International Cartographic Association
Association Cartographique Internationale

# Leveraging machine learning and drone multispectral data for site-specific weed management in tomato agricultural areas

Martin Chari[a,*], Makgaboa Mashala[b], Cecilia Masemola[c]

[a] *Risk & Vulnerability Science Centre (RVSC), University of Fort Hare, Alice, Eastern Cape, South Africa; martinmchari@gmail.com*
[b] *Risk & Vulnerability Science Centre (RVSC), University of Limpopo, Turfloop, Limpopo, South Africa; makgaboa94@gmail.com*
[c] *Precision Agriculture Unit, Advanced Food, Agriculture and Health Cluster, Council for Scientific and Industrial Research (CSIR), Pretoria, South Africa; CMasemola1@csir.co.za*

* Corresponding author

**Keywords:** machine learning, precision agriculture, random forest, South Africa, drones/ UAV, weed management

**Abstract:**

Weeds can significantly reduce productivity in agriculture (Roslim et al., 2021). Mechanical weeding and pesticides for tomatoes under irrigation are some of the weed management methods which are time-consuming and environmentally hazardous (Gerhards et al., 2022). Remote sensing and machine learning techniques offer more precise weed detection and management, using high-resolution satellite data and drones to identify early weed infestations and employ precision agriculture approaches (Su et al., 2022). This study uses machine learning to analyze drone images and map weeds in the Ha-Mphaila tomato irrigation scheme in South Africa, showcasing the potential of drones and machine learning for site-specific weed management in agriculture.

The following figure 1 shows the study site in the Ha-Mphaila irrigation scheme located within 22°53'51.73"S, 30° 8'43.13"E in Makhado in the Vhembe district municipality of Limpopo Province of South Africa. In total, the farm covers a total area of 70.6 hectares and there are 62 households within it, each of which owns an average of 1 ha (Jiyane and Simalenga, 2019). The Ha-Mphaila tomato irrigation scheme is a system designed to provide controlled water supply to crops for optimal growth and productivity and the Mutshedzi river is a dependable source of water (Jiyane and Simalenga, 2019). The area of study site was 3 hectares. Field data was collected using a DJI Matrice 600 Pro drone with Micasense RedEdge-MX+DLS2 multispectral camera system flown at a height of 50m in Ha-Mphaila farm from 31 May 2022 to 04 June 2022 by the Agriculture Research Council (ARC). The camera system has a spectral range of 475-875 nm and 5 spectral bands, providing a spatial resolution of 10 centimeters per pixel and a spectral resolution of 5 bands with a bandwidth of 10 nm for each band. This enables the camera system to capture detailed spectral information of the vegetation and soil in the field. The drone was flown once during the data collection period, providing a single snapshot of the area at that time.

The semi-automated random forest (RF) classifier algorithm was used to classify bare soil, weeds, and tomatoes in Google Earth Engine (GEE) environment using multispectral data obtained from Unmanned Aerial Vehicles (UAVs). The algorithm was trained using digitized training areas of each class (bare soil, weeds, and tomatoes) to identify spectral signatures or patterns that distinguish between them. Data generated from the digitized training areas were split into two datasets 75% and 25%, which is "training" and "testing" respectively. Once the algorithm was trained, it was used to classify each pixel in the multispectral image into one of the three classes based on its spectral signature. This allowed the creation of a map of the field showing the location and extent of the different classes, including the weeds. The kind of weeds detected include blackjack (*Bidens bipinnata*) and tall-khakibush (*Tagetes minuta*). Model testing data was used in the construction of the confusion matrices for the validation of the performance of the RF model. The classification algorithm achieved training overall accuracy of 89% and a Kappa coefficient of 0.43 and overall validation accuracy of 63% and Kappa of 0.44. The validation accuracy of 63% means that the model is performing reasonably well on the validation set and able to make accurate predictions on new data, but there is still room for improvement. This is a common occurrence in machine learning and often requires further optimization of the model or collection of additional data. It is important to note that fair validation is not the same as high accuracy, as a model can achieve high accuracy on the training set but perform poorly on the validation set due to overfitting. Therefore, it is important to evaluate the performance of the model on both the training and validation sets to ensure it is generalizing well to new data.


https://doi.org/10.5194/ica-abs-6-34-2023 

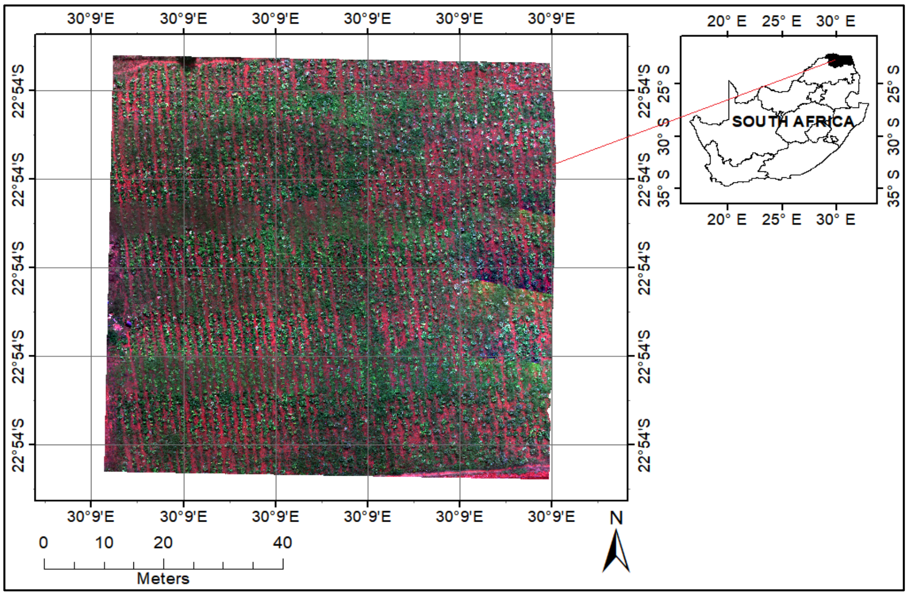


**Figure 1:** Map of study area in in the Ha-Mphaila irrigation scheme in the Vhembe district municipality of Limpopo Province in South Africa

The study was focused on finding more effective methods for weed management, and the proposed method could be scaled up to inform site-specific weed management strategies, reducing costs and promoting environmentally friendly crop farming. While this technology may not be immediately accessible to smallholder farmers, it could be used by larger agricultural operations or government agencies responsible for agricultural extension services. The proposed method has the potential to provide more effective and efficient weed management. By producing a reference map of weed patches and weed-free regions, the results of this study can influence a weed management plan for tomatoes under irrigation. The semi-automated random forest (RF) classifier technique utilized in this study can be used to map weeds and improve crop management practices in different crop fields. The study shows the potential of machine learning and unmanned aerial vehicles (UAVs) for weed management, which could be especially useful for small-holder agricultural production in data-scarce locations. The study's disadvantage is that machine learning algorithms require accurate and representative training data, which may be influenced by ground truth data quality, such as mislabelled or inadequate samples. The study is limited to a case study location in South Africa and may not apply to other regions with various crop varieties, soil types, and weed species. The study does not address the economic feasibility and social acceptability of employing drones and machine learning algorithms for weed management in smallholder agricultural communities, which may require future research.

**Acknowledgements**

This work was a group project during 2022 specialised training in machine learning for earth observation, with application in agriculture through the FAIR Forward-Artificial Intelligence for all Initiative of Deutsche Gesellschaft fur Internationale Zusammenarbeit (GIZ). The authors would like to express their gratitude to GIZ for providing financial support for the program. We also thank University of the Witwatersrand (WITS) GAES, Council for Scientific and Industrial Research (CSIR) Precision Agriculture, Agriculture Research Council (ARC) for their training and financial assistance during the course. The SAF-ADAPT project is also acknowledged for providing conference funding.

**References**

Gerhards, R., Andujar Sanchez, D., Hamouz, P., Peteinatos, G. G., Christensen, S., & Fernandez-Quintanilla, C. (2022). Advances in site-specific weed management in agriculture—A review. *Weed Research*, *62*(2), 123-133

Jiyane, J. & Simalenga, T.E. (2019). Factors influencing under-utilisation of smallholder irrigation schemes and opportunities to improve the schemes' productivity in Limpopo Province, South Africa. WRC Report No. TT 787/19: ISBN 978-0-6392-0100-9

Roslim, M. H. M., Juraimi, A. S., Che'Ya, N. N., Sulaiman, N., Manaf, M. N. H. A., Ramli, Z., & Motmainna, M. (2021). Using remote sensing and an unmanned aerial system for weed management in agricultural crops: A review. *Agronomy, 11*(9), 1809

Su, J., Yi, D., Coombes, M., Liu, C., Zhai, X., McDonald-Maier, K., & Chen, W. H. (2022). Spectral analysis and mapping of blackgrass weed by leveraging machine learning and UAV multispectral imagery. *Computers and Electronics in Agriculture*, *192*, 10662


https://doi.org/10.5194/ica-abs-6-34-2023 | 