

# Quantitative Benchmarking in Agricultural Robotics

Riccardo Bertoglio, Giulio Fontana, Matteo Matteucci  
*Politecnico di Milano - DEIB*

Milan, Italy

{riccardo.bertoglio;giulio.fontana;matteo.matteucci}@polimi.it

Davide Facchinetti, Stefano Santoro  
*Università degli Studi di Milano*

Milan, Italy

{davide.facchinetti;stefano.santoro}@unimi.it

**Abstract**—The aim of ACRE (Agri-food Competition for Robot Evaluation) is to provide a set of benchmarks for agricultural robots and smart implements. While involving capabilities of general application, ACRE puts a special focus on weeding, identified as one of the tasks where it is easier for robotics to demonstrate its potential. ACRE, as the other three robot competitions that are being organised by European project METRICS, is built on the established idea of benchmarking through competitions. In this paper we present the framework of ACRE and examples of its benchmarks.

**Index Terms**—Agricultural Robotics, Robotics Benchmarking, Performance Evaluation

## I. INTRODUCTION

Companies producing agricultural machinery are well aware that artificial intelligence and robotics have the potential to revolutionize their field; yet they are wary of technology that is very far from their area of expertise. Even today, agricultural machines -with few exceptions- include very little in terms of ICT technology, and their users are rarely used to interacting with such systems in the field. Thus, introducing AI and robotics into agricultural machines needs much more than the addition of new systems and components: it requires a radical change in the way such machines are designed, built, and employed.

An additional difficulty is linked to the size of the companies producing machinery for agriculture, which is often (especially in some countries, such as Italy) comparatively small and sometimes bordering on the artisanal. Such companies, in fact, do not have the resources to embark in “clean slate” R&D projects. Finally, the potential benefits that robot-based precision agriculture can bring are significant, but so far theoretical. Farmers, i.e. the market for agricultural robots, are currently on the fence: waiting for such benefits to become clearly visible -and measurable in terms of yields and costs- before committing to radical changes to their modes of operation.

For the reasons outlined above, many companies are currently torn between a strong perceived need to “put their toes in the water”, i.e., to start the development of agricultural robots, and the fear of wasting resources due to lack of expertise and market. At the same time, there is considerable

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement No 871252.

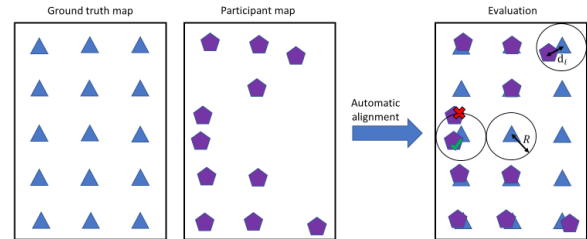


Fig. 1. Extract from the the ACRE Evaluation Plan: Crop mapping TBM.

effort going on in research on the development of algorithms and systems for agriculture, yet so far technology transfer to industry has been scarce, and mainly concentrated in spin-offs of research groups.

Quantitative benchmarking has the potential to change this scenario, enabling companies to quantitatively gauge the performance of research results before committing to implement them into their products, and providing farmers with tools to measure how much agricultural robots can enhance their production before considering their purchase.

## II. AGRI-FOOD COMPETITION FOR ROBOT EVALUATION

The aim of ACRE: Agri-food Competition for Robot Evaluation [1] is to provide a set of benchmarks for agricultural robots and smart implements. While involving capabilities of general application, ACRE puts a special focus on weeding, identified as one of the tasks where it is easier for robotics to demonstrate its potential. ACRE, as the other three robot competitions that are being organised by European project METRICS [2], is built on the established idea of *benchmarking through competitions* [3].

ACRE calls research and industry to demonstrate the capabilities and performance of their machines by executing *benchmarks* that take two forms [4]:

- **Functionality Benchmarks (FBMs)**, focused on specific capabilities of a robot and designed to make the benchmark as independent as possible from other features of the robot not directly involved in the functionality under examination (*Plant discrimination, Field navigation, Leaf area estimation, Weed destruction, Biomass estimation*)
- **Task Benchmarks (TBMs)**, evaluating the execution of complex tasks involving multiple functionalities, where

