MJSIR
Mountain Journal of Science and Interdisciplinary Research
PRINT ISSN: 2619-7855
ONLINE ISSN: 2651-7744
January-June 2025 • 85 (1) :16-37

Geographical Variation, Path Diagram, and Regression Tree of the Incidence and Severity of Potato Late Blight (*Phytophthora infestans*): The Observed Pattern in the Major Growing Areas in Benguet Province

Han B. Bonoy* and Precious M. Valentin

College of Numeracy and Applied Sciences, Benguet State University
**Corresponding author email address: han.b.bonoy@gmail.com*

ARTICLE INFO

Date Received: 09-05-2022
Date Last Revised: 05-26-2025
Date Accepted: 06-19-2025

KEYWORDS

cluster analysis
path analysis
CaRT model
disease tetrahedron
potato late blight

Abstract

There are various multivariate statistical tools for plant disease epidemics to associate multiple epidemiological factors represented by the disease triangle and tetrahedron. However, its applications are limited in Asian countries, let alone in the locality. This descriptive causal–comparative [farm] survey research, therefore, utilized Cluster Analysis, Path Analysis, and Classification and Regression Tree (C&RT). These multivariate tools were used to assess the geographical variations, account the explained variance/ effect size, and model the causal relationship of multiple variables on host variables, environmental factors, and cultural management practices of farmers to visually assessed incidence and severity of late blight of potato farms in major growing areas in Benguet observed in May 2021. There are three (3) clusters of the observed farms that emerged, of which cluster 3 has consistently had the lowest incidence and severity of potato late blight observed. Therefore, the host characteristics and cultural management observed in cluster 3 are worthy of consideration in potato production. Both the best/tuned regression trees and the combined recursive path diagram showed host variables to have contributed a significantly high proportion of effect/ explained variance and emerged most of the significant predictors for the observed incidence and severity of potato late blight.

Introduction

Agriculture has always been considered the backbone of food security and a sustainable economy. It is noteworthy that among non-cereal crops, potato (*Solanum tuberosum*) is the most produced in the world, which is fourth after maize, wheat, and rice (Food and Agricultural Organization of Corporate Statistical Database [FAOSTAT], 2021). The Philippines ranked 159th in potato production in the world, 33rd in Asia, and 5th in Southeast Asia (FAOSTAT, 2021). Particularly, the Cordillera Administrative Region (CAR) is the top producer of potatoes in the Philippines, specifically the municipality of Benguet (Philippine Statistics Authority [PSA] OPENSTAT, 2021). It is observed that potato was the top

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

produced among other highland vegetables in terms of production area and harvest, until it gradually declined in the previous years, 2017-2019 (PSA Openstat, 2021).

Along with other factors, the incidence of plant pests and diseases may have affected the decline of potato production in CAR. Pests and diseases remain one of the major detriments in sustainable crop production. Thus, epidemiological studies and modelling on plant pests and diseases for decision-support systems were among the research areas and priorities stipulated in the DOST—Harmonized National Research and Development Agenda (HNRDA) 2017-2022 and the DA—Research, Development and Extension Agenda and Programs (RDEAP) 2016-2022.

It is through epidemiological studies like field trials, experiments, and farm surveillance that one could be able to understand how disease develops in host crop populations and how other factors influence its development. Epidemiological studies quantify and model the parameters that revolves around an epidemic to determine when, where, what plant, and cultural management practices that suppresses the development of plant pathogens (Nutter, 2007; Xu, 2006; Van Maanen & Xu, 2003).

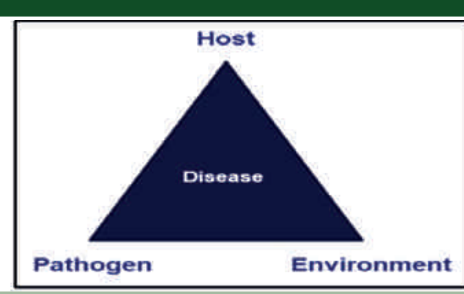
There are many mathematical and statistical models varying in complexity that have been introduced and are already being utilized in plant pathology and epidemiology. For instance, the models that describe the temporal dynamics and spatial patterns of the development of plant diseases (Nutter, 2007; Xu, 2006; Van Maanen & Xu, 2003; Madden & Hughes, 1995) and some nonparametric tests of independence that determine interspecific association of two or more diseases (Turechek, 2004). Additionally, many advanced and multivariate statistical models are being utilized in American and European countries, but limitedly being applied in Asian countries (Nayak et al., 2018), even less so in the locality. Most epidemiological studies on potato late blight in the locality, for instance, are more on experimental trials of control managements, breeding of resistant varieties, and identifying biological controls and cycles of pests and diseases with the utilization of descriptive statistics and basic inferential tests for comparison and correlation. Other than that, nonparametric tests are rarely used and multivariate models are limitedly explored by local plant pathologists. In addition, epidemiological survey data are summarized by mere descriptive statistics, which do not explicitly associate multiple causal factors observed.

Hence, this study embarked on the application of Cluster Analysis, Path Analysis, and Classification and Regression Tree (CaRT) model. Respectively, Path Analysis and CaRT model construct a path diagram and a regression tree, which are user-friendly decision support models to understand the variabilities in the incidence and severity of potato late blight (*Phytophthora infestans*) accounted for or explained by host variables, environmental factors, and cultural management activities and interventions. The path diagram and regression tree are decision support models that have an easy structure of the rule and inherent logic apparent to non-statistically inclined users (Lleras, 2005; Lewis, 2000). With its simplicity, yet captures the broad characteristics of a host-pathosystem, modeling approaches may no longer be a testable proposition but a demonstrated reality that makes a major impact on practical disease management (Jeger, 2004; Cunniffe et al., 2015).

The disease triangle (Figure 1) is the basic concept in plant pathology that represents the fundamental principles involved in disease development from the interaction of a susceptible host, a virulent pathogen, and a favorable environment. Over time, the disease triangle evolved into a three-dimensional disease tetrahedron (Figure 2) where plant pathologists commonly suggested either humans, vectors, or time as another factor (Francl, 2001). In this particular study, human factor was included since human activities like cultivation practices affect

Figure 1

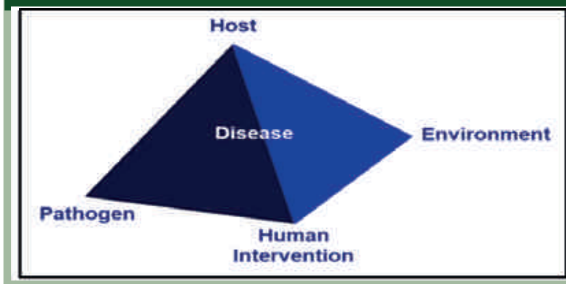
The Disease Triangle



the pathogen's life cycle, and various environmental manipulations affect the incidence and severity of a plant disease.

Figure 2

The Disease Tetrahedron



Anchoring on the principles and concept of the disease triangle and tetrahedron, the paradigm of the study was presented in Figure 3. The variables that were observed for each parameter were identified through a series of consultations with a local plant pathologist and expert on potato late blight. The variables primarily considered for their potential causal effect on the incidence and severity of potato late blight are viable for observational and survey methods of data collection.

Generally, this study conducted a surveillance and assessment of potato farms in Benguet with the utilization of Cluster Analysis, Path Analysis, and Classification and Regression Tree (C&RT) model to account for the variance and model the identified causal variables towards the potato late blight. Specifically, the study sheds light on: (1) What are the clusters of the observed potato farms in the major growing areas in Benguet, and the pattern manifested concerning the incidence and severity of potato late blight, host variables, environmental factors, and cultural management practices?; (2) Which among the observed host variables, environment factors, and cultural management practices have accounted significant effect or explained variance and predictors of the incidence and severity of potato late blight when fitted in the recursive path diagram?; and (3) Which among the observed host variables, environment factors, and cultural management practices are significant predictors of the incidence and severity of potato late blight as determined by the best/tuned regression trees?

Methodology

Research Design

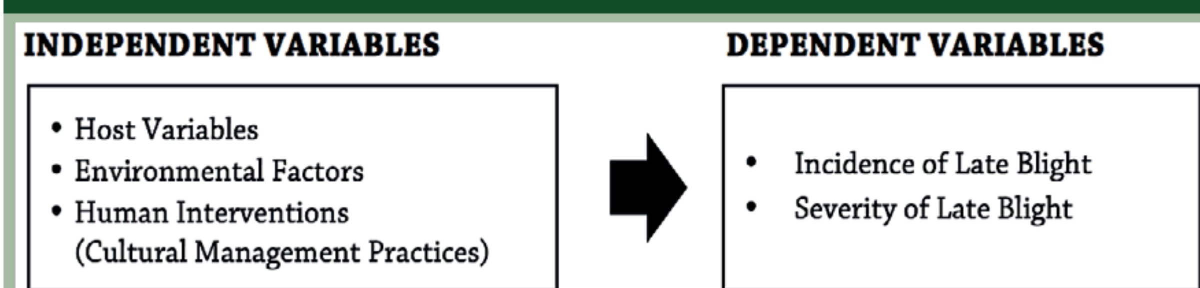
This is a descriptive causal-comparative [farm] survey research study. Primarily, the study assessed geographical variation, accounted for and compared the explained variances, and modeled the causal relationship among the host variables, environmental factors, and cultural management practices to the incidence and severity of potato late blight.

Population and Locale of the Study

The assessment was mainly on the potato farms in the major growing areas in Benguet for being the top producer of potatoes in the Cordillera Administrative Region (PSA Openstat, 2021). Specifically, it is within the municipalities of Buguias, Mankayan, Atok, and Kibungan (Gonzales et al., 2016) (Figure 4). The map was

Figure 3

The Paradigm of the Study



constructed with the geotagged photos of the assessed farms using the DA GeoCamera and uploaded using QGIS Desktop 3.22.9.

There are various rule – of – thumb and simulations conducted on sample size relative to regression modelling (Bentler & Chou, 1987; Nunnally, 1967). Considering the time and resources, 300 potato farms were assessed, considering the 200 minimum independent cases (Boomsma, 1985 & 1982) plus an additional 100 for the cross-validation and independent sample test for the C&RT model.

The assessed farms were selected through a purposive sampling technique with the following inclusion-exclusion criteria:

a. The potato crop must be at least at the vegetative stage because the assessment of the incidence and severity of late blight and host characteristics will be estimated on the leaves and stems;

b. A farmer may own several farm parcels. Hence, a parcel can only be considered as a separate farm if,

i. located in a different area,

ii. adjacent to another owned parcel but planted with a different potato variety,

iii. adjacent parcel planted with the same variety, but it was planted at a different date; and,

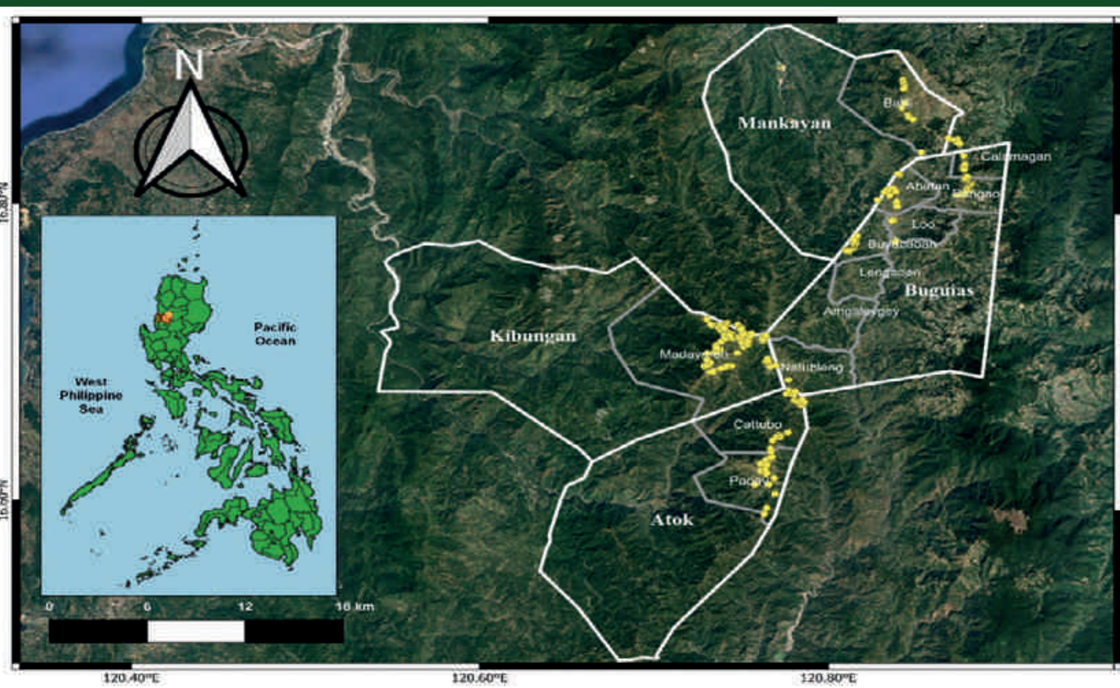
c. Neighboring or adjacent farm parcels planted with potato were treated as separate farms if the owners are different.

Data Collection and Instrument

The developed farm surveillance and assessment form was subjected to face validity. The variables were validated by a local plant pathologist and potato late blight expert. The layout and logical arrangement of the items and appropriateness of the scale of measurements for each variable to the Path Analysis and Classification and Regression Tree (C&RT) model were validated by statistics experts. The study mainly employed farm assessment and farmers' interviews. Of course, willingness and consent from the farmers/ farm owners were sought before the conduct of the farm assessment and farmers' interview.

Figure 4

Map of the Actual Locations of the Potato Farms (Yellow Dots) in the Major Growing Areas in Benguet.



The host variables considered were the number of days from planting to day of assessment (age), number of times the planting material was planted in the same farm, variety, and growth characteristics (canopy cover and plant vigor, and growth habit). The canopy cover (%) was measured using the Canopeo Mobile Application (Patrignani & Ochsner, 2015) at each sample plot and averaged for the entire farm. The plant vigor was the average from individual score of plants in plot samples using the 5 – point rating scale of the International Potato Center (CIP), that is 5 (Highly Vigorous: Plant was strong with robust stems and leaves; light to dark color), 4 (Moderately Vigorous: Plant was moderately strong with robust stems, and leaves were light green), 3 (Vigorous: Better than less vigorous), 2 (Less Vigorous: Plant is weak with few thin stems, and leaves are pale) and 1 (Poor Vigor: Plants are weak with few stems, and leaves are very pale).

The environmental factors include air temperature (°C), relative humidity (%), and elevation. The temperature and relative humidity on the day of the farm assessment were measured using the Plantix mobile application (Rupavatharam et al., 2018). Although this mobile application is mainly for diagnostic purposes using image recognition and a deep learning algorithm, it provides weather forecasts for the location where the user is (Mendes et al., 2020). The elevation (masl) of the farm locations was extracted from the geotagged photos taken by DAGeocamera through QGIS Desktop 3.22.9.

Human interventions were the farming practices, particularly on plot characteristics, planting methods, fertilizer application, and fungicide applications. Plot characteristics were specifically the spacing (cm), distancing (cm), and width (cm) of plots. The planting method includes the single row planting method and plant spacing (cm). Fungicide application includes the type of fungicide (contact and systemic), number of fungicides regularly being used, amount of application (mL/10m plot), frequency of application (per week), and the age of the crops the first fungicide was applied. Fertilizer application includes the types (chicken manure and inorganic), timing (basal and side-dressing), and the amount of application (kg/10m plot).

For the pathogen, the severity of symptoms (e.g., lesions) on the leaf and stem of sample plants was visually assessed by the direct estimation method assisted by the Standardized Area Diagram by Cruickshank et al. (1982) at the plant level, then averaged across the sampled plants to represent the entire farm. Whilst the incidence is the ratio of the infected plants to the total plants for each sample plot, then averaged to represent the entire farm.

There are 20 sample plots for each of the potato farms observed. The sample plots were selected through systematic random sampling. All of the plants within each sampled plot are then individually assessed, particularly on plant vigor, plant spacing, and disease incidence and severity. The canopy cover (%), plot spacing (cm), and plot distancing (cm) are measured at the level of sample plots.

Data Treatment and Analysis

Initial data analysis was conducted to familiarize oneself with the common assumptions in any regression analysis. Importantly, the normal distribution of residuals of the continuous variables was tested through the Shapiro-Wilk test, and the outliers were visualized through the Boxplot.

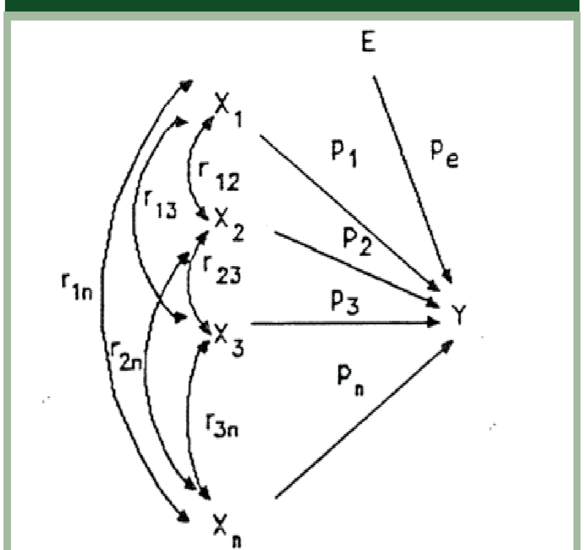
Since the host variables, environmental factors, and cultural management practices were a combination of continuous and dichotomous/ binary categorical, Two-step Clustering Analysis through the Statistical Packages for Statistical Solutions (SPSS) trial version 25 was employed to determine the clusters of the observed farms and the pattern that can be inferred from it (International Business Machines Corporation [IBM], 2016). The continuous variables were standardized, and 25% (default) of the noise/ outliers were handled due to non-normal distribution of residuals and a significant number of outliers observed (IBM, 2016). For cluster evaluation, the Schwarz's Bayesian Criterion (BIC) and Akaike's Information Criterion (AIC) were used to determine the best number of clusters; the silhouette measure that indicates the clarity of the separation among clusters and cohesion within each cluster (IBM, 2016; Thinsungnoen et al., 2015), and; the Analysis of Variance (ANOVA) with Tukey's Honest Significant Difference (Tukey HSD) for unplanned pairwise mean comparison for continuous variables and Chi-square test for Independence for categorical data. The ANOVA was still used for the continuous variables since it remains robust and valid for a single factor, even though normality was not met (Blanca et al., 2017).

The recursive path diagram and regression tree accounted for the explained variance and modeled the causal relationship of the variables to the incidence and severity of potato late blight were constructed through the Path Analysis and the Classification and Regression Tree (C&RT) model. The SPSS Amos trial version 23 was used to construct the recursive path diagram, while the Scikit-Learn Machine Learning in Python using Anaconda IDE (Integrated Development Environment) platform with the utilization of the pandas, numpy, matplotlib, pyplot, and seaborn libraries for the regression tree.

The path analysis method is an extension of multiple regression analysis and estimates the magnitude and strength of effects within a hypothesized causal system/ model, such as the recursive path diagram represented in Figure 5 (Lleras, 2005). The recursive path diagram can be aligned to the disease tetrahedron (Figure 2). A series of parameters is estimated by solving one or more structural equations to test the fit of the correlation matrix of the observed data between the two or more hypothesized causal models.

Figure 5

Path Diagram for n Correlated X's



Being an extension of multiple linear regression, the equation for the recursive path diagram (Figure 5) can be written as (Bondari, 1990):

$$Y = p_1 X_1 + p_2 X_2 + \dots\dots\dots + p_n X_n + p_e E$$

where, $Y = (y - \bar{y})/\sigma y$; $X_i = (x_i - \bar{x})/\sigma_{xi}$; $i = 1,2,\dots,n$; $E = e/\sigma_e$

p_i's, i = 1... n, are standardized partial regression coefficients (path coefficients) to be estimated from the data as:

$$p_i = \beta_i \sigma_{xi}/\sigma_{yi} \qquad i = 1,2,\dots,n$$

There are varieties of goodness of fit statistics that can be used to assess model fit and evaluate the model, i.e. path diagram. Aside from the chi-square statistics, there is also a set of indices like absolute fit index and incremental fit indices in Table 1 (Hu & Bentler, 1999).

The Classification and Regression Tree (C&RT) is a form of binary recursive partitioning that is quite similar to variable selection in regression. The parent node would always split into exactly two child nodes (binary), and each node will in turn become a parent node and split into exactly two child nodes (recursive). It partitions the independent variables into a series of leaves that contain the most homogenous collection of outcomes possible (Moisen, 2008; Lewis, 2000).

There are many decision trees, but considering that the response variables, incidence and severity, are continuous variables, the regression tree will be constructed (Morgan, 2014). The observation X is partitioned by a sequence of binary splits into terminal nodes (Figure 6). In each terminal node t, the predicted response value $y(t)$ is constant.

The first method that was often used in selecting a tree is through growing a maximum large tree successively until for every $t \in \check{T}_{max}$, $N(t) \leq N_{min.}$ Usually, Nmin is taken as five (5). To select the right-sized tree from the sequence T1 > T2 > ..., honest estimates of hyperparameters R(Tk) are needed to provide the minimum error-complexity measure.

The evaluation of the regression tree could be done by independent test sample estimates where the cases are randomly divided into a learning

Table 1

Formulae, Description, and Acceptability Cut-off for Incremental and Absolute Fit Indices for Maximum Likelihood Methods (Hu & Bentler, 1999)

Formula	Acceptability Cut-Off
TLI (or NNFI) = $\frac{\frac{T_B}{df_B} - \frac{T_T}{df_T}}{\frac{T_B}{df_B} - 1}$	≥ 0.95
CFI = $1 - \frac{max\ (T_T - df_T),\ 0}{\max((T_T - df_T),\ ((T_B - df_B,\ 0)}$	≥ 0.95
RMSEA = $\sqrt{\hat{F}/_{df_T}}$, where, $\hat{F} = \max\ (\frac{T_T - df_T}{n-1}),\ 0$	≤ 0.06

Note. *TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = root mean squared error of approximation.*

sample $£_1$ and a test sample $£_2$. The learning sample $£_1$ is used to grow the sequence of pruned trees with the minimum error-complexity measures (MAE, MSE, RMSE), while the test sample $£_2$ will be used to validate its accuracy (r^2).

In most practice, cross-validation is used unless the dataset is quite large. In *V*-fold cross-validation, £ is randomly divided into $£_1$, ..., $£_V$ such that each subsample $£_V$, v = 1, ..., *V*, has the same number of cases as nearly as possible. The *v*th learning sample is $£(V) = £ - £v$, and repeat the tree growing and pruning procedure using $£(V)$. For each v, this produces the tree $T^{(V)}$ (α), which is the minimal error-complexity tree for the parameter value α.

Results and Discussion

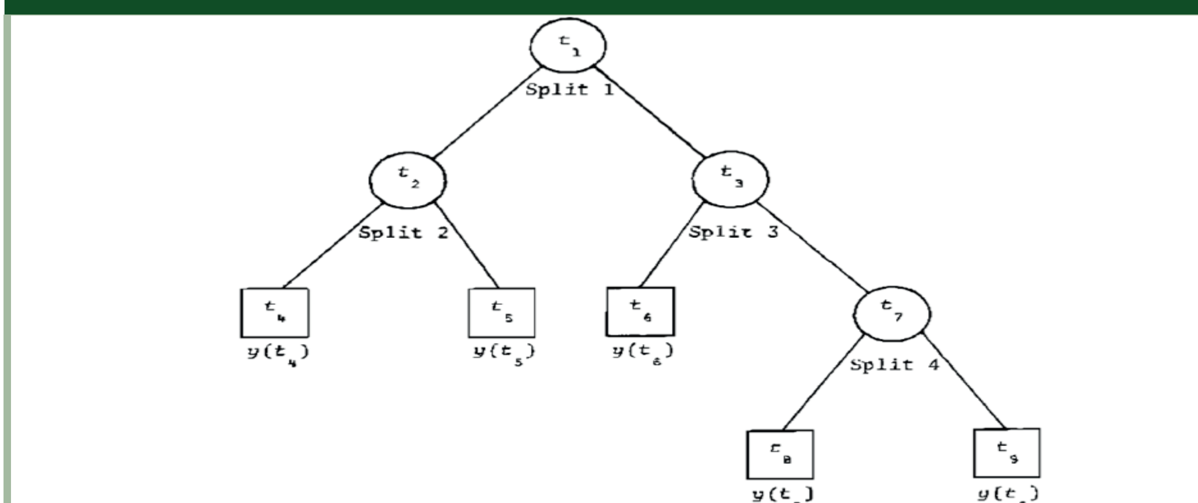
The Clusters of the Observed Potato Farms

Three (3) clusters emerged as the best number of clusters of the observed farms against the simulated 15 clusters. Three clusters have the lowest Schwarz's Bayesian Criterion (BIC) measure and have a significant change, "elbow" point, in Akaike's Information Criterion (AIC) (Figure 7) (IBM, 2016; Mohamed & Awang, 2015).

Cross-tabulating the three clusters against the municipalities where the observed potato

Figure 6

Binary Recursive Partitions of Sample Space X



farms were located, the three clusters were observed to be almost equally distributed into the three distinct major areas (Table 2).

The potato farms assessed in Kibungan (100%) were in cluster 1, while the potato farms assessed in Atok (94%) were in cluster 2, and the potato farms assessed in Buguias (100%) and Mankayan (98.4%) were in cluster 3. Since the clusters are expected to differ significantly in the host variables, environmental conditions, and cultural management practices, the differences among clusters were evaluated by inferential tests, namely, ANOVA with Tukey HSD for continuous variables and Chi-Square Test for Independence for categorical variables.

Potato Late Blight Symptoms

The lesions that were suspected to have been caused by *P. infestans* were assessed visually on the leaves and stems of the potato crops by direct estimation assisted by the Standardized Area Diagram by Cruickshank et al. (1982). The observed incidence (%) and severity (%) among clusters were significantly different (Table 3).

As revealed, potato farms in cluster 3 have been consistently observed with lower incidence and the lowest severity of potato late blight. Hence, it is worth examining the host variables (Table 4), environmental factors (Table 5), and cultural management practices of the farmers

Figure 7

Determining the Best Number of Clusters by BIC and AIC

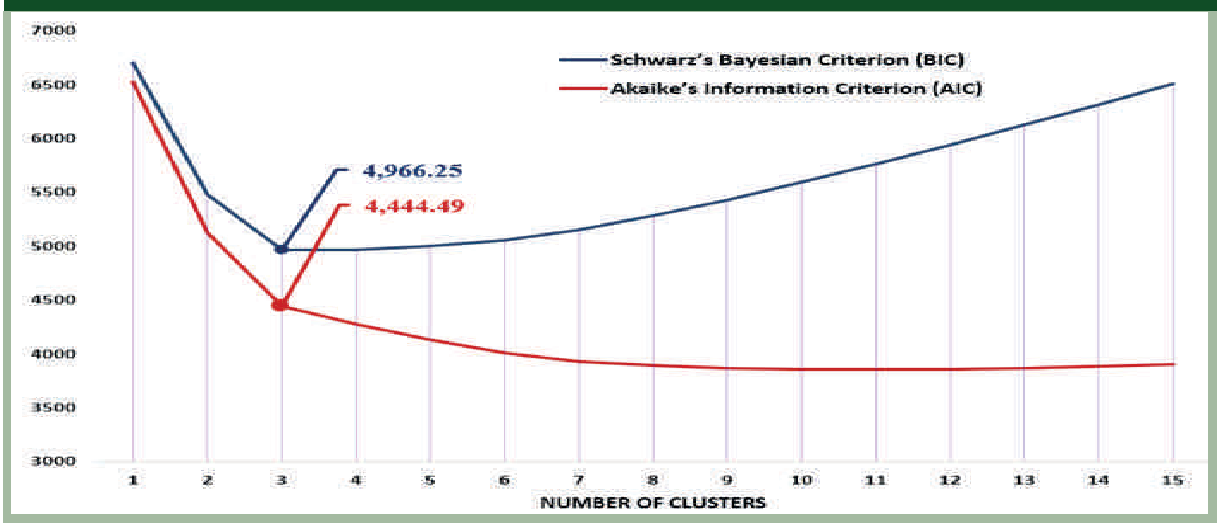


Table 2

Cross-Tabulation of the Municipalities the Farms Location with the Determined Clusters

MUNICIPALITIES	CLUSTER 1	CLUSTER 2	CLUSTER 3	TOTAL
Atok	6	94	-	100
(Paoay and Cattubo)	6.0%	94.0%	-	100%
Buguias	-	-	38	38
(Loo, Buyacaoan, and Bangao)	-	-	100.0%	100%
Kibungan	100	-	-	100
(Madaymen)	100.0%	-	-	100%
Mankayan	1	-	60	61
(Cada, Balili)	1.6%	-	98.4%	98.4%

(Tables 6 to 8) to uncover what makes it consistently low in incidence and in severity compared to the other two clusters.

Host Variables

The age of potato crops observed in cluster 3 is relatively young, at the inflorescence stage (Nemes et al., 2008), which explains the high vigor, implying a less predisposal to disease infections

Table 3

Comparison Among Clusters for the Potato Late Blight Symptoms

STEM AND LEAF LESIONS		CLUSTER 1 *Kibungan (Madaymen*	CLUSTER 2 *Atok (Paoay and Cattubo)*	CLUSTER 3 *Northern Buguias, and Cada, Mankayan*	*P*-Value
Incidence (%)	*μ*	0.793[a]	0.326[b]	0.374[b]	<0.000**
	sd	0.224	0.141	0.461	
Severity (%)	*μ*	14.484[a]	19.907[a]	6.556[b]	<0.000**
	sd	17.084	16.820	15.038	

Means with the same letter indicates no significant difference at 5% level of significance
*μ = mean; sd = standard deviation; ** indicates high significant differences*

Table 4

Comparison Among Clusters for Host Variables

STEM AND LEAF LESIONS		CLUSTER 1 *Kibungan (Madaymen*	CLUSTER 2 *Atok (Paoay and Cattubo)*	CLUSTER 3 *Northern Buguias, and Cada, Mankayan*	*P*-Value
Age (days after planting)	*μ*	74.402[b]	85.064[a]	44.806[c]	<0.000**
	sd	11.021	21.933	16.754	
No. of Times Planting Material was Replanted	*μ*	2.888[a]	2.160 [b]	1.010[c]	<0.000**
	sd	0.945	0.794	0.101	
Canopy Cover (%)	*μ*	66.102[a]	47.124[c]	57.834[b]	<0.000**
	sd	10.491	6.214	24.662	
Average Plant Vigor	*μ*	4.225[b]	4.080[b]	4.559[a]	<0.000**
	sd	0.830	0.655	0.745	
Variety					
Granola	*f*	63	61	98	<0.000**
	%	58.9	64.9	100.0	
Igorota/LBR/Po3	*f*	44	33	0	
	%	41.1	35.1	0.0	
Growth Habit					
Spreading	*f*	16	85	29	<0.000**
	%	15.0	90.4	29.6	
Upright	*f*	91	9	69	
	%	85.0	9.6	70.4	

Means with the same letter indicates no significant difference at 5% level of significance
*μ = mean; sd = standard deviation; ** indicates high significant differences*

(Hosack & Miller, n.d.). A first-time use of planting material observed in cluster 3 also explains the less incidence and severity of potato late blight since utilization of tubers from the previous season or other farms would still be prone, or much worse, less resistant for disease infection (Ristaino et al., 2000). The varieties planted in all clusters are Granola and Igorota/Po3/LBR which are both resistant to potato late blight (Gonzales et al. 2016). Although all farms observed in cluster 3 are planted with Granola variety, most (70.4%) are observed with upright growth habit, which is also necessary for air movement and good drainage that controls the progression of plant disease, as plot and plant spacing and distancing (Restiano et al., 2000). Though not the least dense, cluster 3 has still a less dense canopy cover (57.834%) that somehow possibly controls the disease development (Johnson & Cummings, 2021).

Environmental Factors

Although the air temperature, relative humidity, and elevation were significantly different among the clusters, these conditions recorded in all clusters are within the favorable condition for development of potato late blight (Table 5). The favorable temperature for potato late blight is within 18 to 22°C and relative humidity is <90% (Ristaino et al., 2000).

Commonly, air temperature and relative humidity decrease as the elevation increases. In this case, however, cluster 1 has the highest average elevation, yet has the highest average air temperature recorded. This could be due to mid-day observations of most potato farms in cluster 1, while clusters 2 and 3 were done at early morning and late afternoon.

Cultural Management Practices

Plot Characteristics and Planting Method. A wider plant spacing provides good drainage and air movement around the canopy cover that improves exposure to sun and air current that reduces the incidence and severity of potato late blight (Ristaino et al., 2000; Maloy, 2005). Indeed, the lower incidence and severity of potato late blight observed in cluster 3 is explained by having the widest plant spacing (27.54cm), plot distance (30.45cm), and plot width (46.07cm) (Table 6). Having the widest observed plot size, it is not surprising that there was a significant percentage of farms in cluster 3 that utilized planting method other than single row, which is most likely double row.

Fungicide Application. The use of chemical substances in combating plant disease, such as fungicides for potato late blight, is a common practice. As observed, most farms in the three clusters used contact fungicides, but a significant percentage in clusters 1 and 3 used systemic fungicides, and most farmers in cluster 3 regularly used three (3) contact fungicides (Table 7). The frequency of contact fungicide users is insignificant, implying its common use by almost all farmers in the three clusters.

The farmers in cluster 1 (highest incidence and severity) started to apply fungicide, on average, in the second week after planting (emergence of

Table 5

Comparison Among Clusters for Environment Variables

ENVIRONMENT CONDITIONS		CLUSTER 1 *Kibungan (Madaymen*	CLUSTER 2 *Atok (Paoay and Cattubo)*	CLUSTER 3 *Northern Buguias, and Cada, Mankayan*	*P*-Value
Air Temperature (°C)	*μ*	21.218[a]	19.424[b]	19.385[b]	<0.000**
	sd	1.811	2.377	2.064	
Relative Humidity (%)	*μ*	21.692[b]	30.000[a]	26.459[a]	<0.000**
	sd	10.012	14.482	13.681	
Elevation (masl)	*μ*	2437.92[a]	2306.67[b]	2109.29[c]	<0.000**
	sd	89.181	102.622	225.737	

Means with the same letter indicates no significant difference at 5% level of significance
*μ = mean; sd = standard deviation; ** indicates high significant differences*

Table 6

Comparison Among Clusters for Plant Spacing, Plot Distancing, Plot Width, and Planting Method

PLOT CHARACTERISTICS AND PLANTING METHOD			CLUSTER 1 *Kibungan (Madaymen*	CLUSTER 2 *Atok (Paoay and Cattubo)*	CLUSTER 3 *Northern Buguias, and Cada, Mankayan*	*P*-Value
Plant Spacing (cm)		*μ*	22.005^{c}	24.968^{b}	27.536^{a}	<0.000**
		sd	4.202	0.651	2.288	
Plot Distance (cm)		*μ*	22.849^{b}	29.404^{a}	30.453^{a}	<0.000**
		sd	5.429	2.656	5.167	
Average Plot Width (cm)		*μ*	36.801^{c}	39.331^{b}	46.066^{a}	<0.000**
		sd	5.003	1.702	9.565	
Single Row (Planting Method)	No	*f*	5	2	23	<0.000**
		%	4.7	2.1	23.5	
	Yes	*f*	102	92	75	
		%	95.3	97.9	76.5	

Means with the same letter indicates no significant difference at 5% level of significance
*μ = mean; sd = standard deviation; ** indicates high significant differences*

Table 7

Comparison Among Clusters for Fungicide Application Practices

FUNGICIDE APPLICATION PRACTICES			CLUSTER 1 *Kibungan (Madaymen*	CLUSTER 2 *Atok (Paoay and Cattubo)*	CLUSTER 3 *Northern Buguias, and Cada, Mankayan*	*P*-Value
Contact Application	No	*f*	3.0	0.0	3.0	0.244ns
		%	2.8	0.0	3.1	
	Yes	*f*	104.0	94.0	95.0	
		%	97.2	100.0	96.9	
Number of Contact applied regularly		*μ*	1.215^{c}	1.649^{b}	1.898^{a}	<0.000**
		sd	0.532	0.617	0.507	
Systemic Application	No	*f*	88.0	94.0	86.0	<0.000**
		%	82.2	100.0	87.8	
	Yes	*f*	19.0	0.0	12.0	
		%	17.8	0.0	12.2	
Number of Systemic applied regularly		*μ*	0.178^{a}	0.000^{b}	0.143^{a}	0.000**
		sd	0.384	0.000	0.431	
Plant Age at 1st Fungicide Application		*μ*	11.921^{b}	35.468^{a}	12.857^{b}	<0.000**
		sd	7.844	13.066	4.682	
Average Frequency of Fungicide Application per week		*μ*	1.21^{c}	2.029^{a}	1.574^{b}	<0.000**
		sd	0.435	0.389	0.485	
Amount of Fungicide Application (mL) for 10m Plot		*μ*	22.514^{a}	7.005^{b}	5.960^{b}	<0.000**
		sd	19.250	8.679	3.249	

Means with the same letter indicates no significant difference at 5% level of significance
*μ = mean; sd = standard deviation; ** indicates high significant differences*

sprouts), which is the same as farmers in cluster 3. Yet, most farmers in cluster 3 applied twice more in a week. Additionally, the farmers in cluster 3 recorded the highest number of contact fungicides and a significant proportion of systemic fungicides regularly being used. Thus, a higher number of fungicides being applied early (emergence of sprouts) explains further the significantly lower incidence and lowest severity of potato late blight observed in cluster 3. Moreover, the farmers in cluster 3 use the lowest amount of fungicide, perhaps because the potato crops observed in May 2021 are relatively young and highly vigorous compared to the other two clusters. Having the highest incidence and severity, it is understandable why farmers in cluster 1 use an amount four times higher than the other two clusters.

Fertilizer Application. The application of chicken manure and inorganic fertilizers (e.g., triple 14) was a common farmers' practice in potato crop production (Table 8). Although fertilizer application is being balanced with irrigation practices of the farmers, which was not included in this study, helps in promoting vigorous plants, however is not easy to accomplish as irrigation is a risk factor for disease development (Maloy, 2005). Nonetheless, farmers in cluster 3 applied both inorganic fertilizer and chicken manure at side-dressing, which is a common practice of farmers with abundant water for irrigation, while farmers in both clusters 1 and 2 applied chicken manure at basal and inorganic fertilizer at side-dressing which is a common practice of famers with lacking water for irrigation.

Table 8

Comparison Among Clusters for Fertilizer Application Practices

FERTILIZER APPLICATION PRACTICES			CLUSTER 1 *Kibungan (Madaymen*	CLUSTER 2 *Atok (Paoay and Cattubo)*	CLUSTER 3 *Northern Buguias, and Cada, Mankayan*	*P*-Value
Application of Inorganic Fertilizer at Basal	No	*f*	56.0	94.0	98.0	<0.000**
		%	52.3	100.0	100.0	
	Yes	*f*	51.0	0.0	0.0	
		%	47.7	0.0	0.0	
Application of Inorganic Fertilizer at Side-Dressing	No	*f*	51.0	0.0	0.0	<0.000**
		%	47.7	0.0	0.0	
	Yes	*f*	56.0	94.0	98.0	
		%	52.3	100.0	100.0	
Amount of Inorganic Fertilizer Application (kg) in 10m Plot		μ	2.040^{a}	2.042^{a}	1.399^{b}	<0.000**
		sd	1.174	1.124	0.473	
Application of Chicken Manure at Basal	No	*f*	0.0	1.0	98.0	<0.000**
		%	0.0	1.1	100.0	
	Yes	*f*	107.0	93.0	0.0	
		%	100.0	98.9	0.0	
Application of Chicken Manure at Side dressing	No	*f*	107.0	93.0	0.0	<0.000**
		%	100.0	98.9	0.0	
	Yes	*f*	0.0	1.0	98.0	
		%	0.0	1.1	100.0	
Amount of Chicken Manure Application (kg) for 10m Plot		μ	7.142^{b}	11.863^{a}	11.067^{a}	<0.000**
		sd	3.242	6.621	3.708	

Means with the same letter indicates no significant difference at 5% level of significance
*μ = mean; sd = standard deviation; ** indicates high significant differences*

Excessive use of nitrogen fertilizers increases canopy growth and cover and delays maturity, which may result in more foliage exposure to potential infection for a longer time and can lead to low yield (Ristaino et al., 2000). Cluster 3 has the lowest amount (kg) of inorganic fertilizer (e.g. triple 14), and has the highest amount (kg) of chicken manure application estimated for a 10 m plot. Being observed to have the significantly lowest recorded severity and incidence of late blight, such an amount of inorganic fertilizer and chicken manure is perhaps not excessive enough, as shown by significantly less dense canopy cover observed on the farms in cluster 3, other than the observed crop was at the inflorescence stage during the time of observation.

The Recursive Path Diagrams

The separate recursive path diagram for each of the host, environment, and cultural management practices and the combined recursive path diagram for all variables were constructed and fitted to the observed data. The goodness-of-fit of the path diagrams was compared through the Chi-square and Root Mean Squared Error for Approximation (RMSEA) at $\alpha = 0.05$, and the Tucker-Lewis Index (TFI) and the Comparative Fit Index (CFI). Of which, the separate path diagram of host variables and the combined path diagram of all variables fit the recursive path diagram and met the acceptable fit threshold of ≥ 0.95 (Hu & Bentler, 1999). This means that the covariance among the exogenous host variables may have attributed to the linear relationship with the endogenous variables (disease symptoms) despite non-normal distribution and the presence of outliers. Bondari (1990) explains that the non-linear relationship of exogenous variables to the endogenous variables is sometimes systematically transformed to have a linear relationship throughout the path diagrams with significant covariate exogenous variables.

Comparing the prediction accuracy by the coefficient of determination (r^2) of the separate recursive path diagrams for host variables (Incidence = 0.34; Severity = 0.79) and the combined recursive path diagram (Incidence = 0.53; Severity = 0.81), the latter emerged to be more accurate for predicting the incidence and severity of potato late blight. Thus, the combined recursive path diagram (Figure 8) will be considered further.

Since the estimates and important measures are not visible in the diagram (Figure 8), the matrix of the correlation coefficients, total effects, explained variance, and beta coefficient estimates for each variable to the incidence and severity of potato late blight are presented in Table 9.

Significant Predictors

Host variables. The common significant predictors from the host variables for the incidence and severity of potato late blight were the canopy cover and average plant vigor. Specific to the incidence, age is an additional significant predictor.

Having negative beta coefficients for both incidence and severity, this means that the plant vigor declines as the incidence and severity of potato late blight persist, or the plant vigor remains highly optimal at low incidence and severity of potato late blight. In contrast, the canopy cover has positive beta coefficients, which imply that the incidence and severity of potato late blight persist in potato crops with denser canopy cover. The plant age, a specific predictor for incidence only, has a positive beta coefficient, which proves that potato crops were indeed prone to late blight, or incidence persists as it grows to maturity. The age was not a significant predictor for severity as farmers would likely integrate disease management upon the first observation of symptoms such as foliar and stem lesions, which may have controlled it.

Environmental factors. The air temperature, relative humidity, and elevation were found to be not significant predictors for the observed incidence and severity. As mentioned, the average air temperature (19-21°C) and the average relative humidity, (24–27%) falls within the favorable temperature range of 18-22°C and relative humidity of <0.90 for sporulation of *P. infestans* to produce zoospores capable of initiating an infection (Ristaino et al., 2000). Therefore, the insignificance of the environmental variables as predictors does not necessarily imply insignificance to epidemiological and biological aspect for *P. infestans* to thrive and cause infection. Rather, the variability observed in these environmental variables were not sufficient to be significant predictors in the fluctuating incidence and severity of the observed farms.

Figure 8

The Combined Recursive Path Diagram of All the Variables on Host Variables, Environment Factors, and Cultural Management Practices Towards Incidence and Severity of Potato Late Blight

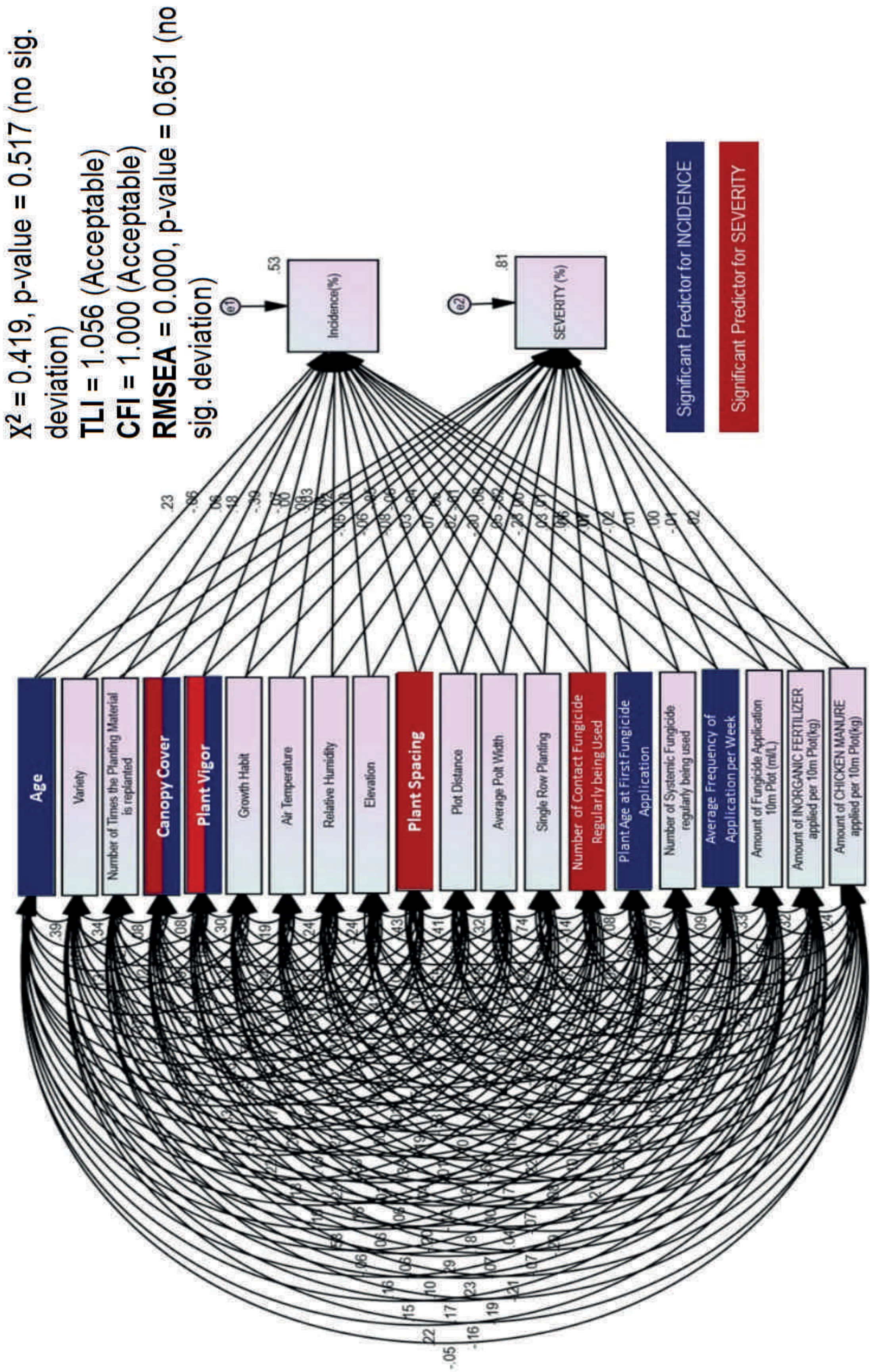


Figure 8

The Correlation Matrix, Total Effect, Explained Variance and Beta Coefficient Estimates of Variables to the Incidence and Severity of Potato Late Blight

VARIABLES		Age	Variety	No. of Times Planting Material was Replanted	Canopy Cover	Average Plant Vigor	Growth Habit	Air Temperature	Relative Humidity	Elevation	Plant Spacing	Plot Distance	Average Plot Width	Single Row Planting Method	No. of Contact Fungicide Regularly Being Used	No. of Systemic Fungicide Regularly Being Used	Plant Age at First Fungicide Application	Average Frequecy of Application per Week	Amount of Fungicide Application for 10m Plot	Amount of Inorganic Fertilizer for 10m Plot	Amount of Chicken Manure for 10m Plot
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
A	Corr	1																			
B	Corr	0.385	1																		
C	Corr	0.398	0.344	1																	
D	Corr	0.127	0.021	0.084	1																
E	Corr	-0.232	-0.015	-0.195	0.081	1															
F	Corr	-0.373	0.007	0.087	0.16	0.302	1														
G	Corr	0.117	0.189	0.33	0.123	0.034	0.19	1													
H	Corr	0.007	-0.114	-0.161	-0.167	-0.036	-0.171	-0.239	1												
I	Corr	0.344	0.261	0.505	0.145	-0.039	0.148	0.048	-0.245	1											
J	Corr	-0.333	-0.355	-0.42	-0.197	-0.033	-0.236	-0.21	0.091	-0.434	1										
K	Corr	-0.195	-0.166	-0.402	-0.124	-0.012	-0.27	-0.407	0.081	-0.303	0.407	1									
L	Corr	-0.22	-0.265	-0.4	0.091	0.016	-0.126	-0.173	0.141	-0.548	0.392	0.322	1								
M	Corr	0.135	0.146	0.213	-0.13	-0.025	0.089	0.023	-0.166	0.404	-0.185	-0.096	-0.742	1							
N	Corr	-0.162	-0.226	-0.333	-0.005	0.01	-0.155	-0.191	0.041	-0.31	0.306	0.238	0.296	-0.139	1						
O	Corr	-0.056	-0.056	-0.059	0.042	-0.008	0.105	0.017	0.039	-0.05	0.055	-0.05	-0.003	-0.055	-0.204	1					
P	Corr	0.53	0.158	0.01	-0.326	-0.192	-0.529	-0.168	0.167	-0.014	0.072	0.236	-0.067	0.097	0.079	-0.065	1				
Q	Corr	0.163	-0.06	-0.2	-0.127	-0.055	-0.361	-0.177	0.236	-0.221	0.224	0.33	0.131	-0.078	0.179	-0.092	0.418	1			
R	Corr	0.146	0.102	0.291	0.182	0.001	0.168	0.224	-0.007	0.237	-0.305	-0.452	-0.269	0.116	-0.148	0.133	-0.214	-0.328	1		
S	Corr	0.224	0.172	0.229	-0.066	-0.044	-0.071	0.079	0.099	0.161	-0.228	-0.195	-0.231	0.024	-0.117	-0.02	0.102	0.05	0.318	1	
T	Corr	-0.049	-0.163	-0.19	-0.208	-0.066	-0.291	-0.096	0.213	-0.271	0.237	0.182	0.174	-0.172	0.114	-0.079	0.237	0.214	-0.088	0.238	1
Total Effects	Incidence	0.231	-0.059	0.055	0.177	-0.386	-0.068	0.094	0.081	-0.052	-0.058	-0.081	-0.027	-0.071	-0.019	0.054	-0.299	-0.234	0.029	0.041	-0.071
	Severity	0.001	-0.03	0.023	0.103	-0.85	-0.064	-0.037	0.053	-0.008	-0.079	-0.02	-0.003	0.008	0.059	-0.02	0.07	0.009	0.003	-0.01	0.02
Explained Variance	Incidence	5.34%	0.35%	0.30%	3.13%	14.90%	0.46%	0.88%	0.66%	0.27%	0.34%	0.66%	0.07%	0.50%	0.04%	0.29%	8.94%	5.48%	0.08%	0.17%	0.50%
	Severity	0.00%	0.09%	0.05%	1.06%	72.25%	0.41%	0.14%	0.28%	0.01%	0.62%	0.04%	0.00%	0.01%	0.35%	0.04%	0.49%	0.01%	0.00%	0.01%	0.04%
Incidence	Coefficient	0.004*	-0.05	0.019	0.004*	-0.186*	-0.051	0.016	0.002	0	-0.006	-0.005	-0.001	-0.088	-0.012	0.058	-0.008*	-0.158*	0.001	0.015	-0.005
	p-Value	<0.0001	0.206	0.314	<0.0001	<0.0001	0.212	0.051	0.071	0.367	0.269	0.129	0.711	0.265	0.676	0.204	<0.0001	<0.0001	0.561	0.384	0.131
Severity	Coefficient	0.001	-1.169	0.374	0.101*	-18.869*	-2.212	-0.284	0.069	-0.001	-0.371*	-0.059	-0.006	0.434	1.617*	-0.998	0.085	0.268	0.003	-0.173	0.067
	p-Value	0.985	0.318	0.511	<0.0001	<0.0001	0.066	0.228	0.067	0.823	0.019	0.566	0.954	0.853	0.048	0.464	0.089	0.784	0.933	0.734	0.508

NOTE: The blue highlight with asterisk (*) indicates significant predictor at 5% level of significance

Cultural Management Practices. The plant age during the first fungicide application and average frequency of fungicide application per week were found to be significant predictors for the incidence of potato late blight. Contact fungicides were supposed to be applied before the incidence of potato late blight; however, farmers tend to apply fungicides at the onset of potato late blight symptoms. Hence, the negative beta coefficient of the plant age at first fungicide application indicates that the high incidence of potato late blight at an early age was due to the earlier onset of the disease and not from the early application of fungicide. The negative beta coefficient of the average frequency of application per week shows that lower incidence of potato late blight in farms with more frequent application of fungicide per week.

For the severity of potato late blight, the plant spacing and number of contact fungicides regularly used were significant predictors. The negative beta coefficient of plant spacing implies that wider spaces between plants help to lessen the severity of potato late blight. The positive beta coefficient of the number of contact fungicides does not necessarily imply that the severity of potato late blight worsens by having more contact fungicides regularly being used. Rather, farmers curbed the severity of disease by applying fungicides either simultaneously or alternately. In other words, the farmers tend to use more contact fungicides, not systemic, as the severity of late blight persists on their crops.

Total Effects and Explained Variances of Significant Predictors

Host Variables. The average plant vigor provides the highest explained variance to both the incidence and severity of potato late blight by 14.89% and 72.25%, respectively. The canopy cover explains 3.13% and 1.06% of the variability of incidence and severity of potato late blight, respectively. The age, being found significant predictor for the incidence of potato late blight only, explains 5.34%. Having less than 1% or near to 0% explained variance, it only means that the observed data for the variables with such explained variance have inadequate regression weights or effects to be significant predictors.

Environmental Factors. The air temperature, relative humidity, and elevation explain less than 1% or nearly zero of the variability of the incidence and severity of potato late blight. Again, such explained variability of environmental variables does not necessarily indicate their insignificant regression weights to plant disease epidemiology but only implies inadequacy in variability of the observed data to explain the fluctuation of the incidence and severity of the observed late blight symptoms.

Cultural Management Practices. The plant age at first fungicide application explains 8.94% in the variability of the incidence of potato late blight, while the average frequency of fungicide application per week explains 5.48%. Whilst plant spacing explains 0.62% in the variability of the severity of potato late blight, and the number of contact fungicides regularly being used explains 0.35%. Although the latter variables explain below 1% variance in the severity of potato late blight, these cultural management practices, as observed data, are considered to have significant regression weights to predict the severity of potato late blight.

The rest of the cultural management practices explain a certain amount of the variability of the endogenous variables (incidence and severity) of which contributed to the accuracy of the recursive path diagram. However, the regression weights are insufficient to be significant predictors of the incidence and severity of potato late blight.

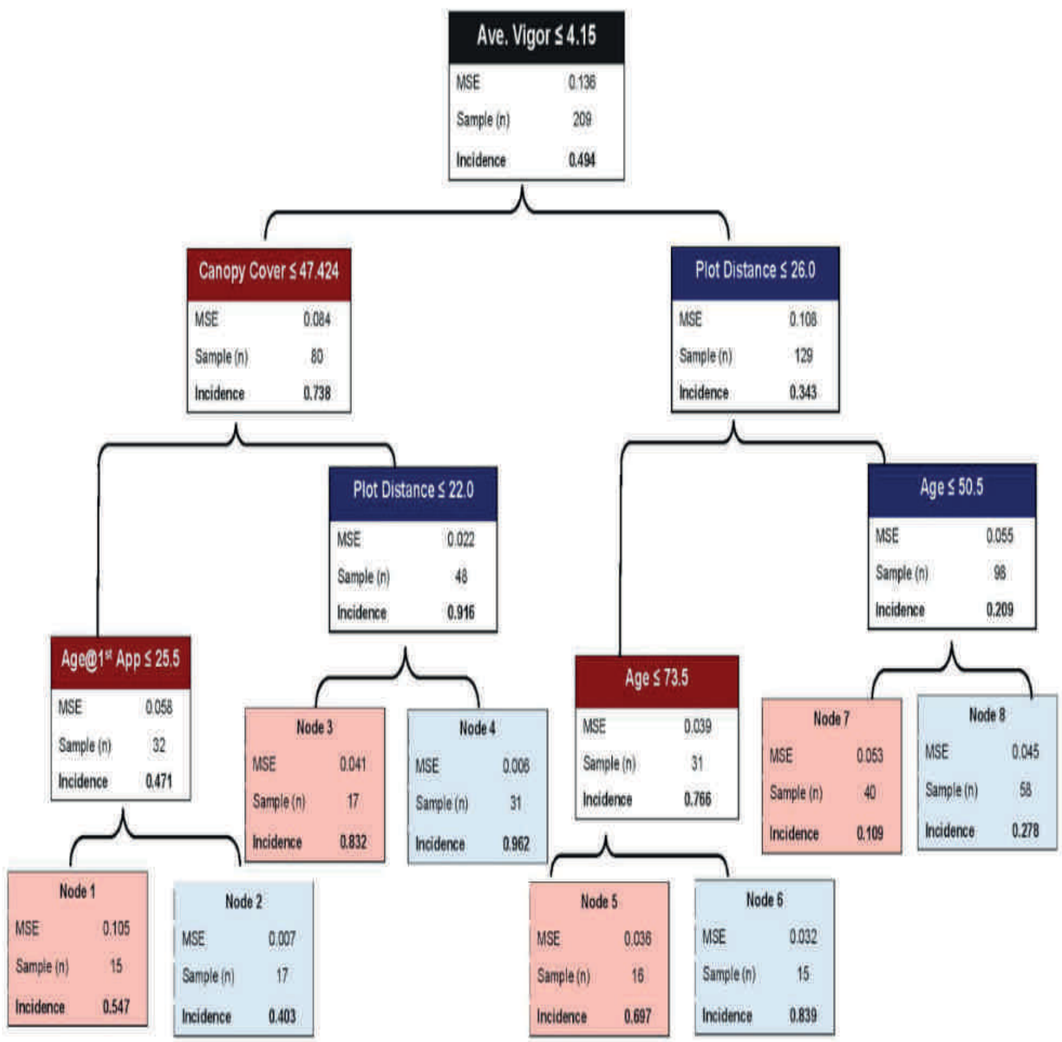
The Regression Trees

In any continuous type of data, a non-parametric test like the Classification and Regression Tree (C&RT) is recommended when there are significant outliers and non-normal distribution. Through C&RT, the regression trees for incidence (Figure 9) and severity (Figure 10) of potato late blight were constructed. The regression trees are capable of identifying the variables on host, environment conditions, and cultural management practices that have the highest significant association with the observed incidence and severity of potato late blight, regardless of their distribution.

Employing the GridSearchCV tuning technique, the best tuned hyperparameters were determined to create the regression trees that yield the minimum error-complexity measures, Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and best tree scores (coefficient of determination, r^2)

Figure 9

The Best Regression Tree for the Incidence of Potato Late Blight

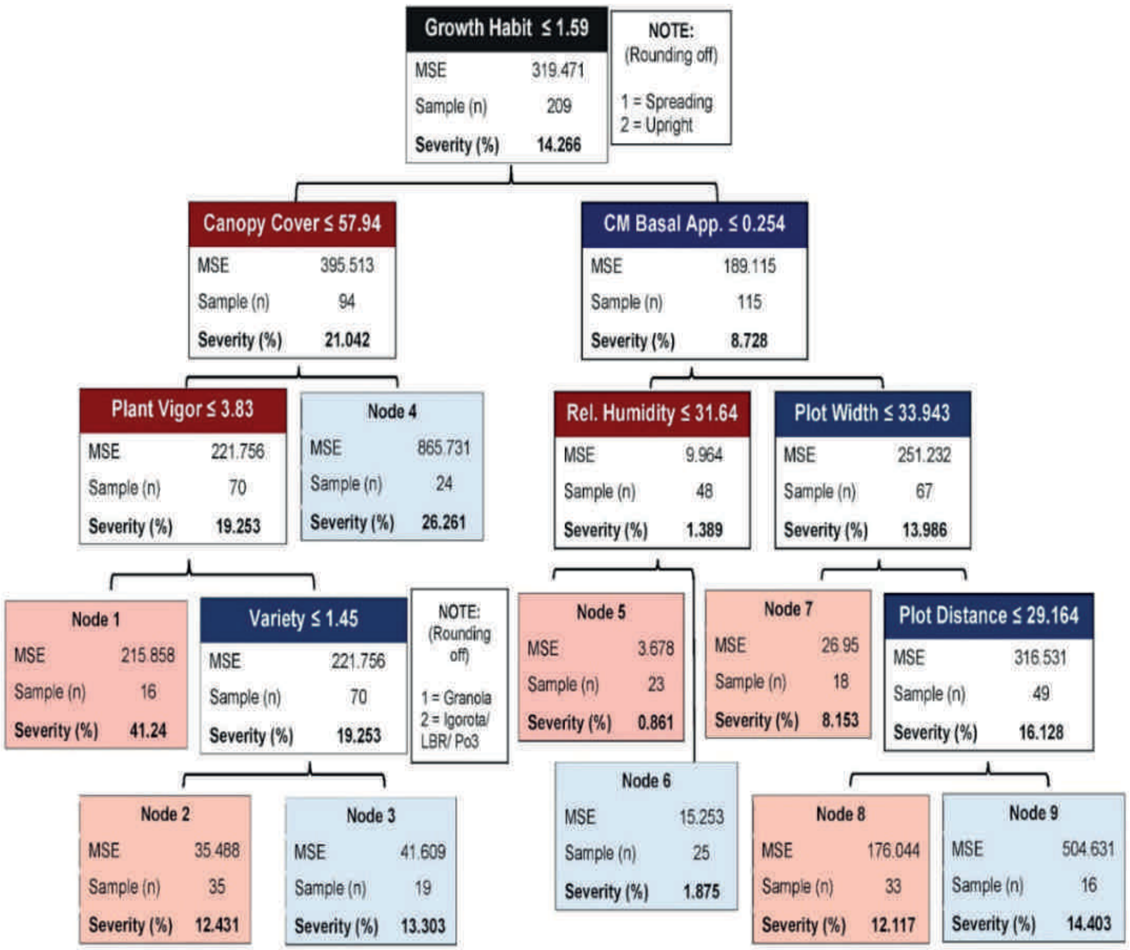


NODE	STATEMENT	PREDICTED INCIDENCE
1	if Ave. Plant Vigor ≤ 4.150 (Moderate), if Canopy Cover ≤ 47.424%, if Age at 1st Fungicide Application ≤ 25.50DAP	0.547
2	if Ave. Plant Vigor ≤ 4.150 (Moderate), if Canopy Cover ≤ 47.424%, if Age at 1st Fungicide Application > 25.50DAP	0.403
3	if Ave. Plant Vigor ≤ 4.149541 (Moderate), if Canopy Cover > 47.424%, if Plot Distance ≤ 22.0 cm,	0.832 (highest)
4	if Ave. Plant Vigor ≤ 4.150 (Moderate), if Canopy Cover > 47.424%, if Plot Distance > 22.0 cm	0.962 (highest)
5	if Ave. Plant Vigor > 4.150 (Moderate), if Plot Distance ≤ 26.0 cm, if Age ≤ 73.50 DAP	0.697
6	if Ave. Plant Vigor > 4.150 (Moderate), if Plot Distance ≤ 26.0 cm, if Age > 73.50 DAP	0.839 (highest)
7	if Ave. Plant Vigor > 4.150 (Moderate), if Plot Distance > 26.0 cm, if Age ≤ 50.5 DAP	0.109 (lowest)
8	if Ave. Plant Vigor > 4.150 (Moderate), if Plot Distance > 26.0 cm, if Age ≤ 50.5 DAP	0.278 (lowest)

Mean Absolute Error **(MAE)** = 0.1926
Mean Squared Error **(MSE)** = 0.0724
Root Mean Squared Error **(RMSEA)** = 0.269
Coefficient of Determination **(R^2)** = 0.4902

Figure 10

The Best Regression Tree for the Severity of Potato Late Blight



NODE	STATEMENT	PREDICTED SEVERITY
1	if Growth Habit ≤ 1.594 (Spreading), if Canopy Cover ≤ 57.949%, if Plant Vigor ≤ 3.827(Moderate),	41.240 (highest)
2	if Growth Habit ≤ 1.594 (Spreading), if Canopy Cover ≤ 57.949%, if Plant Vigor > 3.827 (Moderate), if Variety ≤ 1.478 (Granola),	12.431
3	if Growth Habit ≤ 1.594 (Spreading), if Canopy Cover ≤ 57.949%, if Plant Vigor > 3.827 (Moderate), if Variety > 1.478 (Igorota/LBR/Po3)	13.303
4	if Growth Habit ≤ 1.594 (Spreading), if Canopy Cover > 57.949%	26.261
5	if Growth Habit > 1.594 (Upright), if Chicken Manure Basal Application ≤ 0.254 (No), if Relative Humidity ≤ 31.644%,	0.861 (lowest)
6	if Growth Habit > 1.594 (Upright), if Chicken Manure Basal Application ≤ 0.254 (No), if Relative Humidity > 31.644%,	1.875 (lowest)
7	if Growth Habit > 1.594 (Upright), if Chicken Manure Basal Application > 0.254 (YES), if Plot Width ≤ 33.943 cm,	8.153
8	if Growth Habit > 1.594 (Upright), if Chicken Manure Basal Application > 0.254 (YES), if Plot Width > 33.943 cm, if Plot Distance ≤ 29.164 cm,	12.117
9	if Growth Habit > 1.594 (Upright), if Chicken Manure Basal Application > 0.254 (YES), if Plot Width > 33.943 cm, if Plot Distance > 29.164 cm,	14.403

Mean Absolute Error **(MAE)** = 8.5245
Mean Squared Error **(MSE)** = 187.8564
Root Mean Squared Error **(RMSEA)** = 13.7061
Coefficient of Determination **(R^2)** = 0.7728

among the 96,000 possible combinations of the hyperparameters multiplied to 10-folds cross-validation.

Incidence of Potato Late Blight

The pruned regression tree by the tuned hyperparameter with the minimum error (MSE, MAE, RMSE) and better predictive and classification accuracy (r^2) for the incidence of potato late blight is presented in Figure 10. For guidance and a more comprehensive understanding of the regression tree for the incidence of potato late blight, an algorithm is presented alongside it.

Among all the variables assessed and observed, the average plant vigor, canopy cover, plot distance, plant age, and age of first fungicide application were the variables that had the highest significant association with the variability in the observed incidence of potato late blight.

At the very bottom part of the regression tree emerged eight (8) terminal nodes that classify the incidence of potato late blight observed in Benguet farms. Each terminal node follows branches leading to the parent node with information attached relevant for decision making and understanding the pattern. For instance, the terminal nodes 3 and 4 that emerged from the same branch with respective incidences of 0.832 and 0.962 (highest) were observed on potato farms with average plant vigor of below 4.15 (~moderately vigor) and denser canopy cover (> 47.42%) regardless of the plot distances. Terminal nodes 7 and 8, with respective incidences of 0.109 (lowest) and 0.278, are observed on potato farms with plant vigor of above 4.15 (~highly vigor), and wider plot distance (>26cm) and relatively young (50.5 days old ~Inflorescence stage).

Severity of Potato Late Blight

The pruned regression tree by the tuned hyperparameter with the minimum error (MSE, MAE, RMSE) and better predictive and classification accuracy (r^2) for the severity of potato late blight is presented in Figure 9. For guidance and a more comprehensive understanding of the regression tree for the incidence of potato late blight, an algorithm is presented alongside it.

Among all the variables assessed and observed, the growth habit, canopy cover, average plant vigor, variety, basal application of chicken manure, relative humidity, plot width, and plot distance were the variables that had the highest significant association with the variability in the observed incidence of potato late blight.

Similarly, the very bottom part of the regression tree presented five (9) terminal nodes that classify the severity of potato late blight observed in Benguet farms. Each terminal node follows branches leading to the parent node with information attached relevant for decision making and understanding the pattern. For instance, the terminal node 1 with the highest severity of 41.240% was observed on potato farms with crops having characteristics of spreading growth habit and an average of less than moderate vigor despite a less dense canopy cover (57.95%). Terminal nodes 5 and 6, with respective lowest severities of 0.861% and 1.875%, were observed on potato farms having a crop characteristic of an upright growth habit, farmers' practice of non-application of chicken manure at the basal, and regardless of the relative humidity of the environment.

Conclusions

Vis–à–vis the research problems and the results presented, the following conclusions are: (1) There were three (3) clusters of the observed potato farms derived from all the variables. All (100%) of the observed farms in Madaymen, Kibungan, fall in cluster 1, most (94%) of the observed farms in Atok (Paoay and Cattubo) fall in cluster 2, and all (100%) of the observed farms in Buguias (Loo, Buyacaoan, and Bangao) and almost all (98.4%) of the observed farms in Cada, Balili, Mankayan fall in cluster 3. Cluster 3 has the lowest significant severity (6.56%) and incidence (37.4%) observed, hence worth examining concerning the host, environment, and cultural management practices of the farmers.; (2) The combined recursive path diagram has the highest accuracy measure against the individual recursive path diagram for host variables, environmental factors, and cultural management practices. The significant predictors and accounted the highest total effect and explained variance to the incidence of potato late blight are (a) the host age, (b) the canopy cover, (c) the average plant vigor, (d) the host age at first fungicide application, and (e) the amount of fungicide application per week. Whilst the significant predictors and accounted for the highest total effects and explained variance for the severity of potato late blight are (a) the canopy cover, (b) the average plant vigor, (c) the plant spacing, and (d) the number of contact fungicides regularly applied. Most of the emerged significant predictors for both observed incidence and severity are host variables.; (3) The tuned regression tree for incidence of potato late blight associated with the host variables, environmental factors, and cultural management practices, yields an accuracy of 0.4902, while the tuned regression tree for severity of potato late blight yields an accuracy of 0.7728. The variables that are significant predictors for the incidence of potato late blight are (a) the host age, (b) the canopy cover, (c) the average plant vigor, (d) the average plot distance, and (e) the plant age at first fungicide application. Whilst, the variables that are significant predictors for the severity of potato late blight are (a) the growth habit, (b) the canopy cover, (c) the average plant vigor, (d) the variety, (e) the basal application of chicken manure, (f) the relative humidity, (g) the plot width, and (h) the plot distance. Most of the emerged significant predictors for both the observed incidence and severity were host variables.

Recommendations

Since most of the variables that emerged as significant predictors for the observed incidence and severity of potato late blight by the cluster analysis, combined recursive path diagram, and tuned regression trees are host variables, it is generally recommended that the farmers should focus on the "immune system" of their potato crop. The approach includes the use of planting materials that are resistant varieties with less dense canopy cover, highly vigorous, and upright growth habit, alongside wider plot distance (>29cm) and plot width (>33cm), and early application of contact fungicide to curb the incidence of potato late blight.

Relative to the conclusion vis–à–vis the research problems, the specific recommendations are as follows: (1) Inferred from cluster 3 for having recorded the lowest significant incidence and severity of potato late blight, it is recommended to consider resistant varieties with an upright growth habit and the use of planting materials that are relatively "new". In terms of plot preparation and planting method, single row planting with consideration of wider plot/ mound/hill width (46~50cm), plot distance (~30cm), and plant spacing (~27cm). Both chicken manure and inorganic fertilizers are applied at about or during side-dressing with an amount of 11.08 kg./10m plot and 1.4 kg./10m plot, respectively. Regular application of contact fungicide solutions amounting to ~6 ml/10m plot for at least twice a week at the emergence of sprouts. However, this recommendation may only apply for May or be relative to that cropping season; (2) The combined recursive path diagram, with an accuracy of 0.81, is recommended for the prediction or classification of the severity of potato late blight for independent farms observed in major growing areas in Benguet against the best regression tree for the severity of potato late blight. To maximize efficiency of the path diagram for prediction and classification, consider the variables (a) canopy cover (%), (b) average plant vigor, (c) plant spacing (cm), and (d) number of contact fungicide regularly applied as these are the determined significant predictors and explained the highest effects or variance on the severity of potato late blight; and (3) Excluding the accounted variance of the unobserved variables, the actual accuracy of the combined recursive path diagram is quite lower than the best regression tree for the incidence of potato late blight. It is then recommended that the best regression tree for the incidence of severity for the prediction and classification of the incidence of potato late blight for independent farms observed in major growing areas in Benguet. To maximize the prediction efficacy of the regression tree, consider the variables (a) host age (days after planting), (b) canopy cover (%), (c) average plant vigor, (d) average plot distance (cm), and (e) plant age (days after planting) at first fungicide application.

References

Blanca, M.J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema*, 29(4): 552–557. https://doi.org/10.7334/psicothema2016.383

Bentler, P.M., & Chou, C.H. (1987). Practical issues in structural modeling. *Sociological Methods & Research,* 16, 78-117.

Bondari, K. (1990). Path analysis in agricultural research. Conference on Applied Statistics in Agriculture. https://doi.org/10.4148/2475-7772.1439

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika,* 50, 229-242.

Boomsma, A. (1982). *Robustness of LISREL against small sample sizes in factor analysis models.* In K. G. Joreskog & H. Wold (Eds.), Systems under indirection observation: Causality, structure, prediction (Part I) (pp. 149-173). Amsterdam, Netherlands: North Holland.

Cruickshank, G., Stewart, H.E., & Wastie, R.L. (1982). An illustrated assessment key for foliage blight of potatoes. *Potato Research*, 25(2): 213–214. https://doi.org/10.1007/bf02359807

Cunniffe, N.J., Koskella, B., Metcalf, C.J.E., Parnell, S., Gottwald, T.R., & Gilligan, C.A. (2015). Thirteen challenges in modelling plant diseases. *Epidemics,* 10, 6–10. https://doi.org/10.1016/j.epidem.2014.06.002

Food and Agriculture Organization Corporate Statistical Database. (2021). Food and Agriculture Organization of the United Nation. http://www.fao.org/faostat/en/#data/QC.

Francl, L.J. (2001). The Disease Triangle: A plant pathological paradigm revisited. *The Plant Health Instructor.* https://doi.org/10.1094/phi-t-2001-0517-01

Gonzales, I.C., Kiswa, C.G., & Bautista, A.B. (2016). Sustainable potato production in the Philippine Cordillera Region. International *Journal of Engineering and Applied Sciences*, 3(6): 29–37.

Hosack, P., & Miller, L. (n.d.). Preventing and Managing Plant Diseases. MU Extension. https://extension.missouri.edu/publications/mg13

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal,* 6(1): 1–55. https://doi.org/10.1080/10705519909540118

International Business Machines Corporation. (2016). Two-step cluster analysis. https://www.ibm.com/docs/SSLVMB_23.0.0/spss/base/idh_twostep_main.html.

Jeger, M.J. (2004). Analysis of disease progress as a basis for evaluating disease management practices. *Annual Review of Phytopathology,* 42(1): 61–82. https://doi.org/10.1146/annurev.phyto.42.040803.140427

Johnson, D.A., & Cummings, T.F. (2021). Canopy environment in irrigation a key factor for managing late blight in semiarid environments. *Spudman.* https://spudman.com/article/canopy-environment-in-irrigation-a-key-factor-for-managing-late-blight-in-semiarid-environments/

Lewis, R.J. (2000). An introduction to classification and regression tree (CART) analysis [Paper Presentation]. Annual Meeting of the Society for Academic Emergency Medicine, San Francisco, California.

Lleras, C. (2005). Path analysis. In *Encyclopedia of Social Measurements* (Vol. 3, pp. 25–30). Elsevier Inc.

Madden, L.V., & Hughes, G. (1995). Plant Disease Incidence: Distributions, heterogeneity, and temporal analysis. *Annual Review of Phytopathology,* 33(1), 529–564. https://doi.org/10.1146/annurev.py.33.090195.002525

Maloy, O.C. (2005). Plant Disease Management. *The Plant Health Instructor.* https://doi.org/10.1094/phi-i-2005-0202-01

Mendes, J., Pinho, T.M., Neves Dos Santos, F., Sousa, J.J., Peres, E., Boaventura-Cunha, J., Cunha, M., & Morais, R. (2020). Smartphone applications targeting precision agriculture practices - A Systematic Review. *Agronomy*, 10(6): 855. https://doi.org/10.3390/agronomy 10060855

Mohamed, N., & Awang, S.R. (2015). The multiple intelligence classification of management graduates using two step cluster analysis. *Malaysian Journal of Fundamental and Applied Sciences,* 11(1). https://doi.org/10.11113/mjfas.v11n1.351

Moisen, G.G. (2008). *Classification and regression trees*. In: Jorgensen, Sven Erik; Fath, Brian D., eds. *Encyclopedia of Ecology*, Volume 1. Oxford, U.K.: Elsevier. p. 582-588.

Morgan, J. (2014). Classification and regression tree analysis. https://www.bu.edu/sph/files/2014/05/MorganCART.pdf.

Nayak, P., Mukherjee, A.K., Pandit, E., & Pradhan, S.K. (2018). Application of statistical tools for data analysis and interpretation in rice plant pathology. *Rice Science*, 25(1): 1–18. https://doi.org/10.1016/j.rsci.2017.07.001

Nemes, Z., Baciu, A., Popa, D., Mike, L., Petrus–Vancea, A., & Danci, O. (2008). The study of the potato's life–cycle phases important to the increase of the individual variability. Analele Universitatii din Oradea, Fascicula Biologie, Tom, XV, 60-63.

Nunnally, J.C. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.

Nutter, F.F.W. (2007). The Role of plant disease epidemiology in developing successful integrated disease management programs. *General Concepts in Integrated Pest and Disease Management,* 45–79. https://doi.org/10.1007/978-1-4020-6061-8_3

Patrignani, A., & Ochsner, T.E. (2015). Canopeo: A powerful new tool for measuring fractional green canopy cover. *Agronomy Journal,* 107(6): 2312–2320. https://doi.org/10.2134/agronj15.0150

Philippine Statistics Authority OPENSTAT. (2021). *Philippine Statistics Authority*. https://openstat.psa.gov.ph/PXWeb/pxweb/en/DB/DB__2E__CS/?tablelist=true&rxid=bdf9d8da-96f1-4100-ae09-18cb3eaeb313.

Rupavatharam, S., Kennepohl, A., Kummer, B., & Parimi, V. (2018) Automated plant disease diagnosis using innovative android app (Plantix) for farmers in Indian state of Andhra Pradesh. *Phytopathology,* 108 (10). http://oar.icrisat.org/11014/

Ristaino, J., Schumann, G.L., & D'arcy, C.J. (2000). Late blight of potato and tomato. *The Plant Health Instructor.* https://doi.org/10.1094/phi-i-2000-0724-01

Thinsungnoen, T., Kaoungku, N., Durongdumronchai, P., Kerdprasop, K., & Kerdprasop, N. (2015). The Clustering Validity with Silhouette and Sum of Squared Errors. *The Proceedings of the 2nd International Conference on Industrial Application Engineering 2015*. https://doi.org/10.12792/iciae2015.012

Turechek, W.W. (2004). Nonparametric tests in plant disease epidemiology: Characterizing disease Associations. *Phytopathology,* 94(9), 1018–1021. https://doi.org/10.1094/phyto.2004.94.9.1018

Van Maanen, A., & X.-M. XU. (2003). Modelling plant disease epidemics. European Journal of Plant Pathology. *European Journal of Plant Pathology,* 109(7): 669–682. https://doi.org/10.1023/a:1026018005613

Xu, X. (2006). Modelling and interpreting disease progress in time. *The Epidemiology of Plant Diseases,* 215–238. https://doi.org/10.1007/1-4020-4581-6_8