Tropentag, September 10-12, 2025, hybrid conference

"Reconcile land system changes
with planetary health"

Human behaviour affects the performance of smartphone-based monitoring in coffee crop

Juan Camilo Rivera-Palacio[1], Christian Bunn[2], Masahiro Ryo[1]

[1] *Leibniz Centre for Agric. Landscape Res. (ZALF), Research Area 4 SSimulation and Data Science", Germany*

[2] *Alliance of Bioversity International and CIAT, Climate Action, Colombia*

Abstract

Citizen science presents a promising, scalable approach for agricultural data collection, particularly in low-resource settings. However, concerns about data quality and its impact on artificial intelligence (AI) model performance remain underexplored. This study investigates the factors influencing the accuracy of a deep learning model trained on citizen science–generated data for coffee production.

Specifically, we analysed errors in a (You Only Look Once) YOLO v8 object detection model trained to detect coffee cherries in images captured by coffee smallholder farmers in Colombia and Peru using their own smartphones.

YOLO v8 was trained on over 30.000 annotated coffee cherries and evaluated using 637 new photos. We applied a linear mixed model (LMM) and a decision tree to assess the effect of various factors—photographer identity, protocol adherence, mobile phone specifications, image characteristics, coffee variety, and geographic location—on prediction error e(j). Our findings revealed that photographer identity and adherence to image collection protocols were the most influential factors, with protocol compliance improving model performance significantly (R^2 increase from 0.48 to 0.73). In contrast, factors such as mobile phone type, screen size, or geographic differences did not substantially impact model accuracy.

These results underscore the importance of structured data collection guidelines, localised training, and ongoing feedback in enhancing the reliability of citizen science data for machine learning applications. Furthermore, we demonstrate that virtually any mobile device with a camera can contribute useful data, highlighting the scalability of this method for real-time, farmer-driven agricultural monitoring. Our work contributes to the broader understanding of how participatory approaches and AI can be integrated to support sustainable agricultural development and inclusive digital innovation.

Keywords: citizen science, yield, coffee, deep learning, YOLO

Contact Address: Juan Camilo Rivera-Palacio, Leibniz Centre for Agric. Landscape Res. (ZALF), Research Area 4 SSimulation and Data Science", Rhinstrasse 13, 10315 Müncheberg, Germany, e-mail: juancamilo.rivera@zalf.de