Pesquisa Agropecuária Brasileira

ISSN 1678-3921

Journal homepage: www.embrapa.br/pab

For manuscript submission and journal contents, access: www.scielo.br/pab

Digital Agriculture/ Original Article

Machine learning algorithms to predict the crops most susceptible to weed occurrence in integrated crop-livestock systems

Abstract – The objective of this work was to investigate the use of machine learning algorithms to predict the crops most susceptible to weed occurrence in integrated crop-livestock systems, based on environmental factors of climate, soil, and cropping systems, to establish correlations between these elements and the occurrence of weeds. Three datasets were used for this purpose: the first provided quantitative information on the invasive species, the second contained data about the soil, and the last had records of the region's climate. The algorithms used were Support Vector Machine, Decision Tree, Random Forest, and K-Nearest Neighbors. The application of machine learning algorithms to predict the susceptibility of crops to weed emergence is technically feasible and effective. The Decision Tree and Random Forest algorithms demonstrated the best performance, with both models achieving 99% accuracy. Robust relationships were established between environmental factors (climate, soil, and planting) and the appearance of invasive species in certain crops. The algorithms reproduced the patterns of weed emergence observed under field conditions.

Index terms: artificial intelligence, crop rotation systems, data analysis, weed management.

Algoritmos de aprendizado de máquina para prever as culturas mais suscetíveis à ocorrência de plantas daninhas em sistemas de integração de lavoura-pecuária

Resumo –O objetivo deste trabalho foi investigar o uso de algoritmos de aprendizado de máquina para prever as culturas mais suscetíveis à ocorrência de plantas daninhas em sistemas de integração de lavoura-pecuária, com base em fatores ambientais de clima, solo e sistemas de cultivo para estabelecer correlações entre esses elementos e a incidência de plantas invasoras. Três conjuntos de dados foram utilizados para esse propósito: o primeiro forneceu informações quantitativas sobre as espécies invasoras, o segundo continha dados sobre o solo e o último possuía registros do clima da região. Os algoritmos utilizados foram Support Vector Machine, Decision Tree, Random Forest e K-Nearest Neighbors. A aplicação de algoritmos de aprendizado de máquina para prever a suscetibilidade das culturas à emergência de plantas daninhas é tecnicamente viável e eficaz. Os algoritmos Decision Tree e Random Forest demonstraram o melhor desempenho, com ambos os modelos atingindo 99% de acurácia. Relações robustas foram estabelecidas entre os fatores ambientais (clima, solo e plantio) e o surgimento de espécies invasoras em determinadas culturas. Os algoritmos reproduziram com sucesso os

Ana Letícia Gomes Luz (✉)
Fundação Getúlio Vargas, Rio de Janeiro, RJ, Brazil. E-mail: C3022181@fgv.edu.br

Anita Maria da Rocha Fernandes
Universidade do Vale do Itajaí, Itajaí, SC, Brazil. E-mail: anita.fernandes@univali.br

Fábio Volkman Coelho
Universidade do Vale do Itajaí, Itajaí, SC, Brazil. E-mail: fabiovcoelho05@gmail.com

Maurílio Fernandes de Oliveira
Embrapa Milho e Sorgo, Sete Lagoas, MG, Brazil. E-mail: maurilio.oliveira@embrapa.br

Ramon Costa Alvarenga
Embrapa Milho e Sorgo, Sete Lagoas, MG, Brazil. E-mail: ramon.alvarenga@embrapa.br

✉ Corresponding author

Received
July 5, 2024

Accepted
June 7, 2025

How to cite
LUZ, A.L.G.; FERNANDES, A.M. da R.; COELHO, F.V.; OLIVEIRA, M.F. de; ALVARENGA, R.C. Machine learning algorithms to predict the crops most susceptible to weed occurrence in integrated crop-livestock systems. **Pesquisa Agropecuária Brasileira**, v.61, e03833, 2026. DOI: https://doi.org/10.1590/S1678-3921. pab2026.v61.03833.

This is an open-access article distributed under the Creative Commons Attribution 4.0 International License

padrões de emergência de plantas daninhas observados em condições de campo.

Termos para indexação: inteligência artificial, sistemas de rotação de culturas, análise de dados, controle de plantas daninhas.

Introduction

Weed populations in integrated crop-livestock systems (ICLS) are typically lower than those found in continuous tillage systems (Concenço et al., 2011). This reduction is attributed to forage plants, which act as an important soil cover in intercropped cultures (Concenço et al., 2015; Duarte et al., 2018), preventing the emergence of invasive plants (Schuster et al., 2019). Additionally, rotation systems contribute to reducing the weed seed bank in the soil (Ikeda et al., 2007). Even with these natural benefits, the need for precise and effective monitoring means that the adoption of technologies in weed management is growing (Gomes et al., 2024).

Machine learning (ML) approach is the basis for diverse technologies used to deal with various aspects of weed control (Jha et al., 2019). These ML applications encompass systems that: identify the optimal timing to perform weed control (Monteiro et al., 2021); estimate weed density and distribution (Shorewala et al., 2021); recognize and classify different weed species (Sabzi & Abbaspour-Gilandeh, 2018; Etienne et al., 2021; Costello et al., 2022); and use smart sprayers to confirm local application of herbicides (Yu et al., 2019; Hussain et al., 2021; Raja et al., 2023).

However, the majority of these technologies target post-emergence weed management (Gomes et al., 2024). Consequently, the application of ML algorithms to develop solutions to predict weed occurrence has not yet been widely studied. Furthermore, studies on the use of artificial intelligence applications in integrated crop-livestock systems is unknown (Gomes et al., 2024).

Investigating the use of these algorithms to predict occurrence of invasive species may be appropriate. Firstly, to understand the environmental factors and the cropping systems that favor weed emergence. In addition, the algorithm's predictive results can be used to modify or adopt new techniques in the field, thereby focusing on reducing rates of weed appearance and growth.

The objective of this work was to investigate the use of ML algorithms to predict the crops most susceptible to weed occurrence in integrated crop-livestock systems, based on environmental factors of climate, soil, and cropping system, to establish correlations between these elements and the occurrence of weeds.

Materials and Methods

The study was conducted in the Cerrado biome within the municipality of Sete Lagoas, in the state of Minas Gerais, Brazil (19°29'4.37"S, 44°10'25.66"W, at 755 of altitude). The local climate, classified as Aw (humid tropical) by Köppen, is characterized by an average temperature above 18°C in the coldest month, a dry winter, and a rainy summer season. The average annual rainfall totals 1,350 mm, concentrated from October to March, with a notable occurrence of dry spells often observed in January and February. The soil was classified as Latossolo Vermelho distrófico, according to Brazilian Soil Classification System (Santos et al., 2018), i.e., an Oxisol. The experimental design and the cropping systems information are described by Alvarenga et al. (2007).

The study utilized three datasets: one containing quantitative information on weed species, the second detailing soil properties (pH, organic matter, and nutrients), and the third focused on the local climate. All data collected on weeds, climate, and soil were from the integrated crop-livestock systems experiments belonging to Embrapa (Embrapa Milho e Sorgo, 2018). The weed dataset provided by Embrapa Milho e Sorgo encompasses variables including date, weed common name, leaf morphology (narrow leaf or broad leaf), fresh biomass, dry biomass, sampling period, crop, field number, number of frames, area of each frame, and total area sampled. The precise locations within each field where the weeds were collected were demarcated. The data collection covered two specific time periods: 2006 and from 2015 to 2023.

The rotation system involved four crop types, utilized in the following proportional areas: maize (19.3%), sorghum (4.5%), soy (17.9%), and forage plants for pasture (9.9%). Weed sampling was carried out across four distinct periods corresponding to the crop cycle. The measurements occurred: at harvest, before or immediately after crop harvest; vegetative growth stage, right after planting and prior to the herbicides application; off season, between crop cycles, before seedling, during cattle grazing period; pre-desiccation, a specific stage during off season period, immediately before the desiccation herbicide application.

Some dataset adjustments were required prior to the analysis. The nomenclature of the weed species was standardized and missing dry biomass data were imputed using the median (Bruce et al., 2020). It was also possible to estimate missing values for fresh biomass using the rule of three, based on the principle that dry biomass constitutes approximately 40% of the fresh plant biomass. After these procedures, the total dataset contained 1,103 records.

The soil dataset, provided by Embrapa Milho e Sorgo, encompasses a list of variables. It included identifying contextual information: year, field number, and soil depth, ranging from 0 to 60 cm. The variables also described: soil pH measured in water, aluminum saturation (%), organic matter (dag kg^{-1}), total exchangeable bases ($cmol_c$ dm^{-3}), cation-exchange capacity ($cmol_c$ dm^{-3}), base-cation saturation ratio (%), H+Al ($cmol_c$ dm^{-3}), Al ($cmol_c$ dm^{-3}), Ca ($cmol_c$ dm^{-3}), Mg ($cmol_c$ dm^{-3}), K (mg dm^{-3}), P (mg dm^{-3}), B (mg dm^{-3}), Zn (mg dm^{-3}), Fe (mg dm^{-3}), Cu (mg dm^{-3}), and Mn (mg dm^{-3}). Finally, the dataset included the specific ratios: Ca/Mg ($cmol_c$ dm^{-3}); Ca/K ($cmol_c$ mg^{-1}); Mg/K ($cmol_c$ mg^{-1}); and Ca+Mg/K ($cmol_c$ mg^{-1}).

The variables organic matter, Zn, Cu, Mn, Fe, and B presented missing data. The missing values were replaced with estimates using the median method by Bruce et al. (2020). Overall, this final dataset contained 215 records collected across multiple years: 2005, 2006, 2012, 2014, 2015 to 2020, and 2022.

The third dataset, on climate, was compiled by merging two distinct data sources. The first, from the Brazilian National Institute of Meteorology (INMET), provided weather records for the region from 2017 onwards. The second, from the Sete Lagoas Automatic Meteorological Station (EMA), supplied weather data specific to Sete Lagoas, from 2000 to 2016. This combination was necessary to ensure weather records spanning all years of interest to the present study.

To combine the two sources into the final climate dataset, only the common variables were selected. The variables were: date, air pressure (hPa), average temperature (°C), maximum temperature (°C), minimum temperature (°C), humidity (%), wind speed (m s^{-1}), wind direction (compass degrees), and rainfall (mm). The missing values were replaced with estimates using the median method (Bruce et al., 2020). After datasets merging and processing, the final climate dataset comprised 8,066 climate records.

The soil and climate datasets were submitted to a statistical dispersion analysis to ensure data cleaning (Figure 1). This analysis was performed in order to detect outliers and verify the accuracy of the data (Das & Cakmak, 2018; Bruce et al., 2020).

The weed and climate datasets were merged using the variable date. This merged dataset was then combined with the soil dataset using year as the key variable. The resulting final dataset contained 1,541 records. To prepare the data for the algorithms, the final dataset was normalized and discretized. Some variables, such as year, were only used during the merging process, and were excluded from the final models. Other variables, such as fresh biomass and dry biomass, were eliminated later during the model development since they lowered the values of the performance metrics. The variables retained and used in the final models are detailed in Table 1.

Classification Supervised Learning (CSL) algorithms, a class of machine learning (ML), were employed to predict which crops were most likely to host invasive species, based on the provided climate, soil, and planting conditions. The specific CSL algorithms selected for the present study were Support Vector Machine, Decision Tree, Random Forest, and K-Nearest Neighbors. These algorithms were chosen because they are the most frequently used in weed control problems (Gomes et al., 2024). For all algorithms, the data were split using an 80% proportion for training and 20% for testing, using the same amount of data for each model. A data-shuffling criterion was applied to enhance model performance (Das & Cakmak, 2018; Bruce et al., 2020), expecting that data shuffling may alter the final results.

The performance of the algorithms was evaluated using the following classification metrics: accuracy, precision, recall, and F1 score. These metrics were estimated using the following equations: Accuracy = (TP+TN)/(TP+TN+FP+FN), Precision = TP/(TP+FP), Recall = TP/(TP+FN), and F1 Score = 2×[(Precision×Recall)/ Precision+Recall)], where TP is the number of true positives, TN is the number of true negatives, FP is the number

DOI: 10.1590/S1678-3921.pab2026.v61.03833

of false positives, and FN is the number of false negatives (Das & Cakmak, 2018; Bruce et al., 2020).

The Decision Tree and Random Forest algorithms presented the best performance when using the entropy criterion, which serves as a measure of data disorganization (Das & Cakmak, 2018; Bruce et al., 2020). Support Vector Machine algorithm achieved its optimal performance using linear kernel. Finally, the K-Nearest Neighbors algorithm performed best when the number of neighbors (K) was equal to two.

Results and Discussion

The algorithms demonstrated good overall performance, predicting the crops most likely to host weed species. The metrics results are displayed in Table 2. However, Maize/Pasture and Pasture/Pasture were not represented in the final predictions.

The most critical failure in the prediction occurred when the algorithm neglected to identify an invasive species in a crop that it appeared. As the weed's presence was not expected, no preventive measures

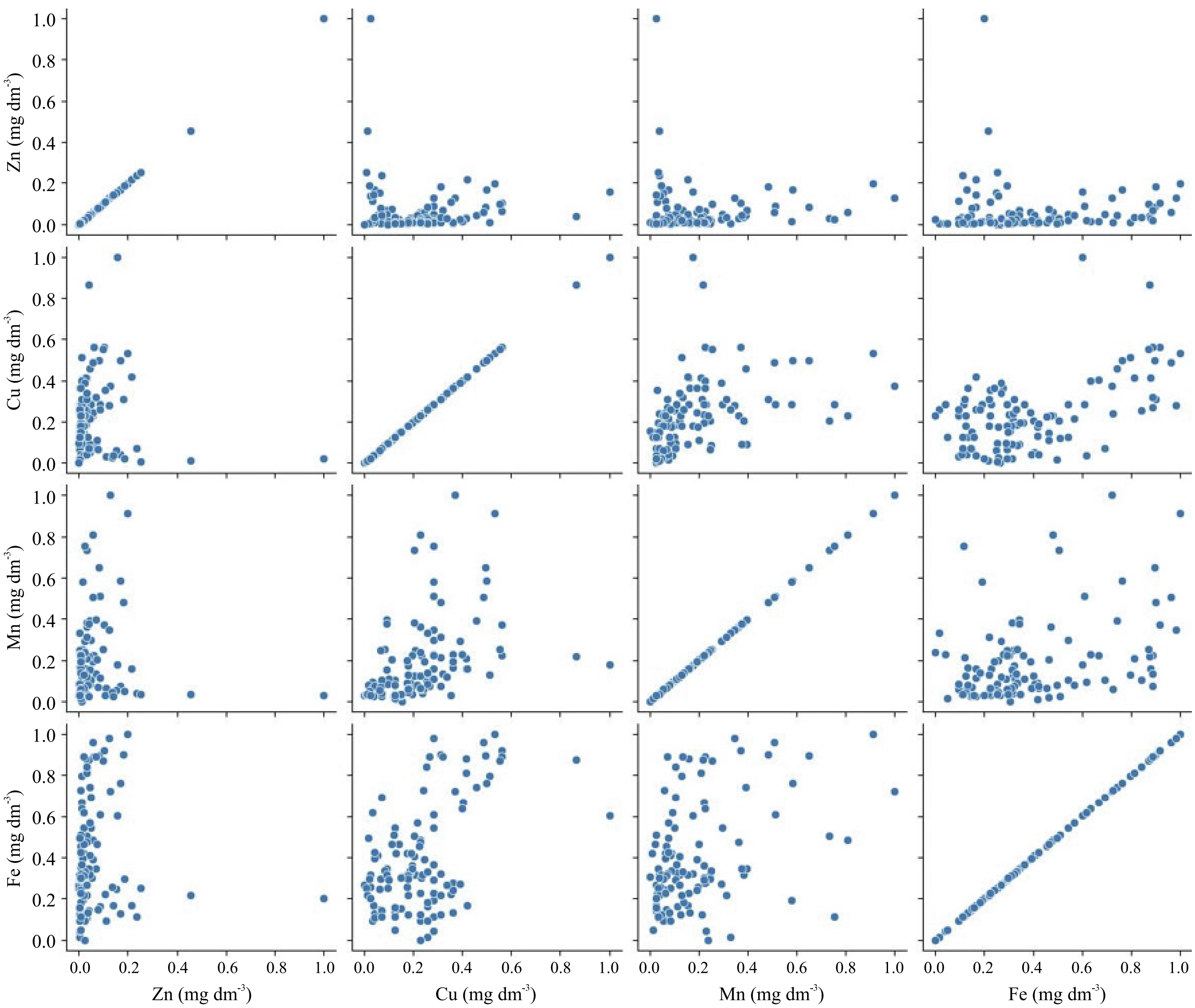

Figure 1. Dispersion statistical analysis of soil data (Zn, Cu, Mn and Fe) in the integrated crop-livestock system, in the municipality of Sete Lagoas, in the state of Minas Gerais, Brazil.

DOI: 10.1590/S1678-3921.pab2026.v61.03833

were taken, resulting in crop losses. High values of the precision and recall metrics indicated that it is not occurring (Das & Cakmak, 2018; Bruce et al., 2020). While K-Nearest Neighbors algorithm presented a low recall for Sorghum/Pasture, all algorithms still presented good performance.

Decision Tree and Random Forest algorithms provided feature importance scores (percentages) for each input variable used in the prediction process (Table 1). The percentages reflected the best

Table 1. Variable importance scores for machine learning algorithms within the integrated crop-livestock system, in the municipality of Sete Lagoas, in the state of Minas Gerais, Brazil.

Variable	Decision Tree (%)	Random Forest (%)
Sampling period	22.42	10.20
Air pressure (hPa)	18.44	8.78
Fe (mg dm^{-3})	14.89	7.50
Organic matter (dag kg^{-1})	7.25	1.10
Wind speed (m s^{-1})	5.90	2.80
Al ($cmol_c$ dm^{-3})	5.31	5.21
Cu (mg dm^{-3})	4.88	7.07
Minimum temperature (°C)	3.22	3.35
Weed name	3.20	2.04
Leaf type	2.20	0.60
Ca+Mg/K ($cmol_c$ mg^{-1})	1.81	1.88
Aluminum saturation (%)	1.77	3.34
B (mg dm^{-3})	1.71	1.95
Wind direction (compass degrees)	1.58	2.29
Mn (mg dm^{-3})	1.41	5.20
Ca/Mg ($cmol_c$ dm^{-3})	1.22	1.96
Base-cation saturation ratio (%)	0.80	1.01
H+Al ($cmol_c$ dm^{-3})	0.73	1.77
Mg/K ($cmol_c$ dm^{-3})	0.32	1.24
P (mg dm^{-3})	0.32	0.35
Ca/K ($cmol_c$ mg^{-1})	0.17	1.26
Maximum temperature (°C)	0.16	6.73
Number of frames	0.16	0.24
Humidity (%)	0.00	6.00
pH H_20	0.00	3.64
Average temperature (°C)	0.00	3.53
Rainfall (mm)	0.00	2.19
Zn (mg dm^{-3})	0.00	1.53
Mg ($cmol_c$ dm^{-3})	0.00	1.23
Ca ($cmol_c$ dm^{-3})	0.00	1.03
K (mg dm^{-3})	0.00	0.98
Total exchangeable bases ($cmol_c$ dm^{-3})	0.00	0.87
Cation-exchange capacity ($cmol_c$ dm^{-3})	0.00	0.73
Soil depth (cm)	0.00	0.16
Total area sampled (m^2)	0.00	0.05
Area of each frame (m^2)	0.00	0.00

Table 2. Performance metrics (precision, recall, accuracy and F1 score) of machine learning algorithms for weed emergence prediction in integrated crop-livestock system, in the municipality of Sete Lagoas, in the state of Minas Gerais, Brazil.

Algor.	Crop(1)	Precision (%)	Recall (%)	Accuracy (%)	F1 Score (%)
Support vector machine	Maize	100	100	98	100
	Maize/Sorghum	97	97	98	97
	Pasture	100	94	98	97
	Pasture/Soy	100	100	98	100
	Soy	98	100	98	99
	Soy/Maize	97	100	98	99
	Sorghum	96	100	98	98
	Sorghum/Pasture	100	93	98	84
Decision tree	Maize	100	100	99	100
	Maize/Sorghum	97	100	99	99
	Pasture	100	94	99	97
	Pasture/Soy	100	100	99	100
	Soy	100	100	99	100
	Soy/Maize	92	99	99	99
	Sorghum	100	100	99	96
	Sorghum/Pasture	100	100	99	100
Random forest	Maize	98	100	99	99
	Maize/Sorghum	100	100	99	100
	Pasture	100	93	99	96
	Pasture/Soy	83	100	99	91
	Soy	100	100	99	100
	Soy/Maize	99	100	99	99
	Sorghum	96	100	99	98
	Sorghum/Pasture	100	80	99	89
K-nearest neighbors	Maize	95	100	95	97
	Maize/Sorghum	94	100	95	97
	Pasture	100	88	95	94
	Pasture/Soy	83	83	95	83
	Soy	96	100	95	98
	Soy/Maize	94	100	95	97
	Sorghum	100	96	95	98
	Sorghum/Pasture	100	27	95	43

(1) The slash notation indicates the sequence of the crop rotation during the off-season: the crop harvested precedes the crop to be planted.

configuration for each algorithm. For both algorithms, the most important variables were sampling period, air pressure, and Fe.

These variables and their importance values can vary according to the shuffling. During optimization, it was observed that the most important climate variable frequently altered due to shuffling. In some results, the algorithm pointed to humidity as the most important, while in others, it indicated average temperature. However, the most important weed variable was consistently sampling period, and Fe was always the most important soil variable throughout the prediction process.

The sampling period variable proved to be extremely important for crop prediction. Changes in a crop's growth stage alter microclimate conditions, soil conditions, and, consequently, there are variations of the sunlight incidence on the crop canopy, which influence the appearance of weeds (Velin et al., 2012).

The variable Fe was the most important soil element influencing algorithm performance, though Cu and Mn were also significant, particularly for the Random Forest model. This aligns with the understanding that weeds make efficient use of micronutrients and require small amounts for growth (Thapa et al., 2021). Notably, micronutrients show a negative correlation with the variable organic matter (Dhaliwal et al., 2019). Therefore, low organic matter levels lead to higher micronutrient values, which was observed in Soy/Maize and Maize crops. These two crops, which lacked forage plants, exhibited the highest amount of weeds: 2,099 and 1,865, respectively, since forage benefits weed reduction. Soy/Maize also recorded the highest average Fe levels. This suggests that micronutrients play a more relevant effect in the absence of forage plants: low organic matter levels increase micronutrients contents in the soil, which in turn correlates with a higher presence of weeds.

Regarding the climate data, air pressure showed a negative correlation with rainfall (Figure 2), that is, higher air pressure was associated with less rainfall, and vice-versa (Ynoue et al., 2017). Since

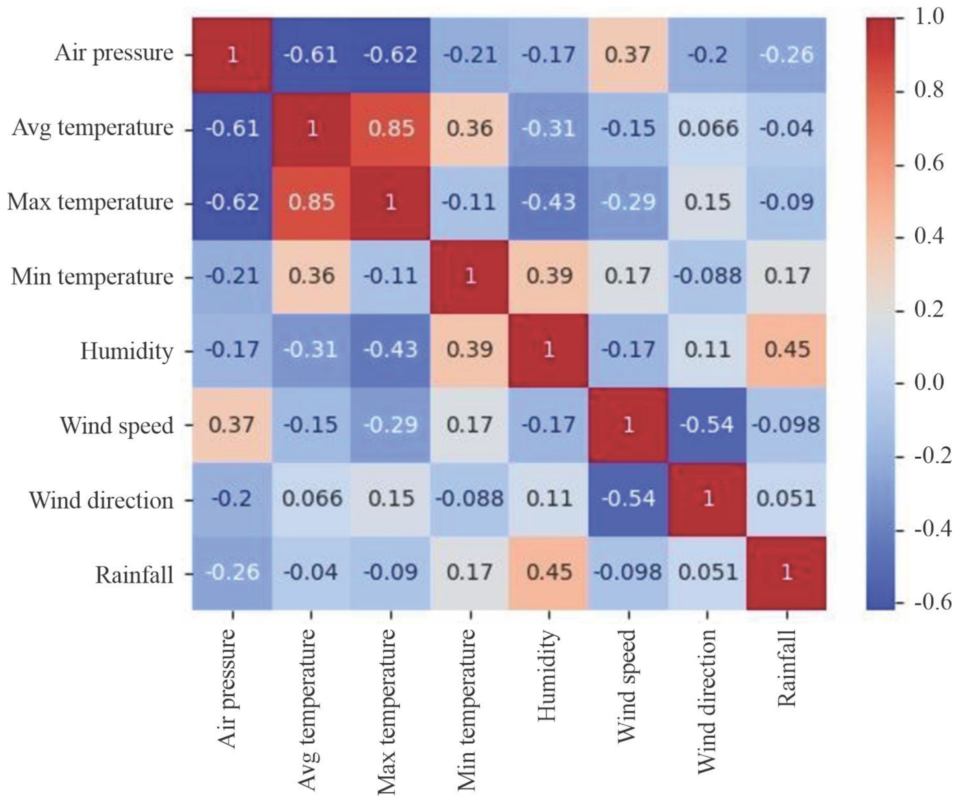

Figure 2. Correlation matrix of climate variables of the integrated crop-livestock system, in the municipality of Sete Lagoas, in the state of Minas Gerais, Brazil.

DOI: 10.1590/S1678-3921.pab2026.v61.03833

field observations and literature confirm that rainfall directly influences weed emergence (Werth et al., 2017), it seems that the algorithms used air pressure as a proxy for rainfall. Furthermore, the variable humidity proved to be important to the algorithms, since most of the weed species found in the studied region are not adapted to dry periods.

The variable average temperature demonstrated its importance, which explains why maximum temperature was given considerable importance by the Random Forest algorithm, since higher temperatures favor seed production and vegetative growth. The variable wind speed was found to be a key variable for the Decision Tree algorithm. The correlation analyses showed a positive correlation between air pressure and wind speed (Figure 2). Higher air pressure resulted in higher wind speed (Ynoue et al., 2017). It allows to establishing a relationship between rainfall and wind speed: significant variations in air pressure increase wind, which leads to rain (Ynoue et al., 2017).

All these identified relationships between the environmental factors and weed appearance are consolidated results. Since the algorithms predicted these relationships, they are confirmed as promising techniques for practical application for agricultural use.

Conclusions

1. The application of machine learning algorithms to predict the susceptibility of crops to weed emergence is technically feasible and effective.
2. The Decision Tree and Random Forest algorithms demonstrated the best performance, with both models achieving 99% accuracy.
3. Robust relationships were established between environmental factors (climate, soil, and planting) and the appearance of invasive species in certain crops.
4. The algorithms reproduced the patterns of weed emergence observed under field conditions.

References

ALVARENGA, R.C.; GONTIJO NETO, M.M.; RAMALHO, J.H.; GARCIA, J.C.; VIANA, M.C.M.; CASTRO, A.A.D.N. **Sistema de integração lavoura-pecuária**: o modelo implantado na Embrapa milho e sorgo. Sete Lagoas: Embrapa, 2007. (Embrapa Milho e Sorgo. Comunicado técnico, 93).

BRUCE, P.; BRUCE, A.; GEDECK, P. **Practical Statistics for Data Scientists:** 50+ essential concepts using R and Python. 2nd. ed. Sebastopol: O'Reilly Media, 2020. 360p.

CONCENÇO, G.; SALTON, J.C.; CECCON, G. **Dinâmica de plantas infestantes em sistemas integrados de cultivo**. Dourados: Embrapa Agropecuária Oeste, 2011a. (Embrapa Agropecuária Oeste. Documentos, 114).

CONCENÇO, G.; SALTON, J.C.; MARQUES, R.F.; PALHARINI, W.G.; ALVES, M.E.S.; SANTOS, S.A.; GALON, L. Weed suppression in sustainable integrated agricultural systems. **Pakistan Journal of Weed Science Research**, v.21, p.1-14, 2015. DOI: https://doi.org/10.28941/pjwsr.v21i1.569.

COSTELLO, B.; OSUNKOYA, O.O.; SANDINO, J.; MARINIC, W.; TROTTER, P.; SHI, B.; GONZALEZ, F.; DHILEEPAN, K. Detection of parthenium weed (*Parthenium hysterophorus* L.) and its growth stages using artificial intelligence. **Agriculture**, v.12, p.1-23, 2022. DOI: https://dx.doi.org/10.3390/agriculture12111838.

DAS, S.; CAKMAK, U.M. **Hands-On Automated Machine Learning**: a beginner's guide to building automated machine learning systems using AutoML and Python. 1.st ed. Birmingham: Packt Pub Ltd, 2018. 284p.

DHALIWAL, S.S.; NARESH, R.K.; MANDAL, A.; SINGH, R.; DHALIWAL, M.K. Dynamics and transformations of micronutrients in agricultural soils as influenced by organic matter build-up: a review. **Environmental and Sustainability Indicators**, v.1, art.100007, 2019. DOI: https://dx.doi.org/10.1016/j.indic.2019.100007.

DUARTE, P.M.; SANTANA, V.T.P.; DALMAS, A.D.; FERRI, I.E.B. Integração lavoura-pecuária (ILP): uma revisão literária. **Uniciências**, v.22, p.106-109, 2018. DOI: https://dx.doi.org/10.17921/1415-5141.2018v22n2p106-109.

EMBRAPA MILHO E SORGO. **Sistema ILP**: produção de grãos, forragens e carne na região central de Minas Gerais. Sete Lagoas: Embrapa, 2018. Available at: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/193357/1/Sistema-ILP.pdf>. Accessed on: June 6, 2025.

ETIENNE, A.; AHMAD, A.; AGGARWAL, V.; SARASWAT, D. Deep learning-based object detection system for identifying weeds using UAS imagery. **Remote Sensing**, v.13, art.5182, 2021. DOI: https://dx.doi.org/10.3390/rs13245182.

GOMES, A.L.B.; FERNANDES, A.M.R.; HORTA, B.A.C.; OLIVEIRA, M.F. Machine learning algorithms applied to weed management in integrated crop-livestock systems: a systematic literature review. **Advances in Weed Science**, v.42, e020240047, 2024. DOI: https://doi.org/10.51694/AdvWeedSci/2024;42:00004.

HUSSAIN, N.; FAROOQUE, A.A.; SCHUMANN, A.W.; ABBAS, F.; ACHARYA, B.; MCKENZIE-GOPSILL, A.; BARRETT, R.; AFZAAL, H.; ZAMAN, Q.U.; CHEEMA, M.J.M. Application of deep learning to detect Lamb's quarters (*Chenopodium album* L.) in potato fields of Atlantic Canada. **Computers and Electronics in Agriculture**, v.182, art.106040, 2021. DOI: https://dx.doi.org/10.1016/j.compag.2021.106040.

IKEDA, F.S.; MITJA, D.; VILELA, L.; CARMONA, R. Banco de sementes no solo em sistemas de cultivo lavoura-pastagem. **Pesquisa Agropecuária Brasileira**, v.42, p.1545-1551, 2007. DOI: https://dx.doi.org/10.1590/s0100-204x2007001100005.

JHA, K.; DOSHI, A.; PATEL, P.; SHAH, M. A comprehensive review on automation in agriculture using artificial intelligence.

Artificial Intelligence in Agriculture, v.2, p.1-12, 2019. DOI: https://dx.doi.org/10.1016/j.aiia.2019.05.004.

MONTEIRO, A.L.; SOUZA, M.F.; LINS, H.A.; TEÓFILO, T.M.S.; BARROS JÚNIOR, A.P.; SILVA, D.V.; MENDONÇA, V. A new alternative to determine weed control in agricultural systems based on artificial neural networks (ANNs). **Field Crops Research**, v.263, art.108075, 2021. DOI: https://dx.doi.org/10.1016/j.fcr.2021.108075.

RAJA, R.; SLAUGHTER, D.C.; FENNIMORE, S.A.; SIEMENS, M.C. Real-time control of high-resolution micro-jet sprayer integrated with machine vision for precision weed control. **Biosystems Engineering**, v.228, p.31-48, 2023. DOI: https://dx.doi.org/10.1016/j.biosystemseng.2023.02.006.

SABZI, S.; ABBASPOUR-GILANDEH, Y.. Using video processing to classify potato plant and three types of weed using hybrid of artificial neural network and partincle swarm algorithm. **Measurement**, v. 126, p. 22-36, 2018. DOI: https://doi.org/10.1016/j.measurement.2018.05.037.

SANTOS, H.G. dos; JACOMINE, P.K.T.; ANJOS, L.H.C. dos; OLIVEIRA, V.Á. de; LUMBRERAS, J.F.; COELHO, M.R.; ALMEIDA, J.A. de; ARAÚJO FILHO, J.C. de; OLIVEIRA, J.B. de; CUNHA, T.J.F. **Sistema brasileiro de classificação de solos**. 5.ed. rev. e ampl. Brasília: Embrapa, 2018. 356p.

SCHUSTER, M.Z.; LUSTOSA, S.B.C.; PELISSARI, A.; HARRISON, S.K.; SULC, R.M.; DEISS, L.; LANG, C.R.; CARVALHO, P.C.F.; GAZZIERO, D.L.P.; MORAES, A. Optimizing forage allowance for productivity and weed management in integrated crop-livestock systems. **Agronomy For Sustainable Development**, v.39, p.1-10, 2019. http://dx.doi.org/10.1007/s13593-019-0564-4.

SHOREWALA, S.; ASHFAQUE, A.; SIDHARTH, R.; VERMA, U. Weed density and distribution estimation for precision agriculture using semi-supervised learning. **Ieee Access**, v.9, p.27971-27986, 2021. DOI: https://dx.doi.org/10.1109/access.2021.3057912.

THAPA, S.; BHANDARI, A.; GHIMIRE, R.; XUE, Q.; KIDWARO, F.; GHATREHSAMANI, S.; MAHARJAN, B.; GOODWIN, M. Managing micronutrients for improving soil fertility, health, and soybean yield. **Sustainability**, v.13, art.11766, 2021. DOI: https://dx.doi.org/10.3390/su132111766.

VELIN, E.D.; TRINDADE, M.L.B.; CARBONARI, C.A. Precisão amostrai na ciência das plantas daninhas. In: CONGRESSO BRASILEIRO DA CIÊNCIA DAS PLANTAS DANINHAS, 28. Campo Grande, 2012. **Anais**. Campo Grande: Sociedade Brasileira da Ciência das Plantas Daninhas, 2012. p.39-48.

WERTH, J.; KEENAN, M.; THORNBY, D.; BELL, K.; WALKER, S. Emergence of four weed species in response to rainfall and temperature. **Weed Biology and Management**, v.17, p.29-35, 2017. DOI: https://dx.doi.org/10.1111/wbm.12113.

YNOUE, R.Y.; AMBRIZZI, T.; REBOITA, M.S.; SILVA, G.A.M. **Meteorologia**: noções básicas. São Paulo: Oficina de Textos, 2017. 184p.

YU, J.; SCHUMANN, A.W.; CAO, Z.; SHARPE, S.M.; BOYD, N.S. Weed detection in perennial ryegrass with deep learning convolutional neural network. **Frontiers in Plant Science**, v.10, p.1-9, 2019. DOI: https://dx.doi.org/10.3389/fpls.2019.01422.

Author contributions

Ana Letícia Gomes Luz: conceptualization, data curation, methodology, software, writing – original draft; **Anita Maria da Rocha Fernandes**: funding acquisition, project administration, resources, supervision, validation, visualization, writing – review & editing; **Fábio Volkman Coelho**: data curation, resources, software, visualization; **Maurílio Fernandes de Oliveira**: data curation, funding acquisition, resources, supervision, validation, writing – review & editing; **Ramon Costa Alvarenga**: data curation, resources, writing – review & editing.

Chief editor: Edemar Corazza
Edited by: Daniel Kinpara

Data availability statement

Data available upon request: research data are only available upon reasonable request to the corresponding author.

Declaration of use of AI technologies

No generative artificial intelligence (AI) was used in this study.

Conflict of interest statement

The authors declare no conflicts of interest.

Acknowledgments

To Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC), for funding this study (project call FAPESC N 54/2022). To Associação Catarinense das Fundações Educacionais (ACAFE), for the support (Research Groups, N 2023TR000875). To Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), for funding this study (CNPq/MCTI No 10/2023, track A, universal notice, N 404755/2023-2). To Universidade do Vale do Itajaí (UNIVALI), for technical support. To Empresa Brasileira de Pesquisa Agropecuária (Embrapa), for the datasets used in the study. To Instituto Brasileiro de Desenvolvimento e Sustentabilidade (IABS), for partial funding (IDB technical cooperation BR-T1409).

Disclaimer/Publisher's note:

The statements, opinions, and data contained in all texts published in Pesquisa Agropecuária Brasileira (PAB) are solely those of the individual author(s) and not of the journal's publisher, editor, and editorial team, who disclaim responsibility for any injury to people or property resulting from any referred ideas, methods, instructions, or products.

The mention of specific chemical products, machines, and commercial equipment in the texts published in this journal does not imply their recommendation by the publisher.