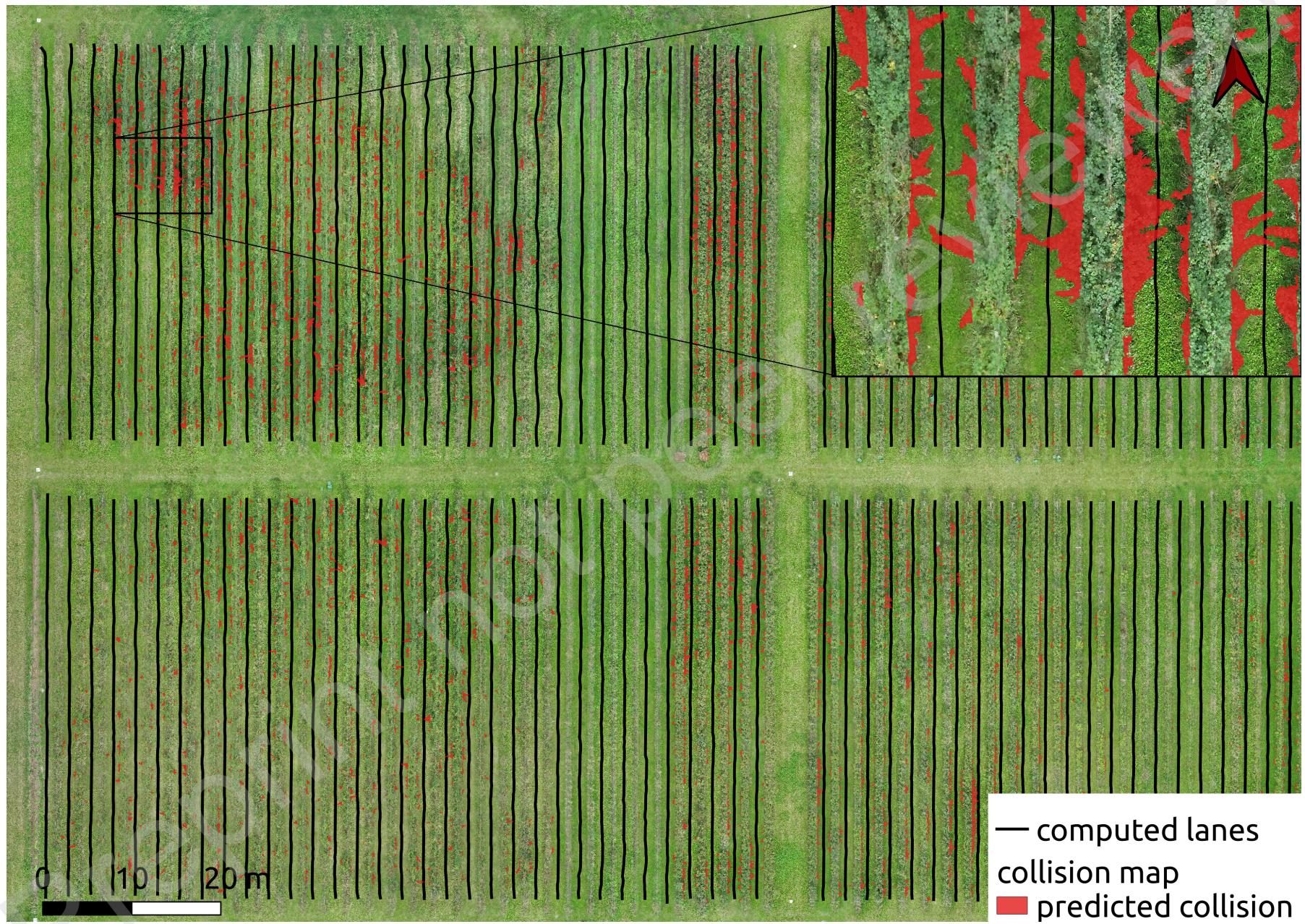
driveable

crop row

obstacles







Relative Area of Collision Over Working Width

