



RESEARCH

The use of machine learning methods in modelling and optimization blueberry plants (*Vaccinium corymbosum* L.) spraying

Katarzyna Pentoś¹ · Beata Cieniawska¹ · Piotr Komarnicki¹

Received: 13 February 2026 / Accepted: 10 March 2026
© The Author(s) 2026

Abstract

This study explores the application of machine learning and genetic algorithm to model and optimize the spraying process of blueberry plants (*Vaccinium corymbosum* L.). The research aimed to determine the relationship between sprayer operating parameters (nozzle type, liquid pressure, and driving speed), weather conditions (temperature, humidity, and wind speed), and spray quality, specifically the coverage of various plant surfaces. Three ML methods, namely Gradient Boosting Regressor, Support Vector Regressor, and Extreme Gradient Boosting were evaluated to develop predictive models based on experimental data collected using water-sensitive papers. The results indicate that all models achieved high accuracy, with the SVR and XGBoost models being particularly effective, showing a MAPE error as low as 4% for certain surfaces. These models were then integrated into a genetic algorithm to identify optimal operating parameters that maximize total spray coverage under specific meteorological conditions. Optimization results revealed that XR nozzles generally achieved higher total coverage compared to AIXR nozzles. Furthermore, a driving speed of 5.5 km/h was found to be optimal for both nozzle types across various conditions. This hybrid approach demonstrates significant potential for developing intelligent decision support systems in precision agriculture, enabling more efficient and sustainable use of plant protection products.

Keywords Machine learning · Genetic algorithm · Precision agriculture · Blueberry spraying · Process optimization

1 Introduction

Since the use of chemical plant protection products has a negative impact on both the environment and human health, a significant volume of scientific research is currently underway with the objective of optimising this process. A variety of methodologies are employed to reduce the utilization of chemicals whilst ensuring the maintenance of optimal treatment standards. These methods encompass a range of modelling techniques including mechanistic models, statistical modelling methods and machine learning (ML) techniques. Moreover, using optimization algorithms for the reduction of agrochemicals is becoming increasingly prevalent.

A mathematical model of spray deposition developed from Young's equation was discussed by Grayson et al. (1993). The weed-herbicide simulation model was proposed by Xie et al. (2024) to estimate the optimal operating parameters of the spraying system with a view to improving the quality of weed control herbicide application. Dong et al. (2025) developed and verified a mathematical model of energy conservation in the droplet–leaf system. This precise model can be utilised as a foundation for the selection of operating parameters for the spraying equipment employed in the production of pepper. Renaudo et al. (2022) proposed a spray drift model, the parameters of which (including spray pressure, nozzle nominal flow rate and spray angle) can be selected for specific nozzles and particular weather conditions. A single-factor test and Box–Behnken response surface method were used by Wang et al. (2022) to model relationships between spraying parameters and Unmanned Aerial Vehicle (UAV)-based spraying performance. As demonstrated by Xiahou et al. (2020) a Computational Fluid Dynamics-based 3D model was established for the air-blast sprayer. The model developed by the authors can be used to reduce the drift of droplets in air-blast spraying technology.

✉ Katarzyna Pentoś
katarzyna.pentos@upwr.edu.pl
Beata Cieniawska
beata.cieniawska@upwr.edu.pl
Piotr Komarnicki
piotr.komarnicki@upwr.edu.pl

¹ Wrocław University of Environmental and Life Sciences, Wrocław, Poland

An interesting methodology for determining the optimal dose of pesticides based on lethal number of droplets values was proposed by Durairaj and Jyoti (2023). They developed an approach in which the pesticide dose is calculated using a model that takes into account a combination of plant, pest, pesticide, and spray equipment. Another illustration is the utilization of canopy detection-based spraying systems. These systems utilise ultrasonic sensors for the purpose of canopy mapping. These maps form the basis for the variable rate application of spray or systems in which the operating parameters of the sprayer are optimised (Londhe and Sujata 2017; Sawant et al. 2023; Tewari et al. 2018; Yongjun et al. 2017). Xue et al. (2023) used the orthogonal method to determine the optimal parameters for spraying with an air-assisted electrostatic sprayer. One of the optimised parameters was spray coverage on the surfaces of citrus leaves. Furthermore, the utilization of vision systems and models that differentiate between crops and weeds, or that detect plant diseases, has the potential to optimise the application of agrochemicals. These mechanisms enable the precise and targeted application of the spray liquid, ensuring that the minimum amount is dispensed for each application. The system developed by Petelewicz et al. (2024), based on the YOLOv3 neural network in combination with an appropriate nozzle system, has significantly reduced herbicide consumption (by approximately 80%). Peng et al. (2022) proposed a model based on the RetinaNet neural network for detecting weeds in rice crops. Singh et al. (2024) developed a system for the spraying of crops using UAVs, in which they integrated an ML model that recognised pineapple, papaya, and cabbage with the crop-spraying component, which determined the amount of pesticide to spray.

In recent years, there has been an increasing prevalence of machine learning methodologies employed for the purpose of modelling relationships within the agricultural sector. These models are then employed by artificial intelligence (AI) optimization algorithms to determine the optimal values for process parameters. Yang et al. (2023) proposed a combination of a neural network and a genetic algorithm (GA) to model the process of drying-assisted walnut breaking, and subsequently select the optimal parameters for this process. A comparable approach was proposed by Braun et al. (2016) for the optimization of energy production. The authors employed neural networks to model individual subsystems in a complex cogeneration system, subsequently utilising these models within a genetic algorithm to optimise energy production whilst minimising fuel consumption. The combination of ML and artificial AI optimization methods is also employed in the optimization of plant spraying. The Particle Swarm Optimization method was employed by Faiçal et al. (2017) for fine-tuning the control rule of the UAV during the spraying of pesticides. The authors of the study asserted that the implementation of their methodology

enhances the precision of pesticide spraying, resulting in up to 86% of the liquid adhering to the plants.

Current research in the field of precision spraying often ends at the stage of assessing the impact of individual factors on the quality of the treatment. However, there is a lack of research that would determine specific operating recommendations based on the relationship between operating parameters and weather conditions and spray quality indicators. Our research fills this gap by proposing a hybrid ML and GA approach. This represents a step forward towards intelligent decision support systems in agriculture.

The objective of the present study was to ascertain the most accurate models describing the relationship between the operating parameters of spraying blueberry plants (nozzle type, liquid pressure, and driving speed) and weather conditions (air temperature, humidity, and wind speed) and the parameters describing the quality of the treatment. Three machine learning methods were utilised for this purpose. A further objective was to integrate the machine learning model with a genetic algorithm, with a view to constructing a system for determining optimal process parameters in specific weather conditions.

2 Materials and methods

2.1 Spraying conditions

Foliar spraying treatments were carried out under field conditions with stable atmospheric parameters, conducive to minimising spray drift and ensuring uniform deposition of the spray liquid on plant surfaces. Spraying was performed at an air temperature of 17–20 °C, relative humidity of 60–65%, and wind speed not exceeding 1 m·s⁻¹. The treatments were conducted using a field sprayer, applying two operating speeds (2.5 and 5.5 km·h⁻¹) and two spray pressures (0.2 and 0.4 MPa). The experiment included two types of flat-fan nozzles: standard XR nozzles and air-induction AIXR nozzles. Before each pass, the nozzle flow rate was checked by measuring the volume of liquid discharged per unit time and comparing the obtained values with the manufacturer's specifications. In the event of deviations, the operating pressure was adjusted to maintain uniform application conditions.

The study was conducted on eight experimental plots measuring 1.5 m in width and 10 m in length, with a 2 m distance between rows. Each plot contained 11 blueberry bushes. For experimental purposes, three zones were designated within each plot. In the first zone (2.5 m), the sprayer reached the target operating speed; the second zone (6 m) constituted the measurement zone; and in the third zone (1.5 m), the sprayer was brought to a stop.

2.2 Assessment of the coverage of sprayed surfaces

The degree of spray liquid coverage was assessed using an indirect method based on water-sensitive papers (WSP). In the measurement section, three dedicated frames were installed, to which the water-sensitive papers were attached in a manner that allowed the formation of vertical and horizontal surfaces representing four zones of spray–plant contact (Fig. 1). Each experimental variant included three replicates, achieved by placing three frames. After spraying, the water-sensitive papers were removed, air-dried, and digitized by scanning. Image analysis was performed using graphic software, with three randomly selected fragments of 1 cm² in area. Each measurement yielded nine results, from which the three most outlying values were removed. Following this operation, a total of six results were obtained for each combination of operating parameters and weather conditions. Based on the analysis conducted using Adobe Photoshop software, the proportion of the surface area covered by the spray liquid relative to the total analysed area was determined.

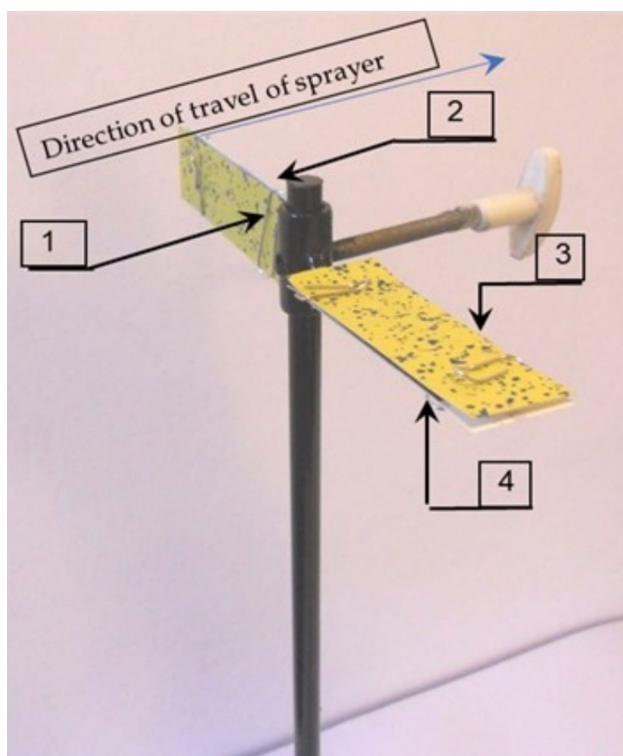


Fig. 1 Frames with water-sensitive papers: 1– vertical approach surface, 2 – vertical leaving surface, 3 – horizontal upper surface, 4 – horizontal bottom surface

2.3 Machine learning methods

2.3.1 Gradient Boosting Regressor (GBR)

This algorithm belongs to the family of ensemble methods, which construct a robust predictive model through the successive integration of numerous weaker models that are subject to training. The boosting process functions by building each subsequent tree to correct the errors (residuals) made by the preceding one. Specifically, each new tree is trained to minimise the loss function using the negative gradient of this function with respect to the current ensemble prediction, which is key to effectively reducing model bias. Precise calibration of GBR necessitates the optimization of several hyperparameters. These include the number of trees in the model ($n_{\text{estimators}}$), where a higher count improves predictive accuracy but increases the computational cost and the risk of overfitting; the learning rate, which scales the contribution of each new tree, acting as a crucial step-size regulator for the model's learning speed; the maximum tree depth (max_depth), which controls the complexity of individual trained trees; the fraction of training samples used for each sequential tree (subsample), which introduces a stochastic component to the gradient descent, effectively preventing overfitting by increasing the variance of the individual trees and improving generalization.

2.3.2 Support Vector Regressor (SVR)

SVR is a regression adaptation of the popular Support Vector Machine (SVM) and is used to predict continuous values. In contrast to conventional regression methodologies, SVR does not seek to minimise the sum of squares of all errors. Instead, it seeks to identify the optimal hyperplane that will minimise error whilst maintaining the simplicity and the generalization capability of the model. This is achieved by the introduction of an error margin. The error margin constitutes an insensitivity zone within which errors are tolerated without consequence. The model is defined solely by support vectors, which are defined as key data points that lie on or outside the boundary of this margin. In the context of non-linear relationships, the SVR model employs a kernel function, such as the radial basis function (RBF), to transform data into a higher-dimensional space where linear regression can be applied. The performance and complexity of the SVR model are fundamentally impacted by three main hyperparameters: regularization parameter (C), kernel parameter (γ), and error margin (ϵ). The regularization parameter is instrumental in determining the equilibrium between the smoothness of the regression function and the precision of its alignment with the training data. Higher C value corresponds to a more substantial penalty for deviations outside the permissible margin. A high C value is indicative of a

more precise fit and greater complexity, while a low C value favours generalization and a simpler model. In the context of the RBF kernel, γ is the critical factor in determining the range of influence of a single support vector. The value of this parameter is indicative of the extent to which the influence of a single data point is propagated. A small γ value signifies a broad range and a more flexible model, whereas a large γ value leads to a more constrained range, resulting in a model that is susceptible to local fluctuations and may exhibit overfitting. The error margin delineates the extent of the insensitivity zone, within which errors are disregarded and data points are considered correctly predicted. A smaller value of ϵ necessitates a highly precise fit, incorporating a greater number of feature vectors into the model. Conversely, a larger value of ϵ results in a more straightforward model that is capable of tolerating a higher level of error.

2.3.3 Extreme Gradient Boosting (XGBoost)

XGBoost is a highly efficient and effective machine learning algorithm based on decision trees, belonging to the family of gradient boosting methods. The core methodology involves the sequential construction of weak predictors (trees), with each new tree correcting the errors made by the aggregate of the previous ones. The results of these trees are summed to create a strong predictive model. The efficacy and scalability of this algorithm rely heavily on the precise calibration of its hyperparameters. Key parameters include the number of trees built ($n_{\text{estimators}}$), which impacts model complexity and the risk of overfitting, the maximum depth of a single tree (max_depth), controlling model complexity, and learning rate determining the weight of new trees, which helps prevent overfitting by moderating the learning speed. The parameter γ_{XGB} introduces tree-node regularization by specifying the minimum loss reduction required to perform a further split. Finally, the selection of the sampling method and the booster type defines the overall architecture and training approach of the model.

2.4 Methodology for the development of machine learning models

To develop optimal machine learning (ML) models for each of the algorithms, multiple training runs were performed for different combinations of hyperparameters. For the purpose of this study, a grid search method was employed to train models for each configuration of hyperparameters, with successive values assigned according to the data presented in Table 1.

The present study utilised five statistical metrics in order to evaluate the performance and accuracy of the neural network models. The metrics employed were the coefficient of determination (R^2), root mean squared error (RMSE),

Table 1 Hyperparameter values for individual models used in the grid search method

Hyperparameter	Values
Gradient Boosting Regressor	
$n_{\text{estimators}}$	100; 500; 1000; 1500
learning rate	0.01; 0.02; 0.03; 0.04
max_depth	4; 6; 8; 10
subsample	0.1; 0.2; 0.5; 0.9
Support Vector Regressor	
regularization parameter C	100; 200; 300; 400; 500; 600; 700; 800; 900; 1000; 3000; 5000; 10000
kernel parameter γ	0.001; 0.01; 0.1; 0.2; 0.3
error margin ϵ	0.01; 0.1; 0.5
kernel function	RBF
Extreme Gradient Boosting	
$n_{\text{estimators}}$	from 100 to 1000 with grain 100
learning rate	0.01; 0.05; 0.2; 0.3
max_depth	from 2 to 10 with grain 2
γ_{XGB}	0; 0.1; 0.2; 0.3; 0.4
sampling method	uniform; gradient_based
booster type	gbtree; gblinear; dart

mean absolute error (MAE), mean absolute percentage error (MAPE), and the Nash–Sutcliffe efficiency coefficient (NSC). The calculation of these evaluation metrics is outlined in Eqs. (1–5).

$$R^2 = \frac{\sum (Y_P - \bar{Y}_M)^2}{\sum (Y_M - \bar{Y}_M)^2} \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_P - Y_M)^2} \quad (2)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Y_P - Y_M| \quad (3)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_M - Y_P}{Y_M} \right| \quad (4)$$

$$\text{NSC} = 1 - \left[\frac{\sum_{i=1}^n (Y_M - Y_P)^2}{\sum_{i=1}^n (Y_M - \bar{Y}_P)^2} \right] \quad (5)$$

where Y_P is the absolute value calculated by the model, $\bar{Y}_P$ is the arithmetic mean of values calculated by the model, Y_M is the absolute experimental value, $\bar{Y}_M$ is the arithmetic

mean of experimental values, and n is the number of vectors in data set.

The data utilised for model training comprised one categorical variable (nozzle type) and five numerical variables (liquid pressure, driving speed, air temperature, air humidity, and wind speed). The categorical variable was encoded using the label encoding method. Prior to the training of the models, the data underwent a process of normalization.

The models were trained in the Python environment [3.12.3]. The GradientBoostingRegressor and SVR from the Scikit-learn package [1.8.0] (Pedregosa et al. 2011), and the XGBRegressor from the XGBoost package [3.1.2] (Chen and Guestrin 2016) were used. To ensure full reproducibility of the results, a fixed random seed of 42 was applied across all models and data splitting procedures. A tenfold nested cross-validation framework was utilised to optimize the model and ensure a robust, unbiased performance evaluation. The data was divided into 10 folds in the outer loop for final performance estimation (the dataset containing 78 observations was partitioned into ten subsets, and the training process was repeated ten times). Within each outer training fold, the optimal hyperparameters were established via grid search method. To prevent data leakage, feature scaling and model training were encapsulated within a computational pipeline. This approach ensured that all preprocessing parameters were derived exclusively from the training subsets in each iteration, guaranteeing that the error metrics reflect the model's true predictive capabilities on unseen data. The models' performance was assessed using the average value of the error metrics obtained from all ten iterations.

Furthermore, a feature importance analysis was performed on the optimal tree-based models obtained, employing the Permutation Importance method.

2.4.1 Optimization

The optimization of parameters of spraying was conducted with the use of genetic algorithm. GAs belong to the class of optimization algorithms inspired by the processes of biological evolution and natural selection. They constitute heuristic search methods for exploring the solution space, being particularly effective for complex, non-linear, discontinuous, or multi-dimensional optimization problems. The search process for an optimal solution begins with the generation of an initial, random population of individuals, where each individual represents a potential solution. Subsequently, in sequential iterations known as generations, the population is evolved using fundamental genetic operators: selection, crossover, and mutation. Selection favors individuals with a better fitness function value, which measures the quality of the solution. Crossover allows for the exchange of genetic

information between two selected individuals to create new offspring, combining the parents' characteristics. Mutation introduces random changes to the genome, thus preventing premature convergence to a local optimum and ensuring exploration of the solution space. This cycle repeats until specified termination criteria are met, leading the population towards an optimal or sub-optimal solution (Immanuel and Chakraborty 2019; Lambora et al. 2019).

The aim of the optimization was to select the spraying process parameters, with the objective of maximization of total coverage of all plant surfaces for a given nozzle and specific weather conditions. Two parameters were optimised: liquid pressure and driving speed (with fixed values for nozzle type, air temperature, air humidity, and wind speed). The objective function was defined as the total coverage of the sprayed surfaces of artificial plants. The surface coverage was calculated using the most accurate machine learning models obtained for each surface. For each nozzle, optimization was performed for each combination of extreme values of weather conditions. The quality of the solution calculated by the genetic algorithm depends directly on the values of the hyperparameters. Therefore, a multitude of configurations of these hyperparameters were subjected to rigorous testing using the grid search method. The population size was modified from 40 to 100 at an interval of 30, the number of iterations was adjusted from 50 to 450 at an interval of 200, the crossover probability was varied from 0.6 to 1 at an interval of 0.2, and the mutation probability was modified from 0.05 to 0.25 at an interval of 0.1. The genetic algorithm was implemented in the Python environment using the DEAP (Distributed Evolutionary Algorithms in Python) library.

2.4.2 Statistical analysis

Statistical analyses were performed to compare the predictive performance of the three evaluated machine learning models. For each model, the error metrics obtained from all 10 folds of the cross-validation process were treated as independent observations ($n = 10$ per model). The normality of the data distribution was assessed using the Shapiro–Wilk test, and the homogeneity of variances was verified with Levene's test. If both assumptions were met ($\alpha > 0.05$), a one-way Analysis of Variance (ANOVA) was conducted, followed by Tukey's Honestly Significant Difference (HSD) post-hoc test for pairwise comparisons. In cases where the assumptions were violated, the non-parametric Kruskal–Wallis H-test was employed, supplemented by Dunn's post-hoc test with Bonferroni adjustment.

To provide a robust estimation of the mean performance, 95% confidence intervals were calculated using a bootstrap method with 10,000 iterations. All statistical procedures

were implemented in Python 3.12.3. A significance level of $p < 0.05$ was used for all tests.

3 Results

The Table 2 provides detailed statistics on the coverage of the areas under study, including the total coverage of all areas. The bottom level surface is not presented in the Table 2 because it has not been covered with spray liquid.

The greatest coverage was obtained for the upper level surface. It ranged from 13.12% to 38.95%, with an average of 24.62%. The surface that was covered with liquid to the least extent was the vertical leaving surface. The mean percentage coverage for this surface was a mere 5.32%. The vertical approach surface was covered with liquid to a slightly greater extent, with an average of 8.58%.

Table 2 The statistics of experimental data

The Parameter	Minimum	Maximum	Mean	Standard Deviation
PAC_ul	13.12	38.95	24.62	6.94
PAC_vta	3.66	15.70	8.58	3.12
PAC_vtl	0.88	11.25	5.32	2.95
PAC_total	21.60	57.72	38.53	8.81

PAC_ul [%]—the coverage of the upper level surface; PAC_vta [%]—the coverage of the vertical approach surface; PAC_vtl [%]—the coverage of the vertical leaving surface; PAC_total = PAC_vta + PAC_vtl + PAC_ul.

Table 3 Hyperparameters of the best ML models

Hyperparameter	Value		
	Upper level surface	Vertical approach surface	Vertical leaving surface
Gradient Boosting Regressor			
n_estimators	1000	1000	1000
learning rate	0.02	0.04	0.02
max_depth	4	6	6
subsample	0.2	0.5	0.5
Support Vector Regressor			
regularization parameter C	200	3000	1000
kernel parameter γ	0.2	0.1	0.3
error margin ϵ	0.1	0.1	0.01
Extreme Gradient Boosting			
n_estimators	700	100	300
learning rate	0.01	0.2	0.3
max_depth	4	4	4
γ _XGB	0.1	0.1	0.1
sampling method	uniform	uniform	uniform
booster type	dart	dart	gbtree

3.1 Machine learning models

To obtain models with optimal accuracy, a range of models were trained using the grid search method. These models differed in their parameter combinations, and the range of these combinations is specified in Table 1. Furthermore, for each model, the data was divided into training and test sets using the tenfold cross-validation method. The Table 3 presents the hyperparameters of the most successful models obtained for each of the machine learning algorithms.

Tables 4, 6, and 8 present the error metric values obtained for the models, the hyperparameters of which are presented in Table 3. Since the tenfold cross-validation method was used in training the models, the tables show not only the mean values and standard deviations of individual metrics, but also the 95% confidence interval, which provides a range of values within which we are 95% confident the true population value of that metric lies.

As shown in Tables 4, 6, and 8, models with very high accuracy were obtained for all the artificial plant surfaces studied. For the upper-level surface, all machine learning algorithms produced models with a mean absolute percentage error (MAPE) of approximately 4% and an R^2 coefficient exceeding 0.96. For the other two surfaces, the MAPE error was below 10% and the R^2 coefficient was above 0.9. All ML methods used in the study produced models of a similar level of accuracy. Statistical analyses showed that there were no statistically significant differences in the error metric values obtained for the individual models (see Tables 5, 7, and 9).

Table 4 Error metrics of the best ML models for test data set for the upper level surface

ML model	Error metrics (mean $\pm$ standard deviation and 95% confidence interval)				
	MAPE	RMSE	MAE	NSC	R ²
Gradient Boosting	0.041 $\pm$ 0.011 [0.035, 0.047]	1.157 $\pm$ 0.312 [0.975, 1.342]	0.980 $\pm$ 0.299 [0.810, 1.160]	0.959 $\pm$ 0.029 [0.941, 0.975]	0.969 $\pm$ 0.0201 [0.956, 0.981]
SVR	0.040 $\pm$ 0.013 [0.032, 0.047]	1.146 $\pm$ 0.369 [0.925, 1.349]	0.961 $\pm$ 0.323 [0.769, 1.145]	0.958 $\pm$ 0.031 [0.939, 0.975]	0.967 $\pm$ 0.026 [0.950, 0.981]
XGBoost	0.042 $\pm$ 0.011 [0.035, 0.048]	1.157 $\pm$ 0.304 [0.981, 1.336]	0.993 $\pm$ 0.290 [0.827, 1.162]	0.960 $\pm$ 0.026 [0.944, 0.975]	0.968 $\pm$ 0.021 [0.956, 0.979]

Table 5 Statistical test results for comparisons of models presented in Table 4

Model 1	Model 2	<i>p</i> -Value				
		MAPE	RMSE	MAE	NSC	R ²
Gradient Boosting	SVR	0.984	0.999	0.989	0.992	0.966
Gradient Boosting	XGBoost	0.993	0.999	0.995	0.998	0.994
SVR	XGBoost	0.957	0.999	0.969	0.983	0.988

To rigorously evaluate the models' stability, learning curves were generated for the best-performing models (Fig. 2). The analysis reveals that as the number of training samples increases, the cross-validation MAPE consistently decreases, progressively approaching the training error curve. This narrowing gap between the two curves is a strong indicator of low variance and suggests that the models have successfully captured the underlying physical patterns of the spraying process rather than merely memorizing the noise in the training data. The observed convergence confirms that the reported high predictive accuracy (with $R^2 > 0.9$ and low MAPE) is statistically robust (Tables 4, 5, 6, 7, 8, and 9).

Figure 3 shows the graphs illustrating the agreement between the target and predicted values of coverage for the best models obtained for each surface for the test set. The SVR model was selected for the upper level surface and the vertical leaving surface, while the XGBoost model was selected for the vertical approach surface. Due to the tenfold cross-validation method used in the training of the models, the plots show the summary results for all 10 test subsets (folds).

To further validate the reliability of models, a feature importance analysis was performed based on models developed for sprayed surfaces (Fig. 4). The results indicated that the models' predictions were primarily driven by key operating parameters, such as nozzle type, liquid pressure and driving speed, rather than weather conditions. Notably, the Permutation Importance analysis confirmed that even for surfaces with lower spray coverage, the models maintained their predictive integrity by relying on the same physical drivers. This consistency across different coverage scales demonstrates that the high R^2 values are a reflection of the models' ability to capture the underlying physics of the

spraying process, rather than an artifact of limited variance in the target variable.

3.2 Optimization

The optimization of the spraying process entailed the identification of process parameter values that would maximise the quality of the spraying process. The quality parameter was the coverage of individual surfaces of sprayed plants. Consequently, the sum of the coverage of three surfaces was adopted as the objective function in the genetic algorithm: $PAC_{total} = PAC_{ul} + PAC_{vta} + PAC_{vtl}$. The experimental trials were conducted under a range of meteorological conditions that could not be modified in the field. Therefore, the optimization process entailed the selection of suitable values for liquid pressure and driving speed, in accordance with the established parameters of air temperature, air humidity, and wind speed. The optimization was conducted in a distinct manner for the XR and AIXR sprayers. The coverage values of individual surfaces in the objective function were calculated using the best ML models, namely the SVR model for PAC_{ul} and PAC_{vtl} and the XGBoost model for PAC_{vta} . The results are displayed in Table 10.

The data presented in Table 10 indicates that it is possible to achieve higher PAC_{total} values with the XR nozzle. For both nozzles, the optimal driving speed is $5.5 \text{ km} \cdot \text{h}^{-1}$. This is the highest value that was observed during the tests. For the XR nozzle, the optimal liquid pressure values are found to be contingent on weather conditions, though no clear correlation can be observed. In general, lower liquid pressure values (up to 0.25 MPa) should be used for higher air humidity. However, at an air humidity of 65%, an air temperature of 20°C and a wind speed of $1 \text{ m} \cdot \text{s}^{-1}$, a high pressure value of 0.4 MPa is

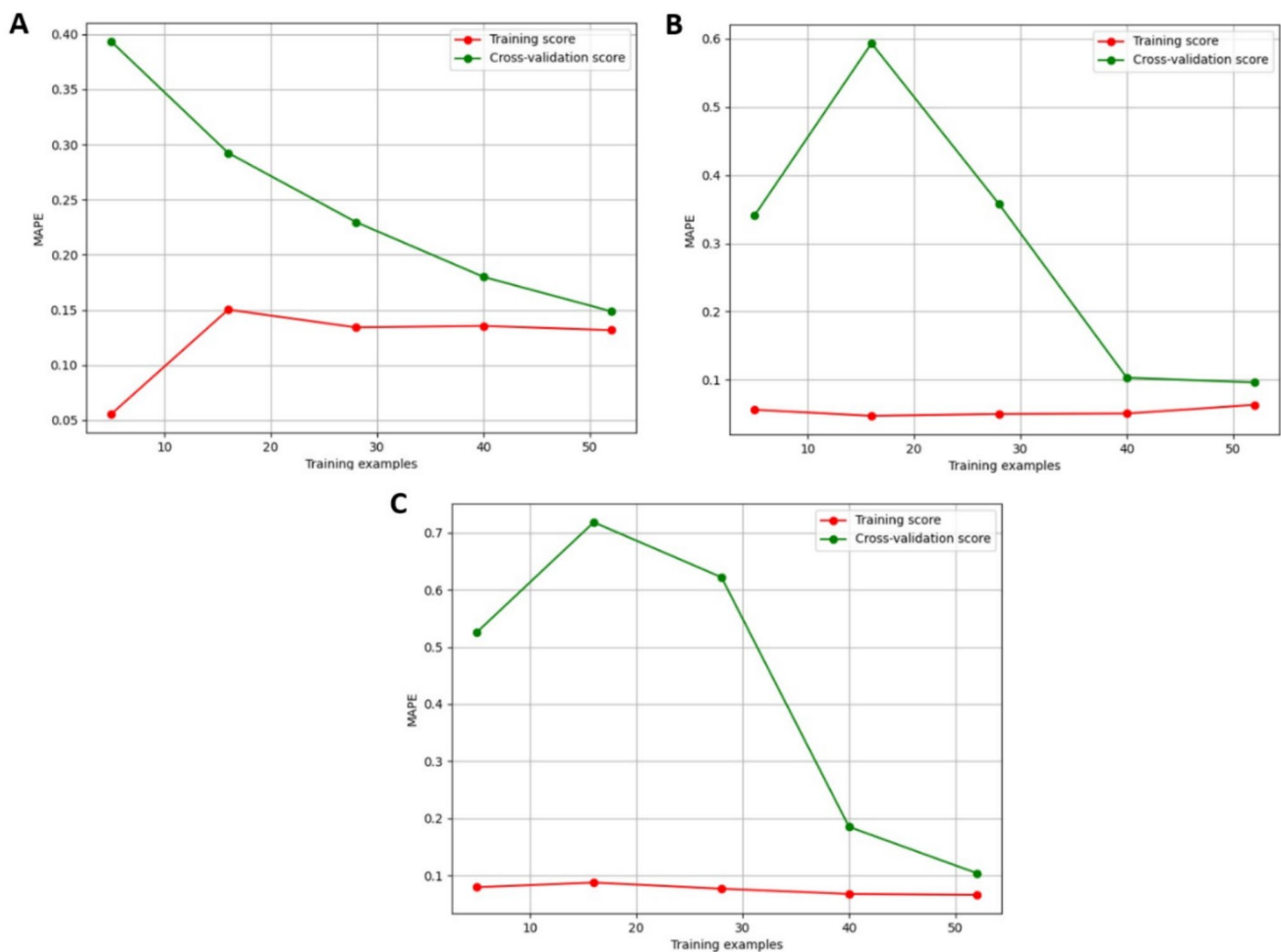


Fig. 2 Learning curves for the best models illustrating the MAPE error as a function of the training examples size. **A**—the upper level surface, **B**—the vertical approach surface, **C**—the vertical leaving surface

Table 6 Error metrics of the best ML models for test data set for the vertical approach surface

ML model	Error metrics (mean $\pm$ standard deviation and 95% confidence interval)				
	MAPE	RMSE	MAE	NSC	R ²
Gradient Boosting	0.093 $\pm$ 0.027 [0.077, 0.109]	0.839 $\pm$ 0.172 [0.736, 0.939]	0.706 $\pm$ 0.167 [0.608, 0.806]	0.824 $\pm$ 0.294 [0.630, 0.940]	0.918 $\pm$ 0.082 [0.862, 0.956]
SVR	0.093 $\pm$ 0.028 [0.076, 0.109]	0.842 $\pm$ 0.161 [0.745, 0.935]	0.694 $\pm$ 0.175 [0.591, 0.799]	0.805 $\pm$ 0.372 [0.562, 0.939]	0.916 $\pm$ 0.086 [0.856, 0.956]
XGBoost	0.092 $\pm$ 0.027 [0.076, 0.107]	0.836 $\pm$ 0.171 [0.734, 0.935]	0.700 $\pm$ 0.165 [0.604, 0.796]	0.831 $\pm$ 0.274 [0.651, 0.939]	0.919 $\pm$ 0.080 [0.865, 0.957]

Table 7 Statistical test results for comparisons of models presented in Table 6

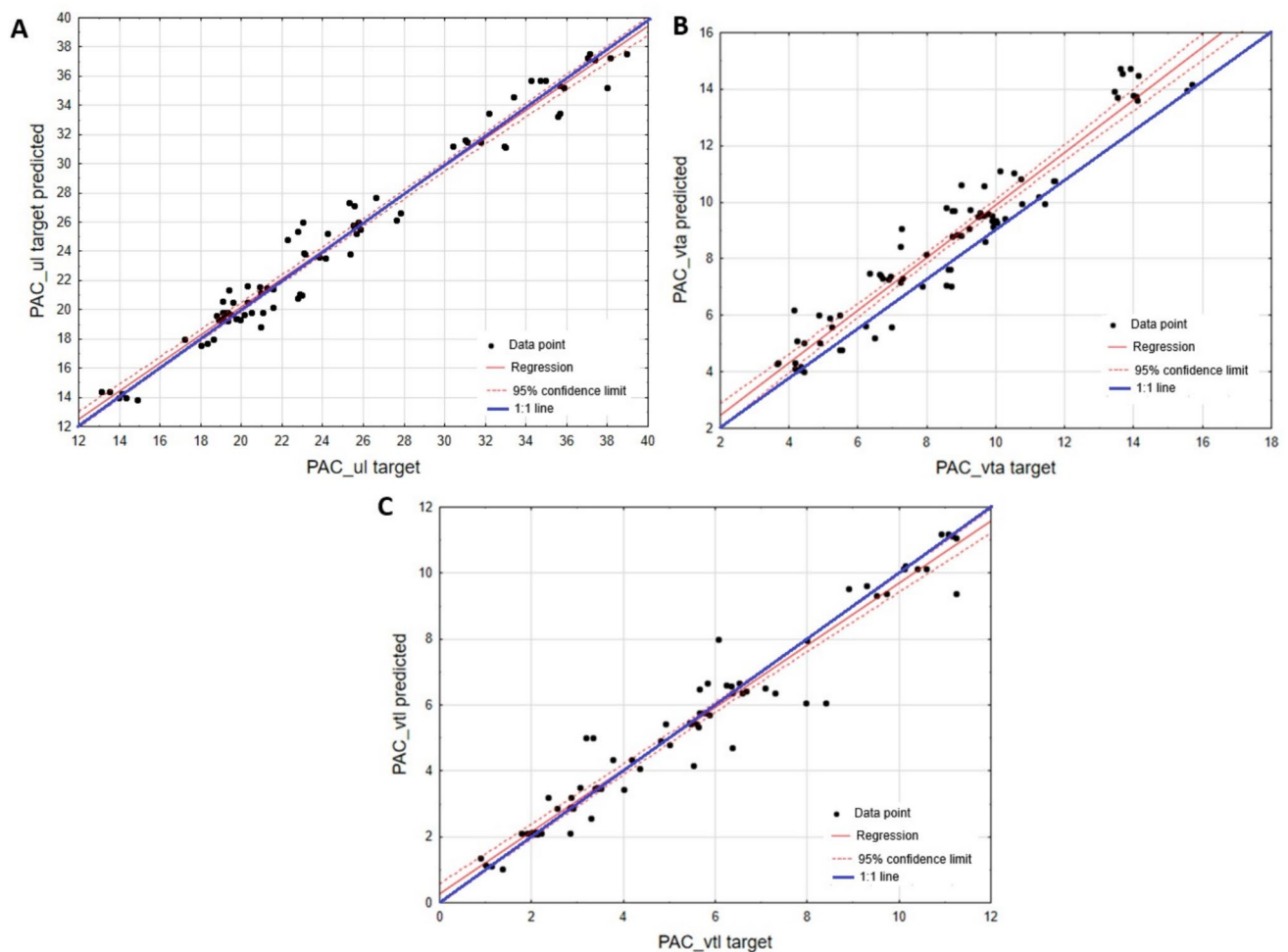
Model 1	Model 2	p-Value				
		MAPE	RMSE	MAE	NSC	R ²
Gradient Boosting	SVR	0.999	0.999	0.987	0.999	0.999
Gradient Boosting	XGBoost	0.993	0.999	0.997	0.999	0.999
SVR	XGBoost	0.992	0.997	0.997	0.999	0.999

Table 8 Error metrics of the best ML models for test data set for the vertical leaving surface

ML model	Error metrics (mean $\pm$ standard deviation and 95% confidence interval)				
	MAPE	RMSE	MAE	NSC	R ²
Gradient Boosting	0.095 $\pm$ 0.026 [0.079, 0.110]	0.628 $\pm$ 0.164 [0.527, 0.717]	0.415 $\pm$ 0.092 [0.358, 0.465]	0.944 $\pm$ 0.033 [0.925, 0.963]	0.957 $\pm$ 0.031 [0.939, 0.975]
SVR	0.092 $\pm$ 0.033 [0.072, 0.111]	0.631 $\pm$ 0.235 [0.498, 0.774]	0.401 $\pm$ 0.143 [0.322, 0.491]	0.946 $\pm$ 0.032 [0.927, 0.965]	0.959 $\pm$ 0.030 [0.941, 0.976]
XGBoost	0.094 $\pm$ 0.027 [0.078, 0.110]	0.638 $\pm$ 0.171 [0.532, 0.730]	0.418 $\pm$ 0.100 [0.358, 0.471]	0.942 $\pm$ 0.034 [0.923, 0.962]	0.957 $\pm$ 0.033 [0.937, 0.975]

Table 9 Statistical test results for comparisons of models presented in Table 8

Model 1	Model 2	<i>p</i> -Value				
		MAPE	RMSE	MAE	NSC	R ²
Gradient Boosting	SVR	0.999	0.999	0.959	0.993	0.992
Gradient Boosting	XGBoost	0.999	0.994	0.998	0.993	0.998
SVR	XGBoost	0.999	0.997	0.938	0.972	0.981

**Fig. 3** The scatter plot of predicted values versus measured values of coverage. **A**—the upper level surface, **B**—the vertical approach surface, **C**—the vertical leaving surface

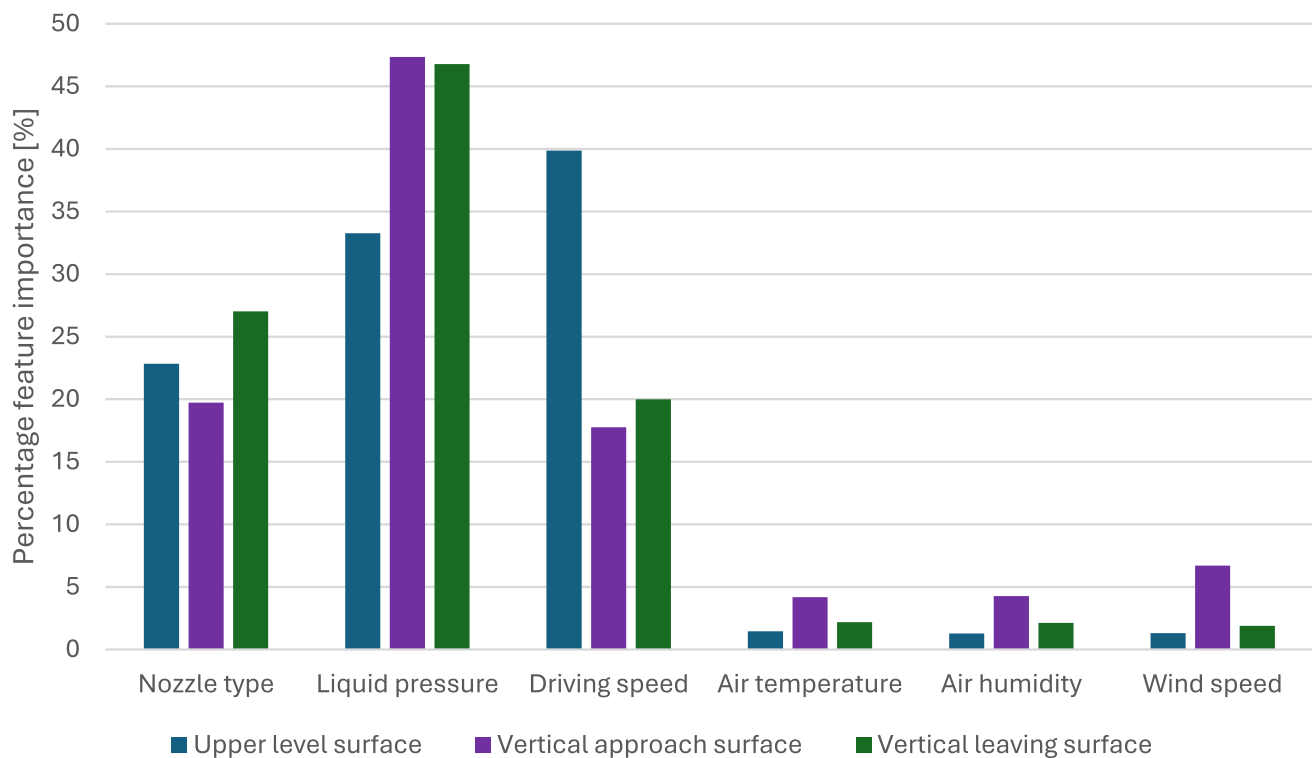


Fig. 4 The percentage feature importance calculated for sprayed surfaces

Table 10 The results of optimization

Nozzle type	Liquid pressure [MPa]	Driving speed [km·h ⁻¹]	Air temperature [°C]	Air humidity [%]	Wind speed [m·s ⁻¹]	PAC_ul [%]	PAC_vta [%]	PAC_vtl [%]	PAC_total [%]
XR	0.30	5.5	17	60	0.5	20.39	8.70	5.65	34.75
XR	0.40	5.5	17	60	1.0	20.76	7.43	3.71	31.89
XR	0.25	5.5	17	65	0.5	25.05	7.22	2.79	35.06
XR	0.22	5.5	17	65	1.0	23.45	7.22	3.86	34.53
XR	0.39	5.5	20	60	0.5	21.36	7.45	3.05	31.87
XR	0.40	5.5	20	60	1.0	21.72	4.32	1.11	27.15
XR	0.22	5.5	20	65	0.5	22.72	7.19	4.66	34.58
XR	0.40	5.5	20	65	1.0	20.90	7.43	5.24	33.57
AIXR	0.26	5.5	17	60	0.5	18.58	6.02	2.73	27.33
AIXR	0.27	5.5	17	60	1.0	15.99	6.11	2.58	24.68
AIXR	0.26	5.5	17	65	0.5	19.74	5.64	2.94	28.33
AIXR	0.24	5.5	17	65	1.0	17.15	5.74	2.80	25.68
AIXR	0.28	5.5	20	60	0.5	16.53	6.12	2.16	24.81
AIXR	0.27	5.5	20	60	1.0	13.94	5.84	2.01	21.79
AIXR	0.24	5.5	20	65	0.5	17.69	5.82	2.37	25.88
AIXR	0.28	5.5	20	65	1.0	15.10	6.11	2.22	23.43

optimal. It is also noteworthy that, in general, high liquid pressures are recommended at high wind speeds. Conversely, for AIXR nozzles, the optimal liquid pressure for all weather conditions is comparatively low, ranging from

0.24 MPa to 0.28 MPa. It should be noted that Table 10 shows the best results achieved. Undoubtedly, alternative operating parameter values can be identified that guarantee results close to optimal.

The performance of the genetic algorithm was evaluated through convergence analysis. The convergence curves for the best and average fitness obtained for XR nozzle, air temperature = 20 °C, air humidity = 60%, and wind speed = 0.5 m·s⁻¹ is presented on the Fig. 5. The optimization process consistently reached a stable solution within 60 to 100 generations, depending on the meteorological scenario. A characteristic sharp increase in the fitness function was observed in the initial phase, followed by a plateau where changes in the best fitness were negligible. This behavior, combined with the convergence of the mean population fitness towards the best individual, confirms the robustness of the GA in navigating the continuous search space of pressure and driving speed.

4 Discussion

Spray coverage of sprayed surfaces under field conditions shows significant variation depending on the orientation of plant leaves, the type of nozzle, and sprayer operating parameters. The highest coverage was obtained on the upper horizontal surfaces, while considerably lower coverage was observed on vertical surfaces – both those facing towards the direction of sprayer travel and those oriented away. No liquid coverage was detected on the bottom horizontal surfaces. Similar relationships were reported by Dekeyser

et al. (2014), Sayinci et al. (2020), and Wang et al. (2025). Researchers have indicated that gravity and spray direction determine the asymmetry of liquid distribution on the target, especially when using single-stream flat-fan nozzles. Analysis of our results reveals a significant influence of meteorological conditions and sprayer operating parameters. Moderate wind speeds ranging from 0.5 to 1 m·s⁻¹ facilitated droplet penetration into the plant canopy. However, even at these low velocities, some deviation of the spray jet occurred, which may have limited coverage of the leaving vertical surfaces. Similar conclusions were drawn by Bock and Hotchkiss (2021), emphasizing that even minor air turbulence can substantially alter deposition patterns within the plant canopy. Similar trends were observed by Cieniawska and Pentoś (2021), who compared single- and dual-stream nozzles. In the case of dual-stream nozzles greater coverage of vertical leaf surfaces was achieved, attributed to improved penetration of spray liquid into deeper canopy layers. In that research, conducted using single-stream nozzles, the results confirm that a unidirectional spray limits coverage efficiency on hard-to-reach vertical surfaces. Regarding the influence of operating pressure, the observed increase in coverage with higher pressure is consistent with the findings of Ferguson et al. (2020) and Subr et al. (2020). Scientists have demonstrated that higher pressure promotes liquid atomization and increases the number of droplets per unit area. However, at higher pressures, the proportion of fine droplets, which are

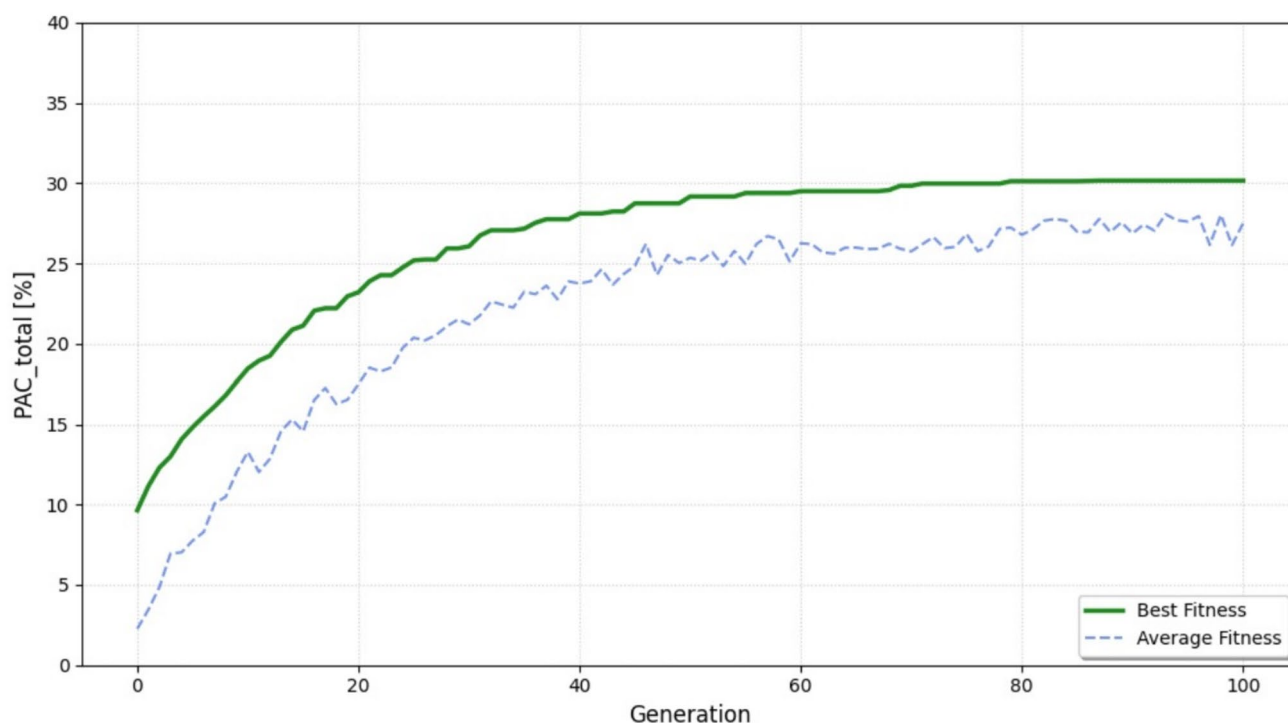


Fig. 5 Example convergence curve of the genetic algorithm for total spray coverage optimization

more susceptible to drift, also increases, potentially reducing actual coverage and spray deposition on plants under field conditions.

The research findings therefore confirm the necessity of a compromise between coverage intensity, environmental safety, and spray stability. The best results were achieved at moderate pressure and operating speed, which is consistent with the recommendations of Bolat and Özlüoymak (2020) and Sayinci et al. (2020). Analysis of coverage distribution also indicates that spray directionality remains one of the key factors limiting application efficacy on vertical surfaces, which can potentially be improved by using dual-stream nozzles or systems with adjustable spray angles. In summary, the obtained results align with current research trends focused on optimizing the technical parameters of field spraying. They confirm that the uniformity and effectiveness of the treatment depend both on the physical properties of the spray liquid and on the selection of an appropriate nozzle type, its operating pressure, spraying speed, as well as on the spatial arrangement of the sprayed surfaces.

In the analysis of the quality of models obtained for individual sprayed surfaces, it was found that the highest quality models were obtained for the upper level surface. For all algorithms utilised, the MAPE metric is approximately 4%, and the determination coefficient R^2 is approximately 0.96. For the remaining surfaces, the error is greater, but its value (approximately 9% for all algorithms used) indicates high prediction quality and the usefulness of the models in practical applications. This assertion is corroborated by other error metrics. According to the classification proposed by Lewis (1982), a MAPE error value below 10% is indicative of highly accurate forecasting. The developed models achieved remarkably high R^2 values; however, it should be noted that this indicator may remain high despite large but proportional differences between target and predicted values (Legates and McCabe 1999). Therefore, the metric is complemented by the Nash–Sutcliffe efficiency coefficient, which is indicative of the degree to which the plot of target versus predicted data aligns with the 1:1 line. The NSC can take values ranging from $-\infty$ to 1 (Moriassi et al. 2007), and for the models obtained in this study, it approaches the optimal value. The RMSE and MAE metric values should be directly related to the range of the measured parameter. As Singh et al. (2004) assert, values of these metrics falling below half the standard deviation of the measured data can be considered low. In the case of the coverage degree analysed in the present study, the standard deviations are 6.94, 3.12 and 2.94 for the upper level, vertical approach and vertical leaving surface, respectively. The obtained RMSE and MAE metric values are significantly lower than half the standard deviation (1.146 and 0.961 vs $0.5SD = 3.47$; 0.836 and 0.700 vs $0.5SD = 1.56$; 0.628 and 0.415 vs $0.5SD = 1.47$, respectively). Consequently, it can be deduced that all

machine learning algorithms employed yielded very high-quality models for all three sprayed surfaces. However, substantial discrepancies were observed in the optimal values of certain hyperparameters of the most effective models, contingent on the predicted parameter.

For the Gradient Boosting Regressor, the optimal maximum depth was found to be low, at 4 or 6, for all three models. This finding suggests that the algorithm effectively employs the concept of weak learners and that the relationships under investigation are non-linear, yet their structure is effectively approximated by moderately complex feature interactions. The optimal subsample value was determined to be 0.2 or 0.5 (utilising between 20 and 50% of the training set instances to construct each subsequent tree). The presence of such low values for this parameter introduces an element of randomness that serves to significantly reduce the variance of the model. The consistent selection of 1000 estimators as the optimal value in all models indicates that the learning process has reached a point of stabilization at the given learning rate parameter values. The consistency of the optimal hyperparameters for all three models confirms the stability of the proposed architecture and the repeatability of the modelling process (Friedman 2001; Hastie et al. 2009). For SVR models, the greatest variability was observed for the regularization parameter C, which ranged from 200 to 3000. A high value for the vertical approach surface model indicates a more complex, non-linear data structure requiring less tolerance for deviations. A low value for the upper level surface model promotes better model generalization and resistance to noise in the data. The presence of an error margin of 0.1 for the upper level surface and vertical approach surface models indicates that the data is subject to a certain degree of noise and that the model permits minor deviations without imposing a penalty in the loss function. Conversely, a value of 0.01 for the vertical leaving surface model may signify low measurement error variance in the data (Cherkassky and Ma, 2004; Smola and Schölkopf 2004). For the XGB algorithm, a strong inverse correlation between the number of estimators and the learning rate can be observed. Differences in the selection of these parameters for individual sprayed areas may indicate that the training data for these models differ in the complexity of the input–output relationship. Despite large differences in learning rates, all three models showed an identical need for tree-node regularization (tree-node regularization = 0.1). This indicates the need to use a constant, albeit moderate, tree complexity reduction mechanism to prevent overfitting. The choice of the uniform booster type suggests that the dataset is characterised by a relatively homogeneous error distribution and that there was no need to use more advanced booster techniques (Probst et al. 2018). The high variability of selected model hyperparameters confirms the need to use optimization methods, such as Grid Search, to ensure

predictive stability and minimise generalization errors for each model individually.

Due to their high accuracy, the models obtained are suitable for determining the values of objective functions in optimization models, e.g. in genetic algorithms. The findings of this study suggest significant disparities between XR and AIXR nozzles with respect to total coverage efficiency. The average total coverage values for XR nozzles were approximately 34–35%, whereas for AIXR nozzles they were around 26–28%. These differences are attributed to variations in droplet size and their behavior within the air stream. XR nozzles produce finer droplets, which are characterised by better deposition on leaf surfaces, thereby increasing coverage, but at the same time increasing the risk of drift (Subr et al. 2023). The optimization results show that total coverage was higher for XR nozzles compared to AIXR nozzles, confirming the importance of droplet size for spraying efficacy (Sapkota et al. 2023). The authors showed that an increase in travel speed led to a decrease in liquid cover regardless of nozzle type. They also found that travel speed and spray pressure are strongly interdependent factors. Their results indicated that higher pressures for XR nozzles resulted in finer atomization and improved coverage, whereas for AIXR no significant improvement was observed with increased pressure. These findings align with the conclusion that spraying efficacy depends not only on nominal travel speed but primarily on maintaining stable flow rate parameters and the appropriate selection of nozzles for operating conditions. It was indicated that under field conditions, the use of XR nozzles allows for achieving higher and more uniform deposition without the need for dynamic pressure adjustment, while for AIXR nozzles application efficacy is more strongly dependent on the interaction between pressure and air velocity. However, it should be noted that the interpretability of the research is significantly limited due to the use of frames with water-sensitive papers. Such artificial plants differ from real plants in terms of leaf surface structure, curvature and ability to move.

Optimization demonstrated that, for both nozzle types, the maximum operating speed ($5.5 \text{ km}\cdot\text{h}^{-1}$) was the most advantageous, resulting from maintaining a stable spray angle and improving spray uniformity under moderate airflow conditions beneath the sprayer boom. Similar results were obtained by Wang et al. (2023), who emphasised that this difference did not affect biological efficacy in weed control in maize and wheat crops. However, it is of great importance for environmental protection.

The integration of precise ML models with optimization algorithms, such as the genetic algorithm, facilitates the identification of the optimal values for specific operating parameters of spraying. Within the optimization algorithm, it is possible to specify the values of certain parameters, such as weather conditions or nozzle type, and thereby determine

the values of the remaining input parameters of the model. This approach provides farmers with accurate guidelines to reduce the use of plant protection products and improve treatment effectiveness. However, it is important to note that this approach is subject to certain limitations. Firstly, it should be noted that the system has been adapted to specific training data, and as a result, it is not universally applicable. Its utilization is contingent upon the conditions that were reflected in the training data. For instance, it is not possible to obtain good quality results in weather conditions that exceed the range of experimental data.

To enhance the versatility of the system, it is imperative to expand the training set with results obtained in a range of weather conditions. Ideally, measurements should be taken across the full range of temperatures and wind speeds at which spraying is permitted by law. Furthermore, it is recommended that measurements are taken using a range of other nozzle types including dual-stream nozzles – both standard and air-induction types – which, according to recent reports, may offer a compromise between coverage accuracy and drift reduction.

Another limitation is the fact that the GA algorithm identified optimal parameters—such as a pressure of 0.22 MPa or 0.25 MPa—that fall between the discrete levels used during the experimental phase. It must be emphasised that the ML models effectively created a continuous response surface by interpolating between the measured data points. While the high predictive performance of the models (low MAPE and high R^2) suggests that the fundamental relationships between operating parameters and spray coverage were successfully captured, these optimised settings represent a theoretical space rather than a directly measured one.

5 Conclusions

The research findings suggest that the type of nozzle significantly affects the efficiency of spray coverage on artificial plant models, with standard (XR) nozzles providing higher and more uniform coverage compared to air-induction (AIXR) nozzles under the tested conditions. These differences mainly result from variations in droplet size and their behavior within the airflow, which are crucial for application efficacy and minimizing drift risk. Machine learning methods allowed for the development of predictive models that delineate potential relationships between operating parameters and treatment quality within the specific range of weather conditions analysed. While the integration of these models into a genetic algorithm helped identify optimal operating parameters for the experimental setup, these results should be interpreted with caution. The finding that a travel speed of $5.5 \text{ km}\cdot\text{h}^{-1}$ was more favorable for both nozzle types is specific to the tested surrogate plants and

controlled environment. Due to the use of artificial plant models, which may not fully replicate the complex architecture and biological variability of real blueberry bushes, and the limited range of meteorological data, these findings cannot be directly implemented on commercial plantations without further validation. Future research is recommended to include field trials on live plants and a broader spectrum of weather conditions, as well as the evaluation of twin-stream nozzles to further balance coverage accuracy and drift reduction.

Author contribution Conceptualization, K.P. and B.C.; methodology, B.C., P.K. and K.P.; software, K.P.; validation, B.C. and P.K.; formal analysis, K.P.; investigation, B.C., P.K. and K.P.; resources, P.K.; data curation, K.P.; writing-original draft, B.C. and K.P.; writing-review and editing, P.K.; visualization, B.C. and K.P. All authors reviewed the manuscript.

Data availability The raw data supporting the conclusions of this article will be made available by the authors upon request.

Declarations

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bock CH, Hotchkiss MW (2021) Effect of tractor speed and spray application volume on spray coverage at different heights in the canopy of tall pecan trees. *Plant Dis* 105(9):2509–2520. <https://doi.org/10.1094/PDIS-11-20-2420-RE>
- Bolat A, Özlüoymak Ö (2020) Evaluation of performances of different types of spray nozzles in site-specific pesticide spraying. *Semin Ciências Agrárias* 41:1199–1212. <https://doi.org/10.5433/1679-0359.2020v41n4p1199>
- Braun MA, Seijo S, Echanobe J, Shukla PK, del Campo I, Garcia-Sedano J, Schmeck H (2016) A neuro-genetic approach for modeling and optimizing a complex cogeneration process. *Appl Soft Comput* 48:347–358. <https://doi.org/10.1016/J.ASOC.2016.07.026>
- Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. In: *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, 785–794
- Cherkassky V, Ma Y (2004) Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Netw* 17(1):113–126. [https://doi.org/10.1016/S0893-6080\(03\)00169-2](https://doi.org/10.1016/S0893-6080(03)00169-2)
- Cieniawska B, Pentoś K (2021) Average degree of coverage and coverage unevenness coefficient as parameters for spraying quality assessment. *Agriculture-Basel* 11(2):151. <https://doi.org/10.3390/agriculture11020151>
- Dekeyser D, Foqué D, Duga AT, Verboven P, Hendrickx N, Nuytens D (2014) Spray deposition assessment using different application techniques in artificial orchard trees. *Crop Prot* 64:187–197. <https://doi.org/10.1016/J.CROPRO.2014.06.008>
- Dong X, Dong L, Gao Z, Wang K, Wang X, Wang S, Qiu B, Wang X (2025) Droplet deposition behavior on the surface of flexible pepper leaves. *Agronomy* 15(3):708. <https://doi.org/10.3390/AGRONOMY15030708>
- Durairaj D, Jyoti B (2023) A method to optimize the pesticide dose considering the combined influence of plant, pest, pesticide, and spray equipment on bioefficacy. *Span J Agric Res* 21(3):e0208–e0208. <https://doi.org/10.5424/SJAR/2023213-19738>
- Façal BS, Freitas H, Gomes PH, Mano LY, Pessin G, de Carvalho A, Krishnamachari B, Ueyama J (2017) An adaptive approach for UAV-based pesticide spraying in dynamic environments. *Comput Electron Agric* 138:210–223. <https://doi.org/10.1016/J.COMPAG.2017.04.011>
- Ferguson JC, Hewitt AJ, O'Donnell CC, Kruger GR (2020) Comparison of water-sensitive paper, Kromekote and Mylar collectors for droplet deposition with a visible fluorescent dye solution. *J Plant Prot Res* 60(1):98–105. <https://doi.org/10.24425/JPPR.2020.132210>
- Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat* 29(5):1189–1232
- Grayson BT, Pack SE, Edwards D, Webb JD (1993) Assessment of a mathematical model to predict spray deposition under laboratory track spraying conditions. II: Examination with further plant species and diluted formulations. *Pestic Sci* 37(2):133–140. <https://doi.org/10.1002/PS.2780370205>
- Hastie T, Tibshirani R, Friedman J (2009) The elements of statistical learning. Springer Ser Stat. <https://doi.org/10.1007/978-0-387-84858-7>
- Immanuel SD, Chakraborty UK (2019) Genetic algorithm: an approach on optimization. In: *Proceedings of the 4th International Conference on Communication and Electronics Systems, ICCES 2019*, 701–708. <https://doi.org/10.1109/ICCES45898.2019.9002372>
- Lambora A, Gupta K, Chopra K (2019) Genetic algorithm- a literature review. In: *Proceedings of the international conference on machine learning, big data, cloud and parallel computing: trends, perspectives and prospects, COMITCon 2019*, 380–384. <https://doi.org/10.1109/COMITCON.2019.8862255>
- Legates DR, McCabe GJ (1999) Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour Res* 35(1):233–241. <https://doi.org/10.1029/1998WR900018>
- Lewis CD (1982) *Industrial and business forecasting methods : a practical guide to exponential smoothing and curve fitting*. Butterworth Scientific.
- Londhe SB, Sujata K (2017) Remotely operated pesticide sprayer robot in agricultural field. *Int J Comput Appl* 167(3):975–8887
- Moriassi DN, Arnold JG, Liew MW, Van Bingner RL, Harmel RD, Veith TL (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans ASABE* 50(3):885–900. <https://doi.org/10.13031/2013.23153>
- Pedregosa F, Varoquaux G, Gramfort A, Vincent M, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
- Peng H, Li Z, Zhou Z, Shao Y (2022) Weed detection in paddy field using an improved RetinaNet network. *Comput Electron Agric* 199:107179. <https://doi.org/10.1016/J.COMPAG.2022.107179>

- Petelewicz P, Zhou Q, Schiavon M, MacDonald GE, Schumann AW, Boyd NS (2024) Simulation-based nozzle density optimization for maximized efficacy of a machine vision-based weed control system for applications in turfgrass settings. *Weed Technol* 38:e25. <https://doi.org/10.1017/WET.2024.7>
- Probst P, Boulesteix AL, Bischl B (2018) Tunability: importance of hyperparameters of machine learning algorithms. *J Machine Learning Res* 20. <https://arxiv.org/pdf/1802.09596>
- Renaudo CA, Bertin DE, Bucalá V (2022) A coupled atomization-spray drift model as online support tool for boom spray applications. *Precis Agric* 23(6):2345–2371. <https://doi.org/10.1007/S11119-022-09923-1>
- Sapkota M, Virk S, Rains G (2023) Spray deposition and quality assessment at varying ground speeds for an agricultural sprayer with and without a rate controller. *AgriEngineering* 5(1):506–519. <https://doi.org/10.3390/AGRIENGINEERING5010033>
- Sawant CP, Jyoti B, Gaikwad BB, Gaikwad N, Kumar N, Kumar M (2023) Optimization of operational parameters of site-specific pesticide spray module for young Pomegranate orchards using RSM and RBFNN-PSO techniques. *J Biosyst Eng* 48:225–241. <https://doi.org/10.1007/s42853-023-00185-x>
- Sayinci B, Demir B, Açık N, Sayinci B, Demir B, Açık N (2020) Comparison of spray nozzles in terms of spray coverage and drop distribution uniformity at low volume. *Turk J Agric for* 44(3):262–270. <https://doi.org/10.3906/tar-1905-112>
- Singh E, Pratap A, Mehta U, Azid SI (2024) Smart agriculture drone for crop spraying using image-processing and machine learning techniques: experimental validation. *IoT* 5(2):250–270. <https://doi.org/10.3390/IOT5020013>
- Singh J, Knapp HV, Demissie M (2004) Hydrologic modeling of the iroquois river watershed using HSPF and SWAT. ISWS CR 2004–08. Champaign, Ill.: Illinois State Water Survey. Available at: www.sws.uiuc.edu/pubdoc/CR/ISWSCR2004-08.pdf
- Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14(3):199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
- Subr A, Al-Ahmadi A, Abbas M (2020) Effect of nozzle type and some locally used surfactants on the spray quality. *Iraqi J Agric Sci* 51(3):856–864. <https://doi.org/10.36103/IJAS.V51I3.1040>
- Subr A, Milanowski M, Parafiniuk S (2023) Effect of nozzle condition and type in the spray quality in wheat plants. *Lecture Notes Civil Eng* 289:131–137. https://doi.org/10.1007/978-3-031-13090-8_14
- Tewari VK, Chandel AK, Nare B, Kumar S (2018) Sonar sensing predicated automatic spraying technology for orchards. *Curr Sci* 115(6):1115–1123
- Wang D, Xu S, Li Z, Cao W (2022) Analysis of the influence of parameters of a spraying system designed for UAV application on the spraying quality based on Box–Behnken response surface method. *Agriculture* 12(2):131. <https://doi.org/10.3390/AGRICULTURE12020131>
- Wang S, Li X, Nuytens D, Zhang L, Liu Y, Li X (2023) Evaluation of compact air-induction flat fan nozzles for herbicide applications: spray drift and biological efficacy. *Front Plant Sci* 14:1018626. <https://doi.org/10.3389/FPLS.2023.1018626>
- Wang S, Wang W, Xu D, Nuytens D, Xu G, Lu D, Li X (2025) Experimental study on the impact outcomes of droplets on rice foliage: influence of liquid properties and leaf angles. *Pest Manag Sci* 81(12):8064–8076. <https://doi.org/10.1002/PS.70114>
- Xiahou B, Sun D, Song S, Xue X, Dai Q (2020) Simulation and experimental research on droplet flow characteristics and deposition in airflow field. *Int J Agric Biol Eng* 13:16–24
- Xie Q, Song M, Wen T, Cao W, Zhu Y, Ni J (2024) An intelligent spraying system for weeds in wheat fields based on a dynamic model of droplets impacting wheat leaves. *Front Plant Sci* 15:1420649. <https://doi.org/10.3389/FPLS.2024.1420649>
- Xue X, Zeng K, Li N, Luo Q, Ji Y, Li Z, Lyu S, Song S (2023) Parameters optimization and performance evaluation model of air-assisted electrostatic sprayer for citrus orchards. *Agriculture* 13(8):1498. <https://doi.org/10.3390/AGRICULTURE13081498>
- Yang T, Zheng X, Vidyarthi SK, Xiao H, Yao X, Li Y, Zang Y, Zhang J (2023) Artificial neural network modeling and genetic algorithm multiobjective optimization of process of drying-assisted walnut breaking. *Foods* 12(9):1897. <https://doi.org/10.3390/FOODS12091897>
- Yongjun Z, Shenghui Y, Chunjiang Z, Liping C, Lan Y, Yu T (2017) Modelling operation parameters of UAV on spray effects at different growth stages of corns. *Int J Agric Biol Eng* 10(3):57–66

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.