



# Intelligent diagnosis of diseases in plants using a hybrid Multi-Criteria decision making technique



Wayne Goodridge, Margaret Bernard\*, René Jordan, Reanne Rampersad

*The University of the West Indies, St. Augustine, Trinidad and Tobago*

## ARTICLE INFO

### Article history:

Received 12 August 2016

Received in revised form 23 November 2016

Accepted 1 December 2016

### Keywords:

Plant disease diagnosis

Analytic Hierarchy Process

Sensitive-Simple Additive Weighting

Multi-Criteria Decision Making

Expert system

AgriDiagnose

## ABSTRACT

This paper describes an Expert System that can intelligently diagnose diseases in plants. The system is dialog-based and uses a Multi-Criteria Decision Making technique that is a hybrid of Analytic Hierarchy Process and Sensitive Simple Additive Weighting. The paper describes an approach for disease modeling that uses a set of characteristics which are weighted for each disease using two types of weights: Relative Weights and Scales. The diagnostic process involves calculating the utility value for each disease based on the utility values of its characteristics. Experimental results show an accuracy of over 95%. The system implemented is called AgriDiagnose and it consists of a web-based pathology tool to model the diseases and a mobile app for farmers to interact with the system for disease diagnosis in the field.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Diseases have the potential to destroy large numbers of crops and can result in significant losses and food shortages if not detected and controlled in time. For example, the Papaya Ringspot virus affected the country of St. Kitts and destroyed about 90% of that country's production (Chin et al., 2007). Many developing countries organize plant clinics for farmers at which farmers can be educated about various pests and diseases and where plant Pathologists can diagnose diseases from samples that farmers bring to the clinic. This is often in addition to visits to the farms by Agriculture Extension Officers. Much work has also been done in trying to automate diagnosis (Barbedo, 2016; Gonzalez-Andujar et al., 2006; Mansingh et al., 2007). These Artificial Intelligence systems generally either apply image processing techniques to images of diseased plants or use a data entry dialog system to attempt a diagnosis.

In this paper, we present a dialog based system for diagnosis of plant diseases. The system uses a multi-criteria decision making technique that is a hybrid of Analytic Hierarchy Process (AHP) (Saaty, 1977) and Sensitive- Simple Additive Weighting (S-SAW) (Goodridge, 2016) to dynamically put forward questions to the

farmers in an optimal way and to reason through their responses returning a diagnosis. A major contribution of the paper is the approach presented for modeling diseases using a consistent set of characteristics (criteria). The AHP is used for determining weights of these characteristics for all diseases in the system. The diagnosis process uses S-SAW for sensitivity analysis. The S-SAW is an extension of the popular SAW method (Hwang and Yoon, 1981) which allows the decision maker to define an objective function which governs the optimization goals of each criterion. This is used in calculating the utility value of each characteristic.

This technique was implemented in a system called AgriDiagnose, a system that consists of a back-end, web-based pathology tool and a front-end mobile app for farmers. The results obtained from experimentation gives a 95.9% accuracy for diagnosing the correct disease and a 100% sensitivity result that the system returns a positive result when the plant is indeed diseased.

The rest of the paper is organized in this manner. In Section 2, we review some of the approaches taken in the literature to intelligent diagnosis of plant diseases. In Section 3, we describe the disease modeling that is configured in the Pathology tool and in Section 4, we describe the diagnostic process. We trace one case study throughout these two sections so that the reader can follow the process with data. In Section 5, we reveal the results obtained from our simulation exercises and introduce four metrics for measuring these results. We conclude in Section 6.

\* Corresponding author.

E-mail address: [margaret.bernard@sta.uwi.edu](mailto:margaret.bernard@sta.uwi.edu) (M. Bernard).

## 2. Background and related work

Farmers would benefit from a diagnostic process that could intelligently act as a human pathologist. This diagnostic process would take the same inputs that farmers typically provide to a pathologist, process them and return real-time diagnoses. Many of the expert systems that have been developed receive input data from images and use image processing methods or through data entry by users through a user interface. Barbedo (2016) provides a comprehensive survey of expert systems applied to plant disease diagnosis.

Here we review a select sample of the different artificial intelligence (AI) methods that have been applied by researchers and illustrated in Table 1.

A system that uses Fuzzy Logic together with image processing was developed to perform disease identification in the pomegranate crop (Sannakki et al., 2013). The application used images of the leaves of the plant taken by farmers, extracted the features of these leaves and processed them. Using Fuzzy Logic, they were then classified as either diseased or healthy and if found to be diseased, they were graded using the application's Fuzzy Inference System. Bashish et al. (2011) used image processing along with a neural network classifier to detect diseases in leaves. K-means clustering was performed on the images to find actual segments of the leaves followed by feature extraction of the diseased part of the leaves. Statistical analysis was performed to choose the best feature in the leaves, and finally, classification was executed using a neural network classifier.

Dewanto and Lukas (2014) developed an expert system for diagnosing pests and diseases in some of the main fruit plants grown in Indonesia. Their system used a rule-based dialog method. To cater for levels of uncertainty, rules had confidence variables applied to them. The inference engine used the backward chaining method as its goal-driven control technique. JAPIEST was another expert system developed by researchers to diagnose diseases and pests in tomatoes grown in hydroponic greenhouses (Lopez-Morales et al., 2008). This system also relied on rule-based reasoning. Knowledge obtained from experts was represented in the form of dependency networks.

Fuzzy Expert Systems, instead of using Boolean type logic, apply a collection of fuzzy logic's membership functions and rules to return a conclusion. A Fuzzy Expert System was developed to diagnose disease in the Chickpea plant (Dubey et al. 2014). Kolhe et al. (2011) presented a Fuzzy Expert System for diagnosis of diseases in Soybean in India. It applied a new approach to rule-based fuzzy logic called rule promotion using both forward and backward chaining.

Bayesian Networks have been used to create a plant disease diagnosis system that both actively and dynamically performs the diagnostic process (Zhu et al., 2013). The authors applied a con-

cept called 'active symptom selection' which uses only the symptoms that are relevant for the active diagnosis at a point in time. The system used the Bayesian network together with a Markov Blanket to determine the symptoms most relevant in the diagnostic process.

Camargo et al. use multi-class Support Vector Machines to classify a crop disorder based on attribute values supplied by the user (Camargo et al., 2012). For computing a diagnosis, the attribute data is mapped to a feature vector and the mc-SVM prediction model is applied to the feature vector resulting in a probability with which the feature vector belongs to the class.

Our approach combines several different multi-criteria decision making techniques to dynamically put forward questions to the farmers and reason through their responses returning a diagnosis for the suspected disease.

## 3. Disease modeling

In the approach presented in this paper, all information about each disease is brought together using the concept of a disease model where a model fundamentally contains all the characteristics of one disease. Each model consists of many characteristics where a characteristic in the simplest terms describes something about a model, for example, spot color. Characteristics may be grouped into categories/types as detailed in Table 2.

Therefore, for each disease there exists a Disease Model  $d_i$  that belongs to the set containing all models  $D$  (Eq. (1)).

$$D = \{d_1, d_2, d_3, \dots, d_n\} \quad (1)$$

All characteristics belong to a pre-defined set  $C$  which holds all characteristics that any member of  $D$  can contain. Having this fixed set of characteristics ensures that there would be one standard, consistent set of vocabulary that would be used during the diagnostic process, avoiding any confusion in future and reducing the likelihood that there would be multiple versions or values of characteristics that essentially have the same meaning. Therefore, letting  $m$  be the total number of distinct characteristics we have (Eq. (2)):

$$C = \{c_1, c_2, c_3, \dots, c_m\} \quad (2)$$

Models have a many-to-many relationship with characteristics which means that a disease model can contain one or more characteristics and vice versa (Fig. 1). This relationship is essential in the diagnosis process as it relies on the fact that a characteristic is associated with more than one model.

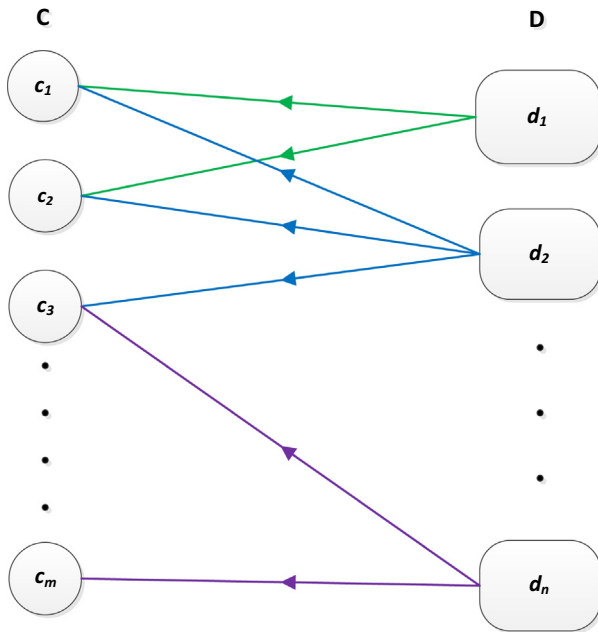
There is one somewhat special type of disease model that is also needed. This is a model that represents a healthy plant. What distinguishes these models from the others is that they have characteristics linked to them that relate to a healthy plant. For each plant type that exists, there will be one instance of this model. This

**Table 1**  
Pest and disease diagnostic systems.

Analytical technique	Plant	Pest/disease	Input	Comments
Fuzzy logic Sannakki et al. (2013)	Pomegranate	Bacterial blight, anthracnose, wilt complex	Images	Diagnosis limited to one plant
Artificial neural network AI Bashish et al. (2011)	Plants (general)	Leaf diseases		Only 5 types of leaf disease diagnosed
Fuzzy expert system Dubey et al. (2014)	Chickpea	Pests and diseases (limited to plants observed)	Text	Limited to the chickpea plant
Expert system Dewanto and Lukas (2014)	Fruit plants			Expert system trained by programmers
Bayesian networks with incremental learning Zhu et al. (2013)	Food crops			Bayesian networks used
Support vector machines Camargo et al. (2012)	Any food crop	Pest disorders		Multi-class SVM using disease attribute vector
AgriDiagnose expert decision making system; AHP and S-SAW	Any food crop	Unlimited		Easily extendable, limited only by the data that has been entered by Plant Pathologists

**Table 2**  
Characteristic types.

Characteristic type	Examples
History	Plant age, age of leaf showing damage, time of day
Symptom	Spot appearance, spot color, spot shape, spot size, spot location, type of leaf damage
Sign	Environmental condition, climatic condition, soil moisture, temperature



**Fig. 1.** Disease models and characteristics relationship.

concept was needed to allow the diagnostic process to detect when the symptoms provided by the farmer may not be the result of a disease but where the infection is either negligible, where treatment may not be needed, or if the plant is indeed healthy.

The key to the decision making process in diagnosis is the calculation of weights or probability measures in setting up the characteristics. Two types of weights are determined for the characteristics. The first weighting scheme, which we call Characteristics Relative Weights, is a probability value that ranks the importance of that characteristic against all other characteristics as a predictor of the presence of that disease in the plant. For example, 'spots on leaves' may have a higher weight than 'stem color' for a given disease model. In the diagnosis, questions related to characteristics with the highest weights are asked first. The second type of weights, which we call Scales, are associated with the values for each characteristic of a given disease. For example, the 'type of leave damage' characteristic may have values {wilting = 100, shot holes = 25, dropping = 75, curling = 10}. These are used in the calculation of the Utility value of a disease model and in the decision making process. We will describe both of these weighting schemes in the subsections that follow.

### 3.1. Characteristics relative weights

For each disease model, each characteristic is assigned a weight, that is, a value that represents its importance in relation to all other characteristics. These weights are calculated using a customized version of the AHP method (Saaty, 1977) according to the following steps:

- (1) Using Saaty's scale given in Table 3, the Pathologist will rank the characteristics pair wise; ranking each characteristic against every other as shown in Table 4. This scale is particularly useful when comparing terms that do not always have the same dimensions; for example, when comparing spot color, which is qualitative, with plant age, which is quantitative and measured in weeks or months.

The ranking will be represented in an  $m \times m$  matrix  $A$ , where  $m$  is the total number of characteristics considered for the disease; each value,  $a_{ij}$ , in the matrix represents the importance of the  $i$ th characteristic relative to the  $j$ th characteristic as a predictor of the disease. If two characteristics have equal importance, then  $a_{ij}$  is set to 1. Table 4 shows this matrix for disease  $D_1$  with  $m = 8$ .

- (2) The normalized pairwise matrix  $A_{norm}$  is then built where the values will be determined using Equation (3) shown below.

For each value in  $A$ ,

$$\bar{a}_{ij} = \frac{a_{ij}}{\sum_{k=1}^m a_{kj}}, \quad (3)$$

where the denominator is the sum of the column values.

- (3) The Characteristic Weight Vector  $w$  would then be determined. This vector is also termed the priority vector or eigenvector and is found by calculating the average value for each row of  $A_{norm}$  (Eq. (4)).

$$w = \frac{\sum_{k=1}^m \bar{a}_{ik}}{m} \quad (4)$$

The priority vector for each characteristic is taken as the characteristics' Relative Weighting (Table 5). The sum of all Relative Characteristics Weightings for a disease model must always be equal to one.

This process is done for each disease at set up. Let us suppose there are just three diseases, Table 6 shows Relative Weights of the 8 characteristics for each disease.

### 3.2. Scales

In the disease modeling, each characteristic has several possible independent descriptors. For example, the 'type of leave damage' characteristic may have values - dropping, wilting, curling, holes. Each of these descriptors will be assigned a value on a scale of

**Table 3**  
Saaty's scale.

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective(s)
3	Weak importance	Experience and judgement slightly favor one activity over another
5	Essential or strong importance	Experience and judgement strongly favor one activity over another
7	Demonstrated importance	An activity is strongly favored and its dominance demonstrated in practice
9	Absolute importance	The evidence favoring one activity over another is one of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values between two adjacent judgements	Where compromise is needed
Reciprocals of the nonzero	If activity $i$ has one of the above nonzero numbers assigned to it when compared with activity $j$ , then $j$ has the reciprocal value when compared with $i$	

**Table 4**

Matrix A of pair wise ranking of characteristics.

Disease Model D <sub>1</sub>	Characteristic	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>	C <sub>8</sub>
Characteristic		Spot location	Spot colour	Spot Appearance	Type of leaf damage	Smell	Plant age	Soil moisture	Humidity
C <sub>1</sub>	Spot location	1	5	3	2	4	4	5	4
C <sub>2</sub>	Spot colour	1/5	1	1	1	1	1	1	1
C <sub>3</sub>	Spot Appearance	1/3	1	1	1	1	1	1	1
C <sub>4</sub>	Type of leaf damage	1/2	1	1	1	1	1	1	5
C <sub>5</sub>	Smell	1/4	1	1	1	1	1	1	1
C <sub>6</sub>	Plant age	1/4	1	1	1	1	1	1/9	1/2
C <sub>7</sub>	Soil moisture	1/5	1	1	1	1	9	1	1
C <sub>8</sub>	Humidity	1/4	1	1	1/5	1	2	1	1

**Table 5**Calculated relative weights for each characteristic for disease D<sub>1</sub>.

Disease Model D <sub>1</sub>	Characteristic	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	C <sub>4</sub>	C <sub>5</sub>	C <sub>6</sub>	C <sub>7</sub>	C <sub>8</sub>
Characteristic		Spot location	Spot colour	Spot Appearance	Type of leaf damage	Smell	Plant age	Soil moisture	Humidity
C <sub>1</sub>	Spot location	0.335	0.417	0.300	0.244	0.363	0.200	0.450	0.276
C <sub>2</sub>	Spot colour	0.067	0.083	0.100	0.122	0.091	0.050	0.090	0.069
C <sub>3</sub>	Spot Appearance	0.112	0.083	0.100	0.122	0.091	0.050	0.090	0.069
C <sub>4</sub>	Type of leaf damage	0.168	0.083	0.100	0.122	0.091	0.050	0.090	0.345
C <sub>5</sub>	Smell	0.084	0.083	0.100	0.122	0.091	0.050	0.090	0.069
C <sub>6</sub>	Plant age	0.084	0.083	0.100	0.122	0.091	0.050	0.090	0.069
C <sub>7</sub>	Soil moisture	0.067	0.083	0.100	0.122	0.091	0.450	0.090	0.069
C <sub>8</sub>	Humidity	0.084	0.083	0.100	0.024	0.091	0.100	0.090	0.069
	Relative weights	0.323	0.084	0.090	0.131	0.086	0.072	0.134	0.080

**Table 6**

Calculated relative weights for each characteristic for all diseases.

Disease models	Characteristics	C <sub>1</sub> Spot location	C <sub>2</sub> Spot color	C <sub>3</sub> Spot Appearance	C <sub>4</sub> Type of leaf damage	C <sub>5</sub> Smell	C <sub>6</sub> Plant age	C <sub>7</sub> Soil moisture	C <sub>8</sub> Humidity
D <sub>1</sub>	Blythe	0.323	0.084	0.090	0.131	0.086	0.072	0.134	0.080
D <sub>2</sub>	Cassava thrips	0.060	0.012	0.234	0.100	0.451	0.050	0.090	0.003
D <sub>3</sub>	Broom's disease	0.110	0.078	0.100	0.234	0.095	0.050	0.100	0.233

1–100 which represents how frequently that value is associated with the characteristic. For example, for Blythe disease, 'dropping' may be assigned a value of 75 as the Pathologist deems that 75% of the time leaves would drop for that disease (therefore meaning that 25% of the time leaves would not drop for that disease) (Table 7). Wilting may be assigned 100% as leaves will always wilt for this disease. These values are not calculated by the system; they are subjective input values entered into the model by the Pathologist based on his expert knowledge.

For each characteristic, we calculate the utility for each descriptor using the scale values entered by the Pathologist and according to the Eqs. (5) and (6). This transforms the decision matrix to a normalized matrix (Table 8) in which the normalized values are in the range (−1...1).

$$z_{ij} = \frac{2(x_{ij} - b_j)}{a_j - b_j} - 1, \quad i = 1, \dots, n; \quad j \in \Omega_b \quad (5)$$

$$z_{ij} = 1 - \frac{2(x_{ij} - b_j)}{a_j - b_j}, \quad i = 1, \dots, n; \quad j \in \Omega_c \quad (6)$$

where  $z_{ij}$  - normalized characteristic values whose values now range from −1 to 1,  $a_j$  -  $\max(x_{ij})$  for characteristic  $j$ ;  $b_j$  -  $\min(x_{ij})$  for characteristic  $j$ ,  $\Omega_b$  - set of benefit characteristics, i.e. characteristics for which the higher the scale value the more likely that disease is the outcome,  $\Omega_c$  - set of cost characteristics, i.e. characteristics for which the higher the scale value the less likely that disease is the outcome.

We now give one other example of a normalized decision matrix (Table 9); this will be used later in the diagnosis.

**Table 7**

Scale values for the "Type of Leaf damage" characteristic.

Disease	Dropping	Wilting	Curling	Holes
Blythe	75	100	10	15
Cassava thrips	60	80	50	55
Broom's disease	20	20	50	40

**Table 8**

Utility matrix showing normalized Scale values for the "Type of Leaf damage" characteristic.

Disease	Dropping	Wilting	Curling	Holes
Blythe	0.44	1	−1	−0.89
Cassava thrips	−0.33	1	−1	−0.67
Broom's disease	−1	−1	1	0.33

**Table 9**

Normalized scale values for "Spot Color" characteristic.

Disease	Red	Purple	Blue
Blythe	0.6	1	−1
Cassava thrips	1	0.4	−1
Broom's disease	0.8	−1	1

### 3.3. Implementing the disease modeling

The disease modeling is implemented in the Pathology Tool. This implementation is initiated by plant pathologists and can be summarized in the series of steps below.

For each plant type and variety:

1. All known diseases are entered.
2. All characteristics that are used by human pathologists to determine plant disease are entered in the system.
  - a. The pathologist must do the pairwise ranking for each pair of characteristics for each disease. The system will calculate the Relative Weights of the characteristics based on the values entered by the pathologist using the AHP method outlined in Section 3.1. The data entry interface is illustrated in Fig. 2.
  - b. For each characteristic, the pathologist must phrase a question which will be posed to the farmer, during disease diagnosis, to capture the information about that characteristic.
  - c. The pathologist must then provide the possible answers to the questions and assign a value on a scale of 1–100 of how frequently that characteristic value occurs in the disease. The system calculates the normalized scale values for the decision matrix as outlined in Section 3.2.

Once the models are carefully configured by pathologists, the value of  $n$  (the total number of disease models) will be set. At this point the configuration is complete and the system is ready for use by farmers.

## 4. Disease diagnosis

Farmers interact with the system by answering a series of questions which describe the appearance of the affected plant. Characteristic Relative Weights and the Sensitive Simple Additive Weighting (S-SAW) are used to determine which questions are asked and how disease diagnosis is determined.

### 4.1. How are questions posed?

The system starts with filter questions on the plant type and variety. The farmer's responses allow the system to narrow down the search to only those disease models associated with that plant. The system continues asking questions associated with characteristics and getting responses from farmers about the values of those characteristics. These questions are first ordered by plant history then symptoms and signs observed (Table 2).

Recall that for each disease, Characteristic Relative Weights exist (Tables 5 and 6). This means that whenever a disease is being considered as the likely diagnosis, the questions posed to the farmer are given according to the weight of that characteristic in the model. Questions on characteristics with the highest relative weight are asked first. Using this order it is expected that the process can eliminate or rule out several diseases at an early stage narrowing down the lists of possibilities and therefore reducing the



Fig. 2. Data entry interface for characteristics relative weights.

number of questions that need to be answered by the user as diagnosis continues.

#### 4.2. How is decision making simulated?

The method used in the decision making process is the Sensitive Simple Additive Weighting (S-SAW) (Goodridge, 2016). In this method, the normalized scale decision matrix is used to calculate the Utility value of each disease model to determine the preference for a model as the most likely diagnosis. At each stage the disease model that has the highest Utility value is the most probable decision output. The Utility value of a disease model is calculated as follows in Eq. (7):

$$U_d(x_1, \dots, x_p) = \sum_{i=1}^p w_i U_i(x_i) \quad \forall d \in D \quad (7)$$

where  $\mathbf{x}$  – The value of input characteristic entered by farmer.  $\mathbf{p}$  – Total number of input characteristics entered by the farmer.  $\mathbf{w}_i$  – Weighting value of characteristic  $i$  (Probability that the characteristic is part of the disease model).  $\mathbf{U}_i$  – Utility value of some characteristic  $i$ .  $\mathbf{U}_d$  – Utility value of a disease model.

When  $U_d$  is found for each member of  $D$ , the model that has the highest utility value will be put at the top of the stack of disease possibilities, and at that point this disease model is the most probable diagnosis.

For example, in the diagnosis, suppose that the system asks the farmer two questions, “What is the type of leaf damage?” and “What is the color of the spot?” and the farmer responds with “Dropping” as the leaf damage and “Red” as the spot color. The system calculates the Utility value for each disease in the stack based on these two input values using Eq. (7) with  $p = 2$ , therefore in Eq. (8),

$$U(\text{disease}) = (w_1 * u_1) + (w_2 * u_2) \quad (8)$$

where  $w_1$  = the relative weight of ‘type of leaf damage’,  $w_2$  = the relative weight of ‘Spot color’,  $u_1$  = the normalized scale value of ‘dropping’,  $u_2$  = the normalized scale value of ‘red’.

This means,

$$U(\text{Blythe}) = (0.131 * 0.44) + (0.084 * 0.6) = 0.108.$$

$$U(\text{Thrips}) = (0.1 * -0.33) + (0.012 * 1) = -0.021.$$

$$U(\text{Broom}) = (0.234 * -1) + (0.078 * 0.8) = -0.1716.$$

Blythe disease is then placed at the top of the stack for continued questioning.

The process then continues to ask the farmer other questions related to the characteristics of the disease model with the highest utility value. This way, only relevant questions are asked as at that point in time that disease is the likeliest disease and the process will attempt to either confirm or eliminate that disease as the diagnosis. The order in which the questions are asked is important. The question linked to the characteristic with the highest Relative weighting will be asked first. As questions continue to be answered the search space becomes narrower as the system would have more information. This process also ensures that no question is asked twice as once information about a characteristic is received from the farmer there is no need to ask about that characteristic again. Throughout the process different disease models could be placed at the top of the stack and some models are also eliminated as the decision process progresses.

## 5. Experimental results

The primary purpose of the experimental simulations is to determine how well the system was able to act as a human Pathologist and correctly diagnose diseased plants. The metrics used to evaluate the system’s diagnostic process are similar to those typically used to evaluate clinical tests (Stojanovic et al., 2014). Four measures were used: sensitivity, specificity, precision and accuracy. We first explain the terms used in these metrics in Table 10 then give the definitions in Equations (9)–(12).

#### Eq. (9) - Sensitivity

This measures the probability of a positive test given that a plant is diseased. Essentially, it measures the diagnostic process’s

ability to correctly detect that there is some problem with the plant.

$$\text{Sensitivity} = \frac{\text{Number of TP}}{\text{Number of TP} + \text{Number of FN}} \quad (9)$$

Eq. (10) - Specificity

This measures the system's ability to correctly return 'not diseased' if a plant is not infected. This is a probability measure of a negative result, if the plant indeed has no disease.

$$\text{Specificity} = \frac{\text{Number of TN}}{\text{Number of TN} + \text{Number of FP}} \quad (10)$$

Eq. (11) - Precision

This measures the system's ability to correctly return 'healthy' when the plant is healthy or to detect a disease when the plant is diseased.

$$\text{Precision} = \frac{\text{Number of TP} + \text{Number of TN}}{\text{Number of Tests}} \quad (11)$$

Eq. (12) - Accuracy

This measures the system's ability to correctly diagnose the disease when a diseased plant is presented.

$$\text{Accuracy} = \frac{\text{Number of correct diagnoses}}{\text{Number of TP}} * 100 \quad (12)$$

For the experiment, the Pathology tool was used to configure the system for four crops – Cassava, Tomato, Cocoa, Hot Peppers – and a total of fourteen disease models with 20 characteristics. To conduct the evaluation, 54 tests were performed. These tests are independent and were used to validate the system's results (see Table 11).

Pathologists were asked to simulate various scenarios that would typically lead them to their diagnosis. The simulations involved identifying the characteristics that would lead them to specific diagnoses. Once they formulated the scenario, they were requested to use the mobile app for farmers, and follow through the question and answer session. The application would then return a result. The result was analyzed by the Pathologist to determine if the diagnosis produced by the system was correct. The results for our metrics are given in Table 12.

The diagnostic process was found to have a perfect (100%) sensitivity because in all instances where the plant was indeed diseased, it was able to diagnose that there was a disease.

In terms of the specificity, 80% of the instances, it correctly determined that based on the symptoms, the plant was healthy. It must be noted that the Pathologist tried as much as possible to simulate a 'healthy' scenario; however, they explained that sometimes the symptoms presented could indicate that a disease is present though at a very early stage. In such cases, treatment is not required as the plant is still considered to be relatively healthy.

The precision metric was excellent as the process was able to determine if a plant was diseased or healthy almost perfectly (98.15%).

**Table 10**  
Terms associated with metrics used.

Name	Description
True Positive (TP)	The plant is infected and the diagnosis returns a disease
False Positive (FP)	The plant is not infected but the diagnosis returns a disease
True Negative (TN)	The plant is healthy and the diagnosis returns this as a result
False Negative (FN)	The plant is infected but the diagnosis returns that the plant is healthy

**Table 11**  
List of crop diseases.

Disease model	Scientific name	Crop
Super elongation disease	<i>Elsinoe brasiliensis</i>	Cassava
Thrips	<i>Corynothrips stenopterus</i>	Cassava
Green spider mite	<i>Green spider mite</i>	Cassava
Chinch bugs	<i>Cyrtomerus bergi</i>	Cassava
Bacterial leaf spot	<i>Xanthomonas campestris</i>	Hot pepper
Shoot fly	<i>Siblua pendula</i>	Cassava
Root rot disease	<i>Phytophthora capsici</i>	Hot pepper
Black pod disease	<i>Phytophthora palmivora</i>	Cocoa
Witches broom disease	<i>Moniliophthora perniciosa</i>	Cocoa
Bacterial leaf spot	<i>Xanthomonas campestris</i>	Tomato
Anthraxnose	<i>Colletotrichum gloeosporioides</i>	Hot pepper
Stem canker	<i>Phytophthora spp.</i>	Cocoa
Southern blythe	<i>Sclerotium rolfsii</i>	Hot pepper

**Table 12**  
Simulation results.

Metric	Inputs	Result
Sensitivity	Number of TP = 49 Number of FN = 0	1.0
Specificity	Number of TN = 4 Number of FP = 1	0.8
Precision	Number of TP = 49 Number of TN = 4 Number of tests = 54	0.9815
Accuracy	Number of correct diagnoses = 47 Number of TP = 49	95.9%

Finally, the accuracy percentage was also very good as in 95.9% of the instances of diseased plants, the system was able to correctly diagnose the disease affecting the plant. This level of accuracy compares well with the results published in the literature using other techniques. For example, Camargo et al. (2012) using SVM technique, reports an accuracy that is below 90% even under the most favorable conditions and deteriorates to about 70% as missing attribute probabilities approach 40%.

## 6. Conclusion

The diagnostic process described in this paper combines components of different multi-criteria decision making (MCDM) techniques to form a dialog-based decision support system for plant disease diagnosis. This automated process can result in fast, real-time diagnoses being provided to farmers in the field allowing for early detection. The techniques have been implemented in a computer system called AgriDiagnose that comprises a web-based Pathology tool for configuring the disease models and a mobile app for farmers to interact with the system to receive a diagnosis. In addition to diagnosis, the app extends its functionality by providing treatment options and general educational information related to the suggested treatment. The experimental results from using the system show excellent performance on several metrics.

The disease modeling that we propose in this paper is very general and can be applied to any plant type and variety. This is in contrast to several other published works that are specific to certain crops or certain environmental conditions. The multi-criteria decision making techniques presented in this paper have generally been used in business and other domains. This paper represents an innovative use of well-established MCDM techniques.

From the inception of production until crops are harvested, pests and diseases threaten food security. AgriDiagnose is one of several systems developed under a larger project called AgriNeTT (AgriNeTT, 2016; Jordan et al., 2016). AgriNeTT has won global

recognition for its contribution to the UN Sustainable Development Goal #2: End hunger, achieve food security and improved nutrition, and promote sustainable agriculture.

## Acknowledgements

This project is funded by the University of the West Indies - Government of Trinidad and Tobago Research and Development Impact fund.

## References

- AgriNeTT Website: <http://sta.uwi.edu/agrinett>; AgriNeTT Wins Global Recognition – Recognized for Its Contribution to SDGs at WSIS Forum 2016. <<http://www.itu.int/net4/wsis/stocktaking/en/Database/Search?pld=1452011099>>.
- Barbedo, J.G.A., 2016. Expert systems applied to plant disease diagnosis: survey and critical view. *IEEE Latin Am. Trans.* 14 (4).
- Bashish, D.A., Braik, M., Bani-Ahmad, S., 2011. Detection and classification of leaf diseases using k-means-based segmentation and neural-networks-based classification. *Inform. Technol. J.* 10 (2), 267–275.
- Camargo, A., Molina, J., Cadena-Torres, J., Jimenez, N., Kim, J., 2012. Intelligent systems for the assessment of crop disorders. *Comput. Electron. Agr.* 85, 1–7.
- Chin, M., Rhodes, L., Tennant, P., 2007. Outbreak of papaya virus in papaya orchards in St. Kitts. *Plant. Pathol.* 57 (4).
- Dewanto, S., Lukas, J., 2014. Expert system for diagnosis of pest and disease in fruit plants. *EPJ Web of Conferences*, vol. 68, p. 00024.
- Dubey, S., Pandey, R., Gautam, s., 2014. Development of multimedia fuzzy based diagnostic expert system for integrated disease management in chickpea. *Int. J. Sci. Modern Eng.* 2 (2), 16–20.
- Gonzalez-Andujar, J.L., Fernandez-Quintanilla, C., Izquierdo, J., Urbano, J.M., 2006. SIMCE: an expert system for seedling weed identification in cereals. *Comput. Electron. Agr.* 54 (2), 115–123.
- Goodridge, W., 2016. Sensitivity analysis using simple additive weighting method. *Int. J. Intell. Syst. Appl.* 5, 27–33.
- Hwang, C.L., Yoon, K., 1981. *Multiple Attribute Decision Making Methods and Application*. Springer, New York, NY, USA.
- Jordan, R., Eudoxie, G., Maharaj, K., Belfon, R., Bernard, M., 2016. AgriMaps: improving site-specific land management through mobile maps. *Comput. Electron. Agr.* 123, 292–296.
- Kolhe, S., Kamal, R., Harvinder, S., Gupta, G.K., 2011. An intelligent multimedia interface for fuzzy-logic based inference in crops. *Exp. Syst. Appl.* 38 (12), 14592–14601.
- Lopez-Morales, V., Lopez-Ortega, Ramos-Fernandez, J., Munoz, L., 2008. JAPIEST: an integral intelligent system for the diagnosis and control of tomatoes diseases and pests in hydroponic greenhouses. *Exp. Syst. Appl.* 35 (4), 1506–1512.
- Mansingh, G., Reichgelt, H., Bryson, K.-M.O., 2007. CPEST: an expert system for the management of pests and diseases in the Jamaican coffee industry. *Exp. Syst. Appl.* 32 (1), 184–192.
- Sannakki, S., Vijay, S., Rajpurohit, Nargund, V.B., Arunkumar, R., 2013. Disease identification and grading of pomegranate leaves using image processing and fuzzy logic. *Int. J. Food Eng.* 9 (4).
- Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. *J. Math. Psychol.* 15 (3), 234–281.
- Stojanovic, M., Apostolovic, M., Stojanovic, D., Milosevic, Z., Toplavac, A., Lakusic, V., Golubovic, M., 2014. Understanding sensitivity, specificity and predictive values. *Vojnosanit. Pregl.* 71 (11), 1062–1065.
- Zhu, Y., Dayou, L., Guifen, C., Haiyang, J., Helong, Y., 2013. Mathematical modelling for active and dynamic diagnosis of crop diseases based on Bayesian networks and incremental learning. *Math. Comput. Model.* 58 (3–4), 514–523.