Machine Learning model to predict total nozzle volume delivery for Pulse Width Modulated flow controllers

[1]Dua S, [1]Sharda A

[1]Carl and Melinda Helwig Department of Biological and Agricultural Engineering, Kansas State University, USA

A paper from the Proceedings of the
16th International Conference on Precision Agriculture
21-24 July 2024
Manhattan, Kansas, United States

Abstract.

Product flow rate in the Pulse Width Modulation (PWM) variable rate technologies depends on the duty cycle. However, the actual product flow rate at any duty cycle depends on pressure rise, stable pressure during the cycle, fall time and pressure drop across the nozzle body. The current controller does not consider the pressure drops and the estimation of actual flow during each cycle at any duty cycle cannot be estimated with capturing high-frequency pressure data. Knowledge of volume delivery during different duty cycles can provide valuable information on product delivery on a nozzle-by nozzle basis.

Therefore, this research aims to estimate total spray volume in agricultural spraying operations, particularly when utilizing PWM system. For optimizing spray volume, the Raven PWM spray system was tied. One nozzle was instrumented with a high frequency pressure transducer to collect data at 1000 Hz. The pressure data was collected when running the spray system at two pressures (275.8 kPa and 448.1 kPa) and at two application rates (112.2 L/ha and 187.1 L/ha). One critical parameter is quantifying actual nozzle volume delivery for a particular PWM system. Different machine learning algorithms such as regression, random forest, XGBoosting were used. Linear regression models provide insights about the linear relationships between the independent variables and the total spray volume. Random forest algorithms offered robustness and interpretability that enable us to discern feature importance and understand the factors contributing most to spray volume variations. XGBoosting was a powerful gradient boosting technique that allows for capturing complex interactions and patterns within the data. By comparing these algorithms, Random Forest algorithm provided the most robust volume estimation.

The results indicated a substantial influence of selected predictor variables on the response variable, total flow providing valuable insights into the precise determination of total volume (>98% correlation) sprayed in the field. By incorporating PWM technology and accounting for constant *variables, our model enhanced resource management and could potentially contribute to easier integration of critical information needed to manage flow rates while contributing to environmental sustainability.*

Implementing this data-driven precision agriculture approach can revolutionize crop protection practices, leading to improved agricultural productivity and reduced environmental impact. It empowers farmers and agricultural professionals with the tools to make informed decisions, ensuring that agrochemicals are applied efficiently and effectively to maximize crop yield and environmental stewardship.

Keywords.
Machine learning, Pulse Width Modulated System, total flow, spraying system.

Introduction

Liquid application is one of the crucial components in agricultural system. Spraying system has seen tremendous transition from broadcasting agrochemicals to use of giant sprayers in field. This continuous development has led to increase crop production, pest and weed management. However, uniform application of agro-chemicals is still a biggest challenge faced by farmers. Over and under use of inputs have increased production cost and cause potential harm to human health and environment. In order to address this issue, precision technologies are being used to improve crop productivity, sustain agricultural practices and optimize resource utilization. Precision technologies in sprayers have become more efficient in its application due to technological advancements in computers, sensors and actuators. Employing these sensor technologies in product application potentially improve crop quality and yield (Fabula et al., 2021; Sharda et al., 2013; Popp et al.,2013). Spraying system utilizes Variable rate technologies (VRT) such as Pulse Width Modulation System which adjust application rates based on specific needs of the different areas within field.

Pulse Width Modulated (PWM) system is the most advanced technology which is implemented in self propelled sprayers. PWM system controls the nozzle flow rate by pulsing an electronically-actuated solenoid valve directly upstream of nozzle by changing duty cycle (Grella et al., 2021). Duty cycle and frequency are the major components of PWM system. Duty cycle is the percentage of total time the signal is in the ON state to complete one cycle and frequency is number of cycles completed per second. PWM control flow nozzle by nozzle basis by opting right duty cycle according to each nozzle' speed and target application rate during parallel and curvilinear passes (Fabula et al., 2021). However, the actual product flow rate at any duty cycle depends on pressure rise, stable pressure during the duty cycle, fall time and pressure drop across the nozzle body. The current controllers do not consider the pressure drops and estimation of actual flow during each cycle at any duty cycle cannot be estimated with capturing high frequency pressure data. It is important to understand the pressure dynamics which provide real time flow changes based on selected target application pressure. Knowledge of volume delivery during different duty cycles can provide valuable information on product delivery on a nozzle-by-nozzle basis.

Therefore, this study aims to estimate the total volume of spray during each cycle in agricultural spraying operations while using PWM system. In this study first total flow per cycle is calculated by considering pressure drop mechanics at varying duty cycle then that calculated data was used to train the model such that we can get total flow per cycle at any sprayers settings (duty cycle, pressure, frequency, etc). For this particular study, Raven Hawkeye PWM spray system was tied. One nozzle was instrumented with high frequency pressure transducer to collect data at 1000 hz. The pressure data was collected when running the spray system at two pressures (275.8 kPa and 448.1 kPa) and at two application rates (112.2 L/ha and 187.1 L/ha). The main objective of this study is to predict the total flow per cycle using different sprayer settings (flow rate, Duty cycle, pressure, frequency) and the impact of these sprayer settings on total flow per cycle.

Material and Methods

This section consists two elements of the study that includes data acquision and data analysis using Machine learning Algorithms.

Data acquision

Raven Hawkeye (Raven Industries, Inc, Sioux Falls, SD) PWM control system was used in data collection for this study. A set of five Teejet nozzle bodies (Teejet Technologies, Springfield,I1) and Raven nozzle control valves were used PWM system (Figure 1). The nozzle selection was based on recommendation for PWM system which includes two flat fan nozzles (Pentair Hypro, Minneapolis, MN). Nozzle was instrumented with high frequency pressure transducer to collect the data at 1000 hz. The pressure data was collected when spray system was running at application rates of 112.2 liters per hectare and 187.1 liter per hectare based on two pressures of 275.8 kPa and 448.1 kPa. The main idea was to estimate total flow at each cycle considering the pressure drop.

Figure 1. TeeJet nozzle body and Raven nozzle control valve

The Raven Viper 4 rate controller had the target application rate dialed in. The regulating valve at the spray system's rear was controlled by a switch that changed the pressure inside the boom. In order to achieve the required pressure of 275.8 or 448.1 kPa, the boom pressure was adjusted during every test prior to data collection. Different procedures were performed to vary duty cycle and system frequency for PWM nozzle control system. The Raven Viper 4 display monitor was used to vary the duty cycle and system frequency. The set up and instrumentation using PWM system for this study is shown in Figure2. The frequency recommended by manufacturer commercially is 10 hz but the PWM system used in this study had functionality to vary the frequency to 10 Hz, 15 Hz and 30 Hz. The pressure and flow rate data were recorded for 30 seconds, which provided 30,000 data points for data analysis.

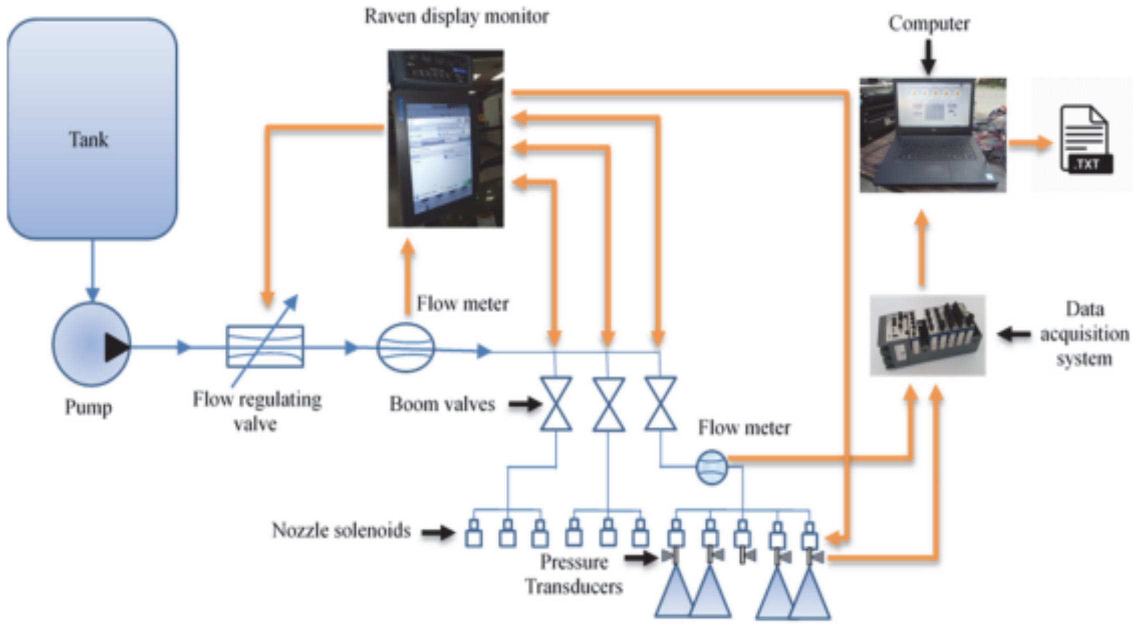


Figure 2. Setup and instrumentation for system

Data analyzation

After data processing, we had our predictor variables flow rate, duty cycle, pressure and frequency and we predicted total flow per cycle using various machine learning algorithms such as Linear Regression, Decsision trees, Random Forest and XGBoosting .

Linear Regression

In this study first linear regression model was used to analyze the relationship between the total flow per cycle and other variables. After the analyzing the relationship the prediction of target variable (total flow per cycle) is being done using linear modeling. Linear regression is a statistical analysis that determine the quantify the relationship between target and input variables. The linear model can be written in the format below:

$Y = \beta 0 + \beta i. Xi + \varepsilon$, i=1….,n

Where Y is target variable, Xi are input variables, β0 is the intercept parameter, that is the value of target variable when predictor is 0. βi is the estimate or slope parameter which is magnitude of change in target variable given one unit change in input variable. ε is the error term that represents the deviation of y from actual value. Unknown parameters β0 and βi are obtained by the method of ordinary least squares. Basically, it uses the methodology of minimizing the sum of squared vertical distances between the observation and fitted line. In linear regression, the target variable is formed from combinations of slope and estimates.

The linear model is simple to use and easily interpreted. It is quite simple to implement using equation. The coefficient interpretation is such that effect on target given a unit change in predictor control for other variables. Here predictions can easily obtain on new data.

Decision tree

Decision trees are supervised machine learning technique. It is basically a predictive modelling technique that takes decision based leaf and nodes. This model is represented by tree like structure It comprises of some components listed below:

- An internal node is a test on an attribute
- A branch represents an outcome of the test
- A leaf node represents a class label or class label distribution.
- At each node, one attribute is chosen to split the training data onto distinct classes as much as possible
- A new instance is classified by following a matching path to a leaf node

Random Forest

It is a supervised machine learning technique that also gives the continous value as the output. It makes no assumptions about relationships between the inputs and output. As of the results from random forest come from some models of decision tree, the three main parts to the random forest are node size, number of trees and number of features sampled. It is usable on both regression as well as the classification problems.

Random forest is based on collection of decision trees. It is an ensemble learning algorithm that is is used in predicting continuous or categorical response variable. Bootstrap aggregating or bagging is a technique which is commonly used by random forest. Samples drawn from the dataset are with replacement in order to create the training set. Thus, some data points may appear multiple times in one sample while other may not appear at all. Decision tree is built for individual training set. For classification tasks, each tree in the forest votes for a class, and the class with the most votes becomes the model's prediction. For regression tasks, the average of all the tree outputs is taken as the final prediction. It gives more accuracy than single decision tree as it combines the results of multiple trees. Random forest is most robust on unseen data due to its randomness. Feature selection and data sampling by randomness help in reduction of overfitting.

XG Boosting

XGBoost is also a ensemble machine learning algorithm. Basically, decision trees are being used in this algorithm as a base learner. It employs regularization techniques in order to enhance model generalization. Boosting is widely known for its computational efficiency, feature importance analysis and handling of missing values, XG boost is widely used for task such as regression, classification and ranking. It is a predictive modeling technique that inculcate the predictions of multiple individual decision tree models in a iterative manner. It adds weak learners to the ensemble with a each learner focusing on correcting the errors made by existing ones. Its uses a gradient descent optimization technique to minimize a predefined loss function during training. The most important features of this algorithm are its ability to handle complex relationships in a data, regularization techniques to prevent overfitting and incorporation of parallel processing for efficient computation. XG boost is widely used due to its high predictive accuracy and versatility across different dataset.

Results and Discussion

This section of this paper consists of some of findings which comes outs of integration of various machine learning models (Regression, Decision tree, Random Forest, XGBoosting) to estimate the total flow per cycle based on various sprayer settings. The predictions made by these models were then compared using various evaluation metrics such as Mean Absolute Error (MAE) , Mean Absolute Percentage error (MAPE) and Akaike Information Criterion (AIC).

Based on dataset collected from Raven Hawkeye system, visuals from the dataset is such that there is increase in total volume per cycle when we tend to increase pressure, duty cycle and flowrate. On the other hand, there is decline in total flow per cycle in case of frequency (Figure 3). Also, with increase in duty cycle, change in total flow per cycle is significant more when we increase our pressure setting (Figure 4).

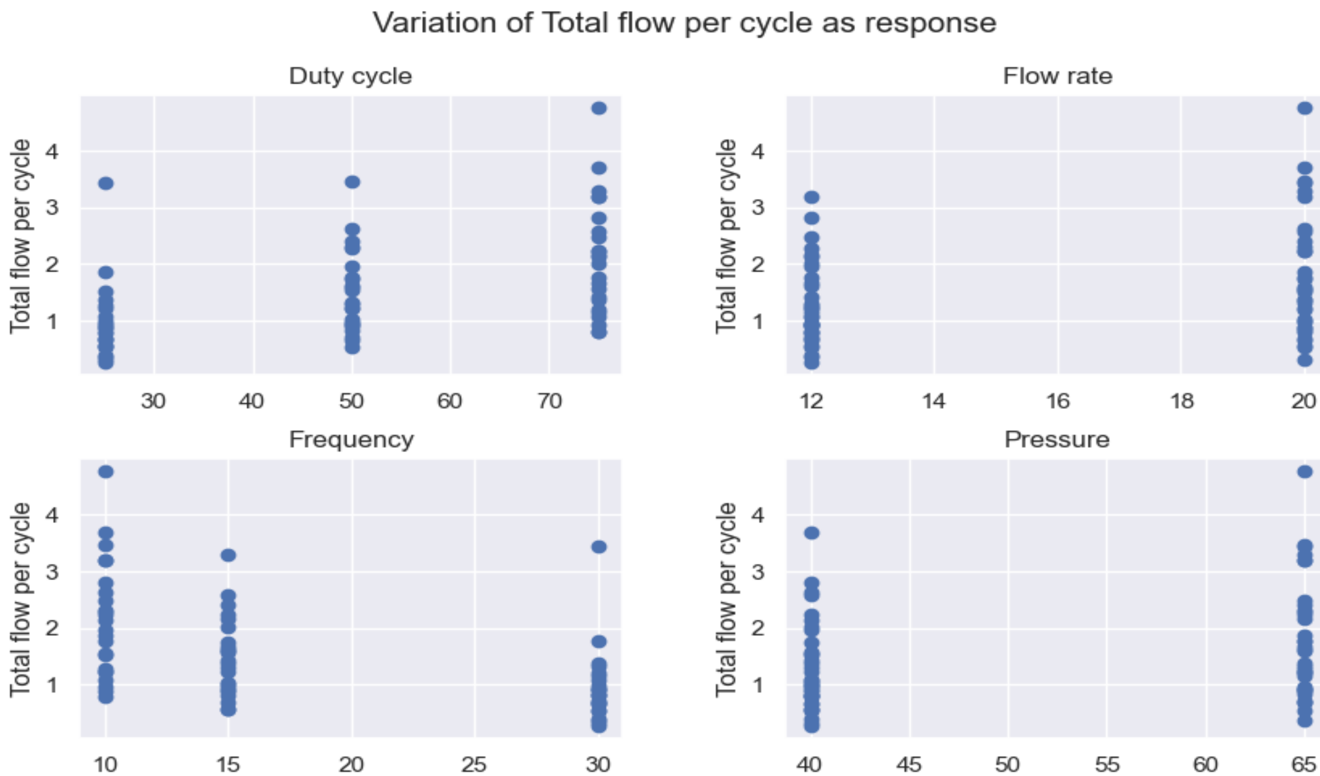


Figure 3. Variation of total per cycle as response

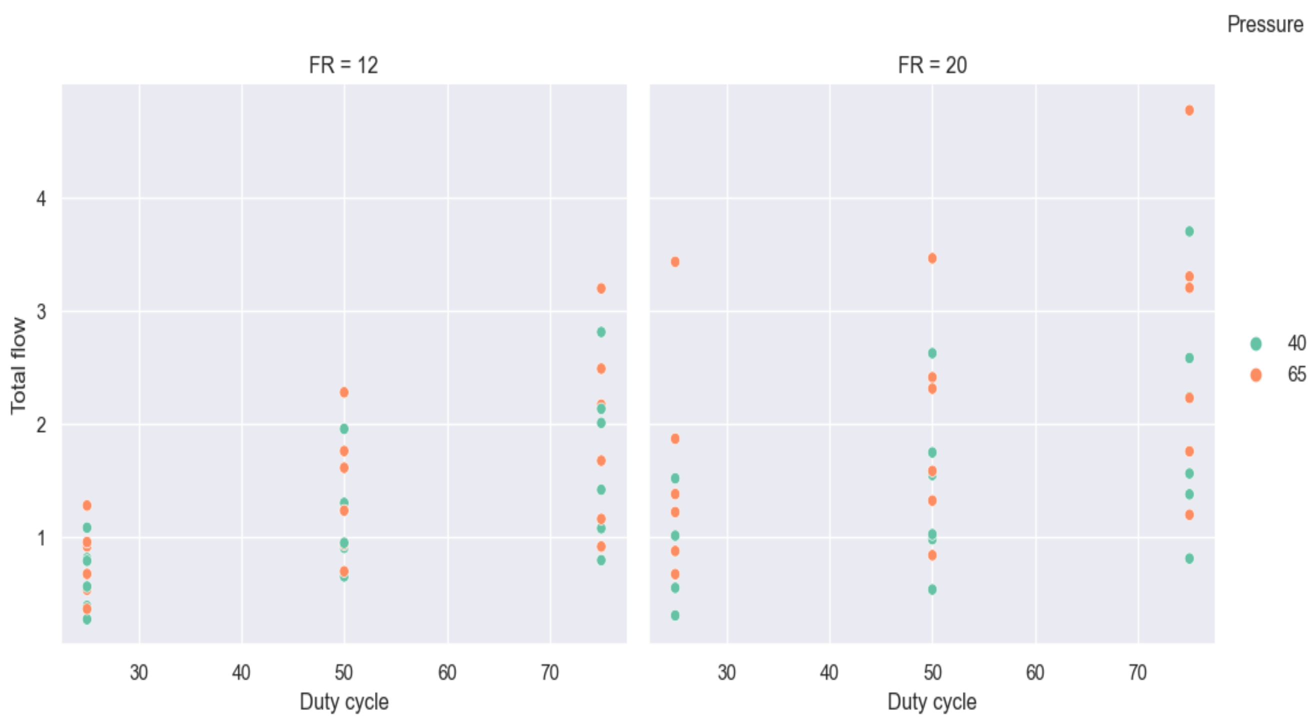


Figure 4. Variation in total flow with duty cycle and flow rate

Linear model with total flow per cycle as a response variable was fitted against the predictors (duty cycle, flow rate, pressure and The results of linear regression model was such that the model had adjusted R square value of 0.884. and the p values of duty cycle, flow rate, pressure and frequency were 0.000, 0.003, 0.016, and 0.000 respectively (<0.05= α) which means these settings are significantly affecting the our response variable. The fitted linear regression plots are shown below (Figure 5).

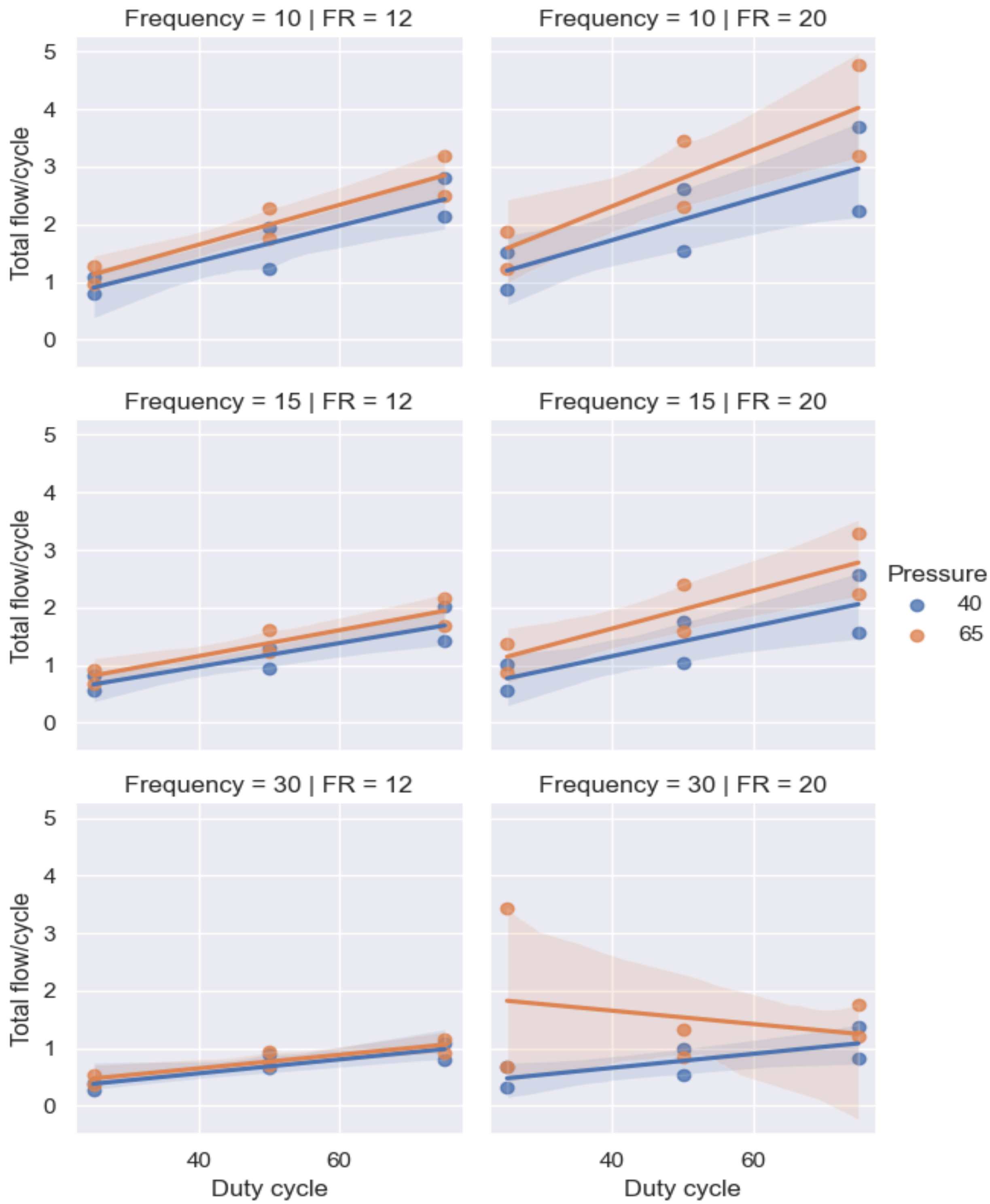
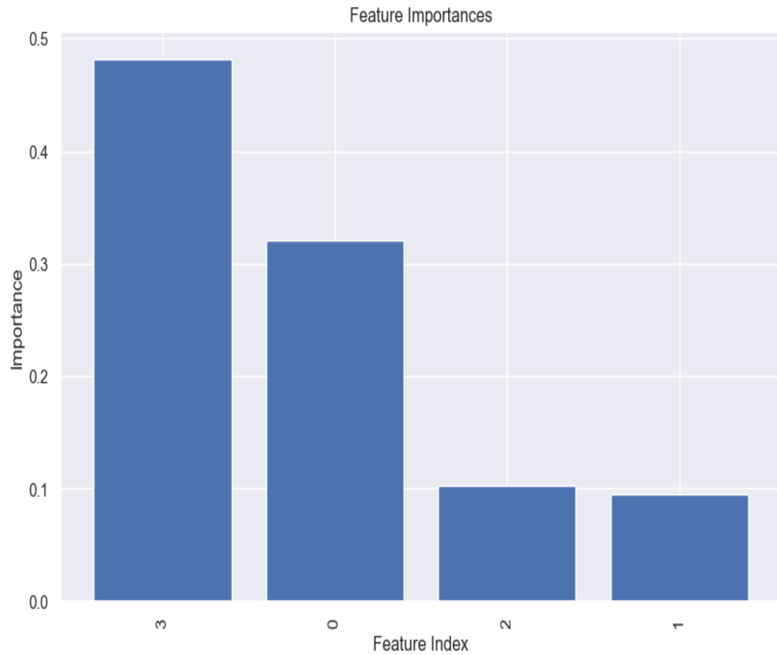


Figure 5. Fitted regression plots

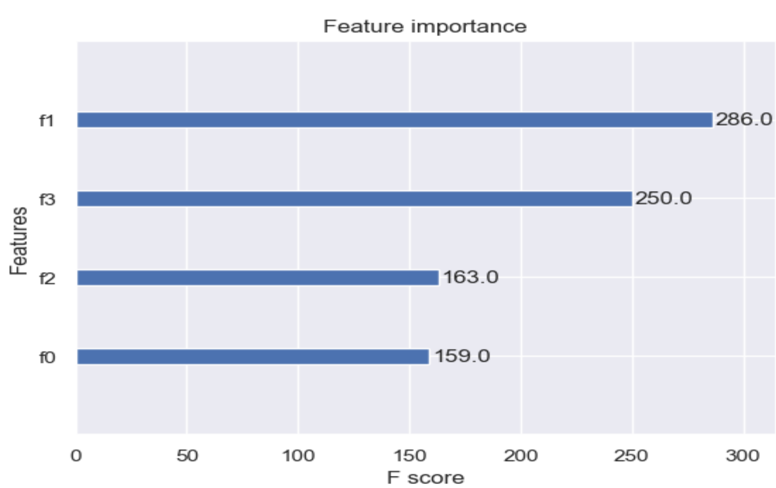
The decision tree regressor was used to fit the same predictors and response variable and the adjusted R square value for decision tree comes out to be 0.8007. For the improvement in modelling random forest as it implies the results of multiple decision trees; R quare value for this model comes out to be 0.804. it gives the feature importance of the fitted model.according to random forest results, frequency is the most influencing factor in predicting the total flow per cycle followed by duty cycle, flow rate and frequency respectively (Figure6).



Feature 0 – Duty Cycle
Feature 1 – Pressure
Feature 2 – Flow rate
Feature 3 – frequency

Figure 7. Feature importance (Random forest)

XGboosting results indicated that Pressure was most influencing factors in predicting followed by frequency, duty cycle and flow rate. However, these are the contracting results from the above models used. So in order to predict total flow per cycle using correct model evaluation metrics for model comparison was done so that we can check most appropriate model for prediction and checking the influence of predictors on total flow per cycle.



Feature 0 – Flow rate
Feature 1 – Pressure
Feature 2 – Duty cycle
Feature 3 – frequency

Figure 8, Feature importance (XG Boosting)

Predictions of total flow per cycle was done using all models (Linear regression, Decision tree, Random forest and XGBoosting but final selection for the prediction was done using performance metrics such as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), Akaike Information Criterion (AIC). Akaike Information Criterion is a mathematical method for checking how well a model fits the data. Basically, it is used to compare possible models and determine which model is best fit for the data. Mean Absolute Error is a statistical metric that is used for evaluating performance of various machine learning algorithm. It is a measure of errors between paired observations expressing the same values. Mean Absolute Percentage error: it is a measure of absolute percentage error and used to measure how accurate a forecast system is. The table below shows various model performance based on these metrics.

Model	**R square**	**MSE**	**MAPE**	**AIC**
Linear Regression	0.884	0.239	36.584	-21.59
Random forest	0.804	0.220	32.626	-12.70
Decision tree	0.8007	0.253	34.474	-10.58
XGboosting	0.807	0.260	33.848	-10.17

Based on performance metrics shown in table Random Forest is most accurate for prediction of total flow per cycle. However, there are numerous performance metrics which can be used in future studies to improve model performance and hence increase the accuracy of prediction model. Using this trained model a dash bord has been made using streamlitt app for prediction of total flow per cycle. Prediction can be done on unseen input based already trained data.

References

Alam, M., Alam, M. S., Roman, M., Tufail, M., Khan, M. U., & Khan, M. T. (2020). Real-time machine-learning based crop/weed detection and classification for variable-rate spraying in precision agriculture. *2020 7th International Conference on Electrical and Electronics Engineering (ICEEE)*, 273–280. https://doi.org/10.1109/ICEEE49618.2020.9102505

Butts, T. R., Luck, J. D., Fritz, B. K., Hoffmann, W. C., & Kruger, G. R. (2019). Evaluation of spray pattern uniformity using three unique analyses as impacted by nozzle, pressure, and pulse-width modulation duty cycle. *Pest Management Science*, *75*(7), 1875–1886. https://doi.org/10.1002/ps.5352

Fabula, J. (n.d.). Nozzle flow dynamics during control system response of pulse width modulated (PWM) technology- equipped agricultural sprayer. Kansas State University.

Fabula, J., Sharda, A., Kang, Q., & Flippo, D. (2021). Nozzle flow rate, pressure drop, and response time of pulse width modulation (Pwm) nozzle control systems. *Transactions of the ASABE*, *64*(5), 1519–1532. https://doi.org/10.13031/trans.14360

Fabula, J., Sharda, A., Luck, J. D., & Brokesh, E. (2021). Nozzle pressure uniformity and expected droplet size of a pulse width modulation (Pwm) spray technology. *Computers and Electronics in Agriculture*, *190*, 106388. https://doi.org/10.1016/j.compag.2021.106388

Grella, M., Gioelli, F., Marucco, P., Zwertvaegher, I., Mozzanini, E., Mylonas, N., Nuyttens, D., & Balsari, P. (2022). Field assessment of a pulse width modulation (Pwm) spray system applying different spray volumes: Duty cycle and forward speed effects on vines spray coverage. *Precision Agriculture*, *23*(1), 219–252. https://doi.org/10.1007/s11119-021-09835-6

Hemant Kumar Singh, Bhanu Pratap, S. K. Maheshwari, Ayushi Gupta, Anuradha Chug, Amit Prakash Singh, & Dinesh Singh. (2023). Spray prediction model for aonla rust disease using machine learning techniques. *Journal of Agricultural Science and Technology B*, *13*(1). https://doi.org/10.17265/2161-6264/2023.01.001

Indu, Baghel, A. S., Bhardwaj, A., & Ibrahim, W. (2022). Optimization of pesticides spray on crops in agriculture using machine learning. *Computational Intelligence and Neuroscience*, *2022*, 1–10. https://doi.org/10.1155/2022/9408535

Jiang, H., Zhang, L., & Shi, W. (2016). Effects of operating parameters for dynamic pwm variable spray system on spray distribution uniformity. *IFAC-PapersOnLine*, *49*(16), 216–220. https://doi.org/10.1016/j.ifacol.2016.10.040

P. A. Larbi & M. Salyani. (2011). Model to predict spray deposition in citrus airblast sprayer applications: Part 1. Spray dispersion. *Transactions of the ASABE*, *55*(1), 29–39. https://doi.org/10.13031/2013.41245

S. GopalaPillai, L. Tian, & J. Zheng. (1999). Evaluation of a flow control system for site-specific herbicide applications. *Transactions of the ASAE*, *42*(4), 863–870. https://doi.org/10.13031/2013.13265

Sharda, A., J.P. Fulton, T.P. McDonald, and C.J. Brodbeck. 2011. Real-time nozzle flow uniformity when using automatic section control on agricultural sprayers. Computers and Electronics in Agriculture. 79(2011): 169-179.

Sharma, A., Jain, A., Gupta, P., & Chowdary, V. (2021). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, *9*, 4843–4873. https://doi.org/10.1109/ACCESS.2020.3048415