Arboriculture & Urban Forestry 2023. 49(1):38–53
https://doi.org/10.48044/jauf.2023.004

ARBORICULTURE & URBAN FORESTRY
Scientific Journal of the International Society of Arboriculture

Integrated Visual Assessment Method for Infestation by Lebbek Borer (*Xystrocera globosa*) in Rain Trees of Singapore

By Derrick Nguyen Hoang Danh, Muhammad Izaan bin Istijab, Muhammad Azhari bin Mohammad Zain, Xue Huiyu, Diana Chiu, Varughese Philip, Mei Lai Yap, and Abdul Hamid bin Mohammad

Abstract. Background: In Singapore, determining the level of infestation by lebbek borer (*Xystrocera globosa*)(Olivier, 1795)[Coleoptera: Cerambycidae] is the crucial first step in control of this pestra in rain tree (*Albizia saman* [Jacq.] Merr.)[Fabales: Fabaceae]. Current assessment methods rely on symptoms such as canopy colour, defoliation, dieback, and actual estimation of borer population via counting of larvae or exit holes created by adults. Currently, there is a lack of systematic approach to integrate different tree health indicators and symptoms to quantify infestation level. This gap poses challenges in assessment of treatment efficacy as managers could not quantitatively determine whether infestation level has changed following treatment. Thus, this study aimed to develop a visual assessment method that can integrate all mentioned symptoms to quantify infestation level. Methods: We surveyed a total of 388 rain trees and used principal component analysis (PCA) to investigate the correlation between *X. globosa* infestation and different borer infestation symptoms. Borer Infestation Score (BIS) formula was developed based on the linear combinations of the statistically significant principal component. Results: Infestation level was strongly associated with bark peeling, exit holes, and proximity of bark peeling and/or exit holes to trunk base and weakly associated with defoliation, dieback, and canopy colour. Developed BIS formula generated numerical values that distinguished between noninfested and infested trees, reflected infestation level in surveyed areas and temporal progression of infestation. Conclusions: Described integrated visual assessment method can be executed quickly on field. BIS formula generates quantitative scores easy to be interpreted, tracked, and compared.

Keywords. *Albizia saman*; Borer Infestation Score; Integrated Pest Management; Integrated Visual Assessment; Rain Tree Borer Infestation; Tree Health; Urban Forestry; *Xystrocera globosa*.

INTRODUCTION

In Singapore, rain tree *Albizia saman* [Fabales: Fabaceae] is among the 10 most common tree species on the island, planted abundantly along roadsides, parks, and gardens. Despite their great adaptability to the local climate, rain trees are susceptible to various insect pests, the most destructive of which is *Xystrocera globosa* [Coleoptera: Cerambycidae]. Larvae of this species bore into the wood of branches and trunks as they feed, causing extensive damage to internal structure and disrupting nutrient transport (Beeson 1941; Suharti et al. 1994; Matsumoto et al. 2000). As the infestation progresses, trees show constant decline, which over time can manifest into visible symptoms such as bark peeling, dieback, defoliation, and development of adventitious roots.

As the first step of *X. globosa* management, assessment of infestation by this borer is essential to determine appropriate intervention measures and to monitor the effectiveness of said measures in controlling the infestation. There are 2 common approaches employed in assessment of borer infestation: (1) measurement of symptoms and actual number of borers, and (2) measurement of tree health.

Measurements of signs such as exit holes and the actual number of borers can directly estimate the population of borers in infested trees. A higher density of exit holes indicates a higher number of emerging borer adults, which means a higher number of borer larvae actively feeding on the internal wood structure of infested trees (McCullough et al. 2005; Anulewicz et al. 2007; McCullough and Siegert 2007; Pontius et al. 2008). Quantification of exit holes can be in terms of numerical categorization (Pontius et al. 2008) or direct density measurement expressed as the number of exit holes per unit area (McCullough et al. 2005; Anulewicz et al. 2007; McCullough and Siegert 2007). Actual counts of larvae and estimation of larval mortality rate can be done by cutting and splitting branches and trunks of infested trees (Mercader et al. 2013). Whole branches or entire trees can be cut to allow accurate measurement of exit holes and larval density. Although these methods can provide relatively accurate estimations of borer infestation level, branch removal can affect tree form and canopy balance.

Meanwhile, measuring tree health is based on the principle that prolonged infestation causes tree decline over time, which manifests into symptoms such as defoliation, canopy thinning, dieback, leaf yellowing, and, in extreme cases, death of the infested trees. To quantitatively capture this information, categorical scoring or percentage infestation are commonly used techniques. For instance, Smitley et al. (2008) established pictorial examples for different percentages of canopy thinning. Health classes based on canopy conditions can be operationally defined as seen in Murfitt et al. (2016). At larger scale, tree mortality rate can be used to indicate the infestation level of an area (Morin et al. 2017). However, a decline in tree health indicated by canopy thinning, dieback, or leaf yellowing is only observed when the infestation has become severe. Delayed manifestations of the mentioned symptoms imply that these variables are generally not sensitive to the progression of infestation in early stages. Furthermore, besides *X. globosa*, rain trees can be infested with multiple pests (e.g., defoliator moths such as *Pandesma quenavadi* and root rot fungi *Ganoderma* sp.) which can also induce decline symptoms.

Currently in Singapore, park and streetscape managers rely on signs of exit holes and adult *X. globosa*, as well as symptoms such as bark peeling, dieback, leaf yellowing, and canopy thinning, to assess rain tree infestation. The major gap in the current protocol is the lack of a systematic approach to integrate all of this information into a unified calculation to quantitatively assess the infestation level of rain trees. Consequently, managers face challenges in determining appropriate control measures and ascertaining the effectiveness of such measures.

To bridge this gap, the present study aimed to develop a mathematical expression that integrates visual signs and symptoms associated with *X. globosa* infestation (e.g., bark peeling, exit holes, dieback, leaf yellowing, and canopy thinning) to give a unified scoring that (1) distinguishes between infested and noninfested trees, (2) reflects the general level of infestation of a tree or trees in the area, and (3) allows monitoring of the temporal progression of *X. globosa* infestation. To achieve these objectives, tree surveys were conducted on 388 rain trees located in 5 different locations in Singapore: Siglap Link (SL), Bedok South Avenue 1 (BSA), East Coast Park Service Road (ECP), Penjuru Road (PR), and Geylang East Central (GEC) from May 2020 to July 2020. For each tree, we recorded a quantitative estimate of defoliation, dieback, exit holes, bark peeling, and proximity of observed exit holes and/or bark peeling to the trunk base. Correlation between these variables and infestation was analyzed using principal component analysis (PCA), and a formula to estimate infestation level from the linear combinations of the most informative principal components was derived.

MATERIALS AND METHODS

Tree Survey

A total of 388 rain trees along roadsides were assessed at 5 different locations from May 2020 to July 2020 (Table 1). Locations were selected based on alerts of an incursion of borer infestation. Of all assessed trees, 48 infested trees along East Coast Park Service Road (ECP) with a Borer Infestation Score (BIS)(the definition and calculation of which will be explained in a later part of the paper) within the interquartile range of BIS distribution were selected. These infested trees were reassessed in October 2020 to investigate the progression of infestation. These 48 trees were not treated with any chemicals or pruned from July 2020 to October 2020.

Recorded Variables

A total of 7 variables were recorded (Table 2). As discussed earlier, larval stages of *X. globosa* bore deep inside branches and/or trunks of rain trees and cannot be visually detected. Therefore, trees showing no signs

Table 1. List of surveyed areas, their corresponding coordinates, and number of trees assessed.

Serial number	Location	Coordinates	Number of trees assessed
1	Bedok South Avenue 1 (BSA)	1.3177878, 103.9307563	67
2	East Coast Park Service Road (ECP)	1.3073860, 103.9278246	144
3	Geylang East Central (GEC)	1.3190039, 103.8846806	84
4	Penjuru Road (PR)	1.3143540, 103.7328387	56
5	Siglap Link (SL)	1.3070307, 103.9253042	37
Total			388

Table 2. List of recorded variables during tree survey. Variables 2, 3, and 4 were recorded as ratios to account for different canopy sizes in rain trees. Variables 5, 6, and 7 were recorded as numerical values based on predefined categories for quick estimation (extent defoliation) or to provide basis for quantitative calculation (canopy colour and lowest position of bark peeling or exit holes).

Serial number	Variable	Unit
1	Infestation status	Infested or noninfested
2	Number of major branches with bark peeling symptom:Number of major branches	Ratio
3	Number of major branches with exit hole symptom:Number of major branches	Ratio
4	Number of major branches with dieback symptom:Number of major branches	Ratio
5	Canopy colour	(1) Green or (2) chlorotic
6	Extent of defoliation	(1) 0% to 20% (2) 21% to 40% (3) 41% to 60% (4) 61% to 80% (5) 81% to 100%
7	Lowest position of bark peeling or exit holes	(1–6) Refer to Figure 1 (7) No observable bark peeling or exit holes

of bark peeling and/or exit holes may still be infested. To overcome this limitation, the infestation status of each tree was determined based on the inspection history of the tree. That means trees with an inspection record indicating no symptoms of infestation for at least 1 year before the survey and 1 year after the survey were considered to be noninfested. As described by Matsumoto et al. (2000), the duration of the immature period (egg, larvae, and pupae) for *X. globosa* lasts 112.5 ± 6.3 days in males and 104.6 ± 10.6 days in females. In Singapore, rain trees infested with *X. globosa* show symptoms within 1 year. Although destructive sampling is the desired method to determine infestation status of trees, this sampling method was not possible in Singapore as surveyed rain trees were planted along roadsides. Removal or pruning of healthy trees is strictly controlled to minimize negative impacts on tree health and city landscape. Nonetheless, considering *X. globosa* biology and inspection records (1 year before and after survey date), the risk of a false negative due to asymptomatic infestation could be completely mitigated. This information was not known to the assessors who did the tree surveys to avoid confirmational bias.

The number of major branches per tree was determined by moving an imaginary horizontal plane from the trunk base up to the canopy. At the height where branches split into at least 4, the number was recorded as the number of major branches. Symptoms such as bark peeling, exit holes, and dieback were determined based on the visual detection of these symptoms. The extent of defoliation was visually estimated and assigned a score ranging from 1 to 5: (1) 0% to 20% defoliation, (2) 21% to 40% defoliation, (3) 41% to 60% defoliation, (4) 61% to 80% defoliation, and (5) 81% to 100% defoliation. Canopy colour was scored as (1) green or (2) chlorotic.

Proximity of bark peeling and/or exit holes to trunk base was a new variable that was investigated in this study. It stemmed from the consistently observed trend of *X. globosa* infestation to start in higher branches and progress downward to the trunk base. When symptoms such as bark peeling and exit holes were observed nearer to the trunk base, the infestation level was more severe. To quantitatively capture this information, rain trees were broken down into 6 regions based on distance to trunk base (Figure 1). During tree assessment, each observation of bark peeling and/or exit holes was assigned a number based on the regions described, and the lowest number was retained for final recording. Trees with no exit holes or bark peeling were assigned with a score of 7.

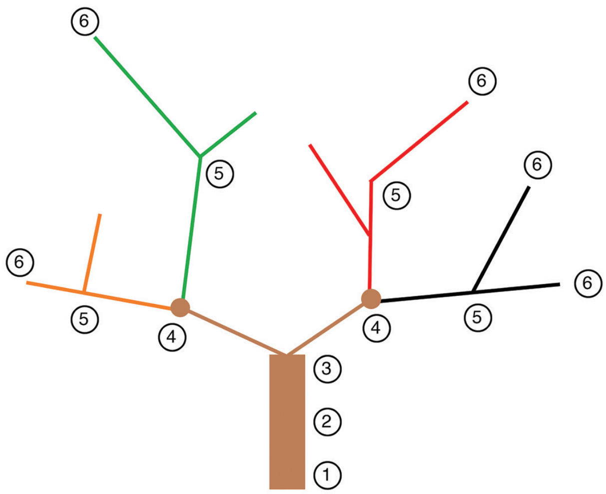

Figure 1. Schematic presentation of a rain tree and its corresponding positions with regards to proximity to trunk base. (1) Base of trunk; (2) middle of trunk; (3) top of trunk; (4) base of major branch; (5) middle of major branch; (6) tip of major branch. Different colours were used to denote different major branches.

Index Construction

Checking of Linearity, Normality, and Outliers

All statistical analyses in this study were performed using statistical software R v4.1.1 (The R Foundation for Statistical Computing, Vienna, Austria). Data linearity was assessed using scatter plots between variable pairs using the *pairs* function. To test for multivariate normality, a Shapiro-Wilk multivariate test was conducted using the *mshapiro.test* function from the *mvnormtest* package. Multivariate outliers in the data were detected based on Mahalanobis distance using the *mahalanobis_distance* function from the *rstatix* package.

Deviation from Normality and Presence of Outliers

Principal component analysis (PCA) is a linear orthogonal data transformation technique that uses Pearson correlation coefficients to find uncorrelated rotation axes (principal components)(PCs) capturing maximal variance in the data (Bro and Smilde 2014). For multivariate normally distributed data, independence between components can be guaranteed when there is zero correlation between them. For multivariate non-Gaussian data, PCA components are uncorrelated but not independent (Kim and Kim 2012). However, deviation from normality does not invalidate the application of PCA, especially as a descriptive tool, on non-Gaussian data (Jolliffe and Cadima 2016). Wang and Du (2000) showed that PCA conducted on both Gaussian and non-Gaussian data sets yielded useful results. Nonetheless, lack of independence between PCA components could potentially reduce the technique robustness due to contamination between components (Kim and Kim 2012). Therefore, for non-Gaussian data, independent component analysis (ICA) could be conducted to compare with PCA to study the effect of nonindependence on the interpretation of PCA results (Kim and Kim 2012). In this study, ICA was conducted using the *fastICA* function from the *fastICA* package.

To study the effects of outliers, PCA was performed on the data set with and without outliers. The PCA results were compared.

Training of Data with PCA and Visualization

The original data set was first split into 2 sets of data, each containing only infested trees or noninfested trees. Each of these sets was then partitioned into 2 nonoverlapping subsets, (a) 75% of original data, and

(b) 25% of original data via random selection. Two subsets (a) were combined to form the training data set, while two subsets (b) were combined to form the testing data set. For the training data set, PCA with scaling was done on recorded variables except for infestation status using the in-built *prcomp* function. Location of trees was not included in PCA. Visualization of dimensions and calculation of their corresponding eigenvalues and variances were done using the *fviz_eig* function from the *factoextra* package. Using the *pca3d* function from the *pca3d* package, 3D visualization of PCA score plots for the first 3 components was performed.

Dimension Selection and Stopping Rules

To determine the number of statistically significant principal components, we employed functions within the *PCDimension* package to perform a Pseudo-*F* ratio test (Ter Braak 1990), an eigenvalues *P*-value test (Ter Braak 1988), a broken stick statistical test (Barton and David 1956), and an Auer-Gervini method (Auer and Gervini 2008). Pseudo-*F* ratio and eigenvalues *P*-value tests were applied using the *rndLambdaF* function for 1,000 iterations at 0.05 significance level. The *bsDimension* function was used to perform the broken stick test. The *AuerGervini* function was used to perform the Auer-Gervini method with twicemean, kmean, spectral clustering, and changepoint criteria.

Borer Infestation Score (BIS)

Statistically significant principal components were identified. Linear combinations of variables along these principal components were determined to derive the formula to calculate the scores for each tree, which are defined as their Borer Infestation Score (BIS). Distribution of BIS values for all trees in the training data set were calculated and plotted using the *ggdensity* function from the *ggpubr* package for infested and noninfested trees. The intersection between BIS density plots of infested and noninfested trees was calculated using the in-built *intersect* function. The value of this intersection was used as the classification threshold to distinguish between infested and noninfested trees.

Index Validation

Classification Accuracy Between Infested and Noninfested Trees

The developed BIS formula was applied to calculate BIS values for trees in the test data set. Using the classification threshold determined prior, trees were classified as either infested or noninfested. The class predictions were tallied against actual classification of trees in the test data set. A confusion matrix was constructed, and classification accuracy was determined based on the number of correct classifications.

Ranking Accuracy for Infestation Level of Surveyed Areas

Based on feedback from managers of the surveyed areas, the levels of infestation were considered low for GEC and PR, moderate for SL, and high for ECP and BSA. Using the developed BIS formula, average BIS values for each area were obtained, ranked, and compared to the general level of infestation based on feedback from area managers.

Sensitivity to Temporal Progression of Infestation

BIS values for the selected 48 infested trees along the ECP area were calculated for July and October. Violin plots for BIS values in July 2020 and October 2020 were generated using the *ggplot* function from the *ggbiplot* package. Temporal changes in BIS mean value between July and October were tested using the built-in *aov* function.

RESULTS

Index Construction

Linearity, Normality, and Outliers

As seen from 15 scatter plots between variable pairs (Figure 2), there was no nonlinear correlation detectable between any variable pair.

The Shapiro-Wilk multivariate normality test had a P-value = 3.752×10^{-12} leading to rejection of null hypothesis. The data set did not have normal distribution. Consequently, ICA was performed and the result was compared with PCA. As seen from Figure 3a, the 3D plot of PCA based on the top 3 principal components indicated that infested and noninfested rain trees were clearly separated along PC1 with most of the data variation also explained along this component. Meanwhile, the 3D plot of ICA (Figure 3b) also found similar distribution of points since infested and noninfested trees clearly separated along ICA1. Since ICA was designed to separate independent components from multivariate non-Gaussian data, more clusters detected in the 3D ICA plot were expected. For both 3D plots obtained from PCA and ICA, PC2, PC3, ICA2, and ICA3 did not give clear separation between infested and noninfested trees. Thus, with respect to the ability to distinguish between infested and noninfested trees and account for variation in the data set, PCA and ICA yielded similar results, indicating that deviation from normality did not considerably affect the interpretation of PCA components for the data set used in this study.

The outlier detection method based on Mahalanobis distance returned 33 out of 388 observations that

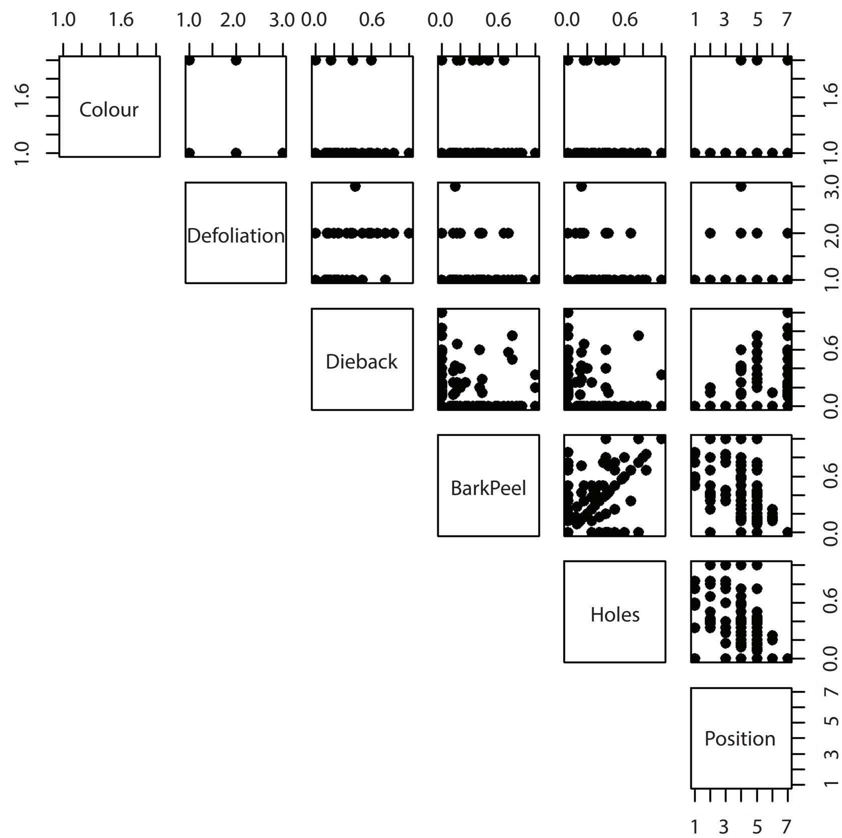

Figure 2. Scatter plot matrix for pairs of 6 recorded variables (canopy colour, defoliation, dieback, bark peeling, exit holes, and lowest position of bark peeling and/or exit holes) to test linearity of data. Scatter plot matrix was obtained using *pairs* function in R (The R Foundation for Statistical Computing, Vienna, Austria).

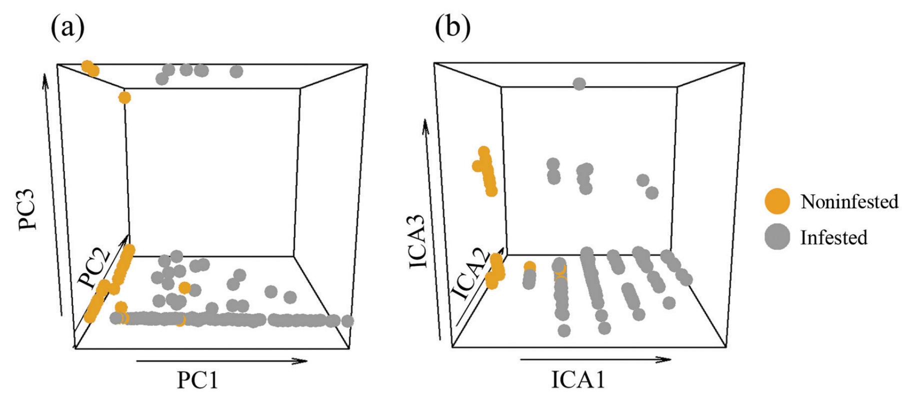

Figure 3. The 3D score plots for visualization of (a) principal component analysis (PCA) and (b) independent component analysis (ICA) to study natural data clusters and investigate the effect of non-normality on interpretation of principal component (PC). PC scores and independent component scores were obtained from respective PCA and ICA objects in R. The obtained scores formed xyz coordinates of points plotted using *scatterplot3d* function from *scatterplot3d* package in R (The R Foundation for Statistical Computing, Vienna, Austria).

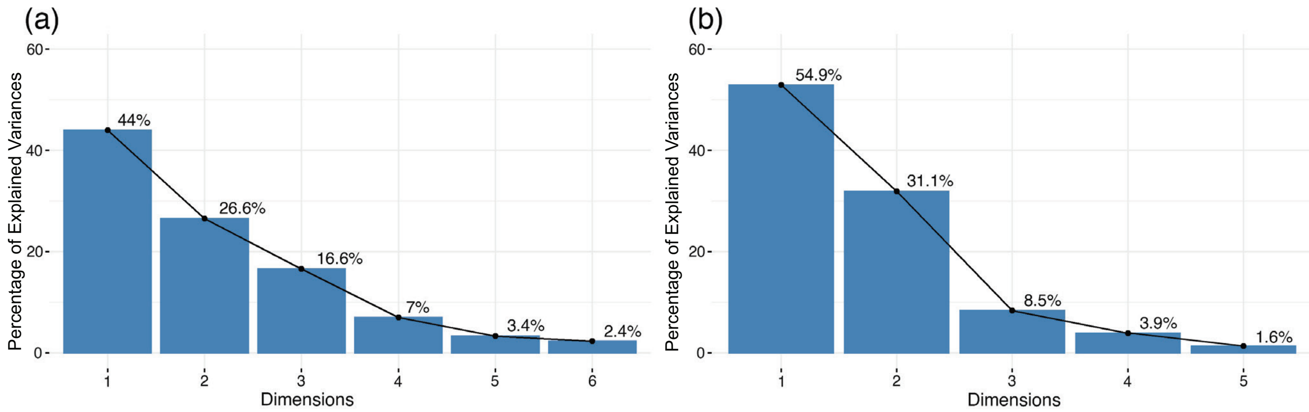

Figure 4. Scree plots for percentage of explained variance when PCA was performed on (a) original data set and (b) data set with removed outliers to study the effect of outliers on the interpretation of principal component (PC) also referred to as "Dimensions." Plot was obtained using *fviz_eig* function from *factoextra* package in R (The R Foundation for Statistical Computing, Vienna, Austria). Note: when outliers were removed, all the remaining observations had the same value for colour variable (colour score = 1) leading to reduction in number of possible PCs (or "Dimensions") from 6 to 5.

were statistically considered outliers. In 388 observations, there were 14 observations of chlorotic canopy colour (colour score = 2) and 21 observations of defoliation higher than 20% (defoliation score > 1). As seen from the scores of different recorded variables for these 33 outliers (Table S1), these 33 outliers included all 14 observations of chlorotic canopy colour and 11 out of 21 observations of defoliation higher than 20%. Scores for other variables such as bark peeling, exit holes, and dieback were generally high as compared to the total average of the entire data set. PCA results derived from the full data set (all 388 observations) and the data set without outliers (355 observations) showed increased percentage of explained variance in PC1 and PC2 when outliers were removed (Figure 4). Such an increase was partly because removing the outliers also removed all chlorotic canopy colour observations in the data set, resulting in no variation in colour variables and a subsequent reduction in the number of possible PCs. The 3D plots of PCA with and without outliers yielded similar outcomes, as infested and noninfested trees remained clearly separated along PC1 (Figure 5). In general, we found that the presence of outliers in our data set did not significantly change the interpretation of PCs. Thus, we retained outliers in our data set for subsequent analyses.

Principal Component Analysis (PCA)

Among 6 principal components (PCs), the first, second, and third PCs had eigenvalues more than or equal to 1 (Figure 6a). Together, these 3 components accounted for 86.3% of variability (Figure 6b). Table 3 showed that all stopping rules except for the Pseudo-*F* ratio found one statistically significant PC. Poor performance of the Pseudo-*F* ratio stopping rule aligned with findings by Wang et al. (2018), when large matrices were used as input. The PCA score plot was constructed based on the first 3 components (Figure 7). Healthy and infested rain trees were clearly separated along the first principal component (PC1). For infested rain trees, most variations in PCA scores were also explained by PC1. Along this principal component, symptoms such as exit holes, bark peeling, and lowest position of observed bark peeling and/or exit holes had the highest absolute values of rotations (Table 4). Therefore, PC1 was the sole component used for subsequent construction of the infestation index.

Borer Infestation Score (BIS)

The values of scale, rotation, and center of the PCA analysis were obtained to reconstruct a formula to calculate the principal score along PC1 which is defined as the Borer Infestation Score (BIS) from this point onwards. This formula was further simplified to give the following linear equation,

$$\mathrm{BIS} = 1.075 - 0.085 \times C - 0.076 \times D1 - 0.648 \times D2 + 2.075 \times B + 2.155 \times E - 0.333 \times P$$

where C = colour of canopy; $D1$ = defoliation; $D2$ = dieback; B = bark peeling; E = exit holes; and P = lowest position of bark peeling and/or exit holes.

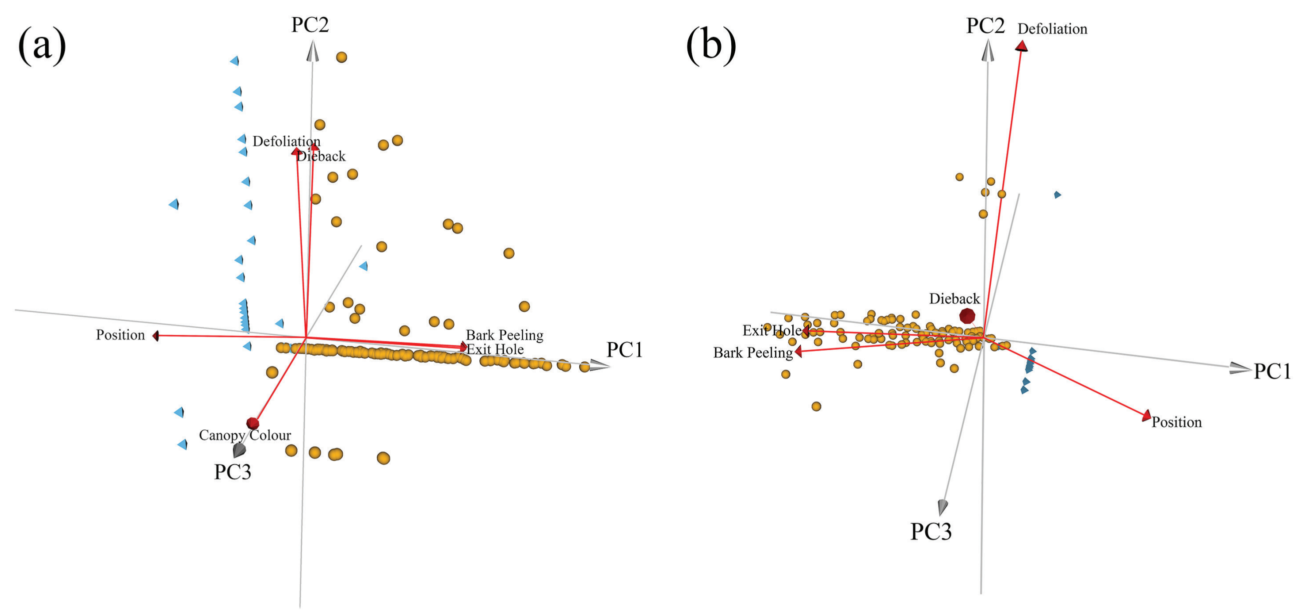

Figure 5. The 3D Principal Component Score plots for 3 principal components with the highest eigenvalues/with the highest percentage of explained variance when PCA was performed on (a) original data set and (b) data set with removed outliers. Plotting was performed using *pca3d* function from *pca3d* package in R (The R Foundation for Statistical Computing, Vienna, Austria). Blue pyramids represent noninfested trees while yellow spheres represent infested trees.

Table 3. Number of statistically significant principal components (PCs) using different stopping rules.

Stopping rule	Number of statistically significant PCs
Pseudo-*F* ratio	4
Eigenvalues *P*-value	1
Broken stick	1
Auer-Gervini: TwiceMean	1
Auer-Gervini: K-means ($k = 2$)	1
Auer-Gervini: K-means ($k = 3$)	1
Auer-Gervini: spectral clustering	1
Auer-Gervini: changepoint	1

Table 4. Rotation values for each variable along principal component 1 (PC1). After principal component analysis (PCA) was conducted on the training data set using *prcomp* function, PC1 was the principal component that showed the clearest separation between infested and noninfested trees and also accounted for the most variability in the training data set based on 3D PCA score plot. Corresponding rotation values of 6 variables along PC1 were obtained to show the correlation between these variables and PC1.

Variable	Rotation value
Colour of canopy	+0.01634139
Defoliation	+0.02019316
Dieback	+0.09567381
Bark peeling	–0.57198696
Exit holes	–0.58179903
Lowest position of bark peeling and/or exit holes	+0.56966001

Refer to Table 2 for positive values of each sign and symptom. The described formula was used to calculate BIS values for all trees within the training data set.

Figure 8 shows box-and-whisker plots for BIS values of the 290 trees in the training data set. Noninfested trees had minimum, first-quartile, medium, third-quartile, and maximum BIS values of –2.1432, –1.4198, –1.4198, –1.4198, and 0.6306. The overall mean ± standard error BIS of all noninfested trees was –1.4100 ± 0.0286. Surveyed infested trees had minimum, first-quartile, medium, third-quartile, and maximum BIS values of –0.5529, 0.0993, 0.8336, 1.9343, and 4.4781 respectively. The overall mean ± standard error BIS of all infested trees was 1.1945 ± 0.1005.

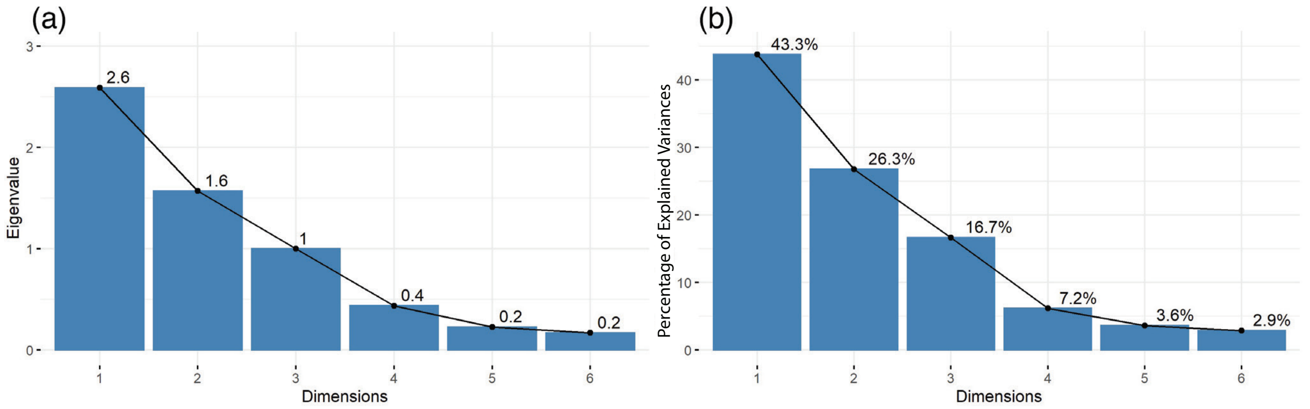

Figure 6. PCA scree plots for principal components (also referred to as "Dimensions"): (a) eigenvalues of 6 principal components and (b) percentage of explained variance of 6 principal components. Principal component analysis was performed for 6 variables (bark peeling, exit holes, dieback, canopy colour, defoliation, and lowest position of bark peeling and/or exit holes) on training data set of 290 trees using *prcomp* function. Corresponding eigenvalues and percentage of explained variances were obtained and plotted using *fviz_eig* function from *factoextra* package in R (The R Foundation for Statistical Computing, Vienna, Austria).

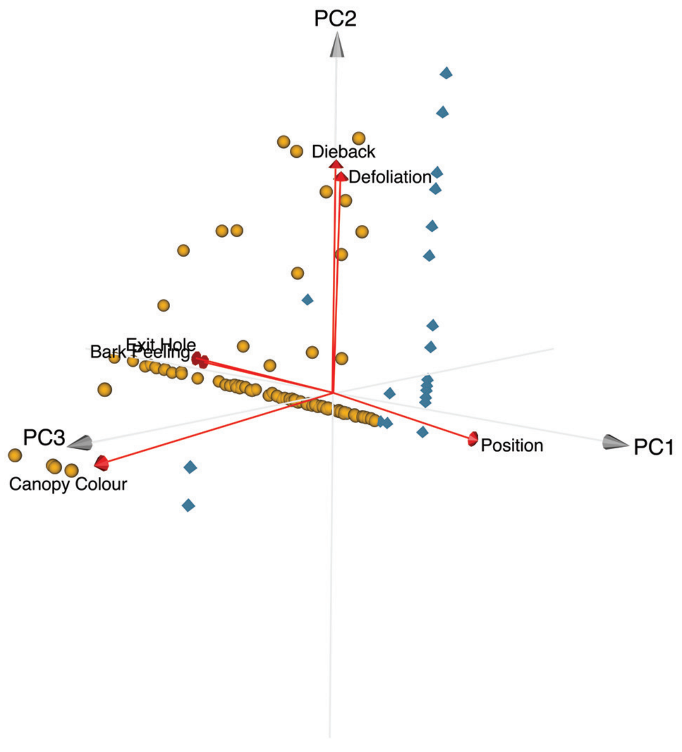

Figure 7. The 3D Principal Component Score plot for 3 principal components with the highest eigenvalues/with the highest percentage of explained variance as seen from Figure 6. Plotting was performed using *pca3d* function from *pca3d* package in R (The R Foundation for Statistical Computing, Vienna, Austria). Blue pyramids represent noninfested trees while yellow spheres represent infested trees.

Index Validation

Classification Accuracy Between Infested and Noninfested Trees

Based on BIS density curves of infested and noninfested trees in the training data set (Figure 9), the intersection between the 2 density curves was found to be –1.0132. Using this intersection as the threshold, the BIS formula was fitted into the test data set to predict the infestation class of trees. Trees were classified as infested (BIS ≥ Threshold) or noninfested (BIS < Threshold). The predicted classification was compared with the actual classification of trees. The results were presented by confusion matrix (Table 5), and 97 out of 98 trees had their infestation status correctly classified based on BIS values.

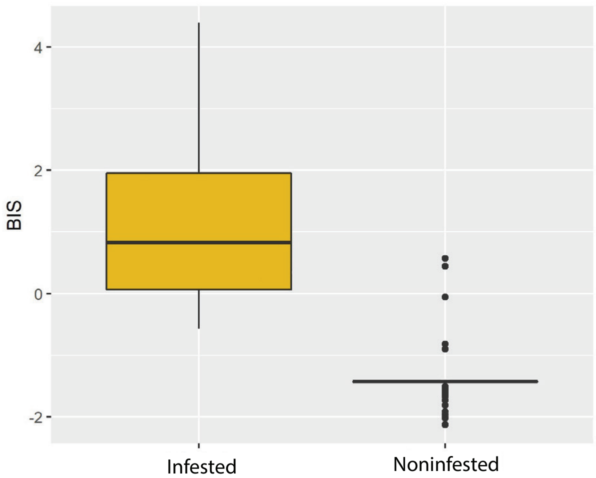

Figure 8. Box-and-whisker plots for distribution of Borer Infestation Score (BIS) values of 290 trees in training data set. The BIS formula developed from principal component analysis was applied to all 290 trees to calculate individual BIS values for trees in the training data set. Plotting was performed using *ggplot* function from *ggbiplot* package in R (The R Foundation for Statistical Computing, Vienna, Austria).

Ranking Accuracy for Infestation Level of Surveyed Areas

The BIS formula was fitted into the original data set. Average BIS scores by surveyed areas were summarized in Table 6.

Table 5. Confusion matrix for infestation status of all 98 trees in test data set. Borer Infestation Score (BIS) formula developed from principal component analysis (PCA) conducted on training data set was applied to test data set. Based on calculated BIS formula, trees were either classified as noninfested (BIS < –1.0132) or infested (BIS ≥ –1.0132).

		Predicted infestation status	
		Infested	Noninfested
Actual infestation status	Infested	20	0
	Noninfested	1	77

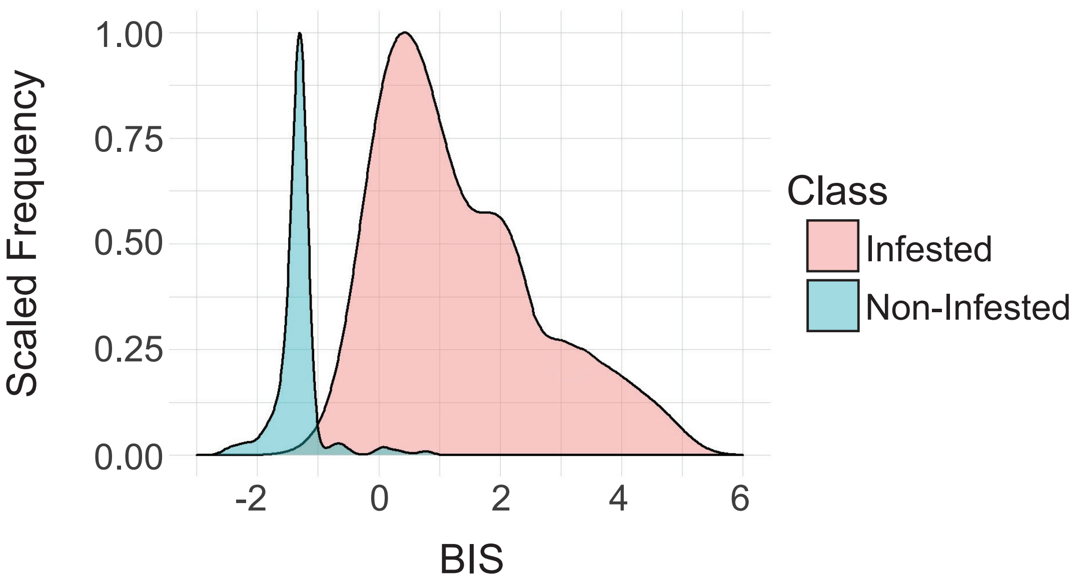

Figure 9. Density vs. BIS values for 290 trees in training data set by infestation status. Density plot for BIS values of infested and noninfested trees was obtained using *ggdensity* function from *ggpubr* package in R (The R Foundation for Statistical Computing, Vienna, Austria).

Table 6. Summary of number of infested and noninfested trees and the average Borer Infestation Score (BIS) values by surveyed areas. Developed BIS formula was applied to calculate BIS values for all trees during July 2020 survey. Average BIS values for trees for each respective area were obtained and compared against the general feedback of infestation level by managers.

Serial number	Location	No. of infested trees	No. of noninfested trees	Average BIS score	Feedback
1	Bedok South Avenue 1 (BSA)	35	32	+0.0862	High
2	East Coast Park Service Road (ECP)	86	58	+0.2932	High
3	Geylang East Central (GEC)	1	83	–1.4218	Low
4	Penjuru Road (PR)	15	41	–0.9846	Low
5	Siglap Link (SL)	16	21	–0.5218	Moderate
Total		153	235	-	-

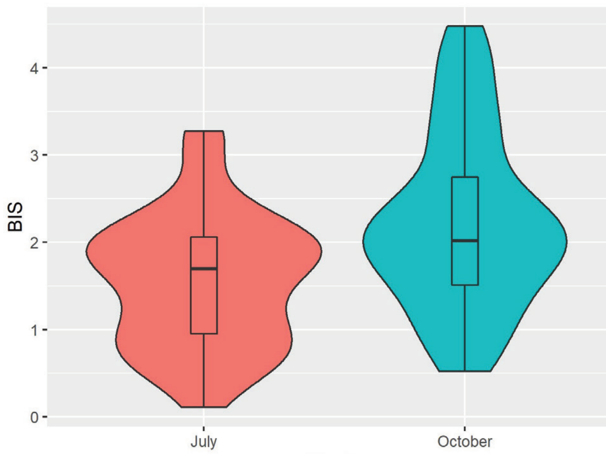

Figure 10. Violin plots of BIS values for selected 48 trees along ECP in July and October 2020. The BIS formula developed from PCA analysis was applied to survey data from July and October 2020 surveys of the selected 48 trees to calculate BIS values for these trees. Distribution of BIS values for the selected 48 trees in July and October 2020 were plotted using *ggplot* function from *ggbiplot* package in R (The R Foundation for Statistical Computing, Vienna, Austria).

Sensitivity to Temporal Progression of Infestation

The BIS value distribution of 48 selected trees along ECP was summarized in Figure 10. In July 2020, minimum, first-quartile, medium, third-quartile, and maximum BIS scores for these trees were 0.1088, 0.9511, 1.6970, 2.0609, and 3.2732. In October 2020, minimum, first-quartile, medium, third-quartile, and maximum BIS scores for the same trees were 0.5178, 1.5088, 2.0146, 2.7470, and 4.4780. Analysis of variance indicated significant difference ($F = 13.42$; df = 1.94; P-value = 0.0004) between the mean July BIS value (1.5662) and the mean October BIS value (2.2293).

DISCUSSION

As demonstrated in this study, although dieback, canopy colour, defoliation, bark peeling, exit holes, and proximity of bark peeling and/or exit holes to trunk base were common symptoms used to assess *X. globosa* infestation, they had greatly different associations with the actual infestation. Canopy colour and defoliation were found to be weakly associated with infestation of *X. globosa* as seen from the small absolute values of rotations of these variables along PC1 in the 3D PCA plot (Table 4 and Figure 7). These symptoms often signify tree decline which is not specific to this pest in rain trees. In Singapore, there are multiple stress factors, both biotic and abiotic, that can cause such symptoms to manifest in rain trees. For instance, dieback in rain trees can be caused by fungal pathogens such as *Phomopsis* sp. (Chareprasert et al. 2006; NParks 2019) and *Botryodiplodia* sp. (Boa and Lenné 1994; Hossain 2004). In Singapore, defoliation in rain trees has also been observed to be associated with the increase in larval populations of *Hypopyra* sp. (Lepidoptera: Erebidae) and *Pandesma quenavadii* (Guenée 1852)(Lepidoptera: Noctuidae) moths. The defoliation caused by the larvae of these moths is cyclical, with approximately 2 major outbreaks a year in rain trees, making the parameter insufficient alone in assessing infestation caused by *X. globosa*. In addition, environmental conditions such as water stress during extensive drought periods (King 2008) and root damage due to construction activities (Hauer et al. 1994) can also cause rapid decline in rain trees and symptoms of leaf yellowing, defoliation, and dieback.

Meanwhile, exit holes, bark peeling, and proximity of observed bark peeling and/or exit holes to trunk base were highly indicative of *X. globosa* infestation as seen from PCA. The absolute values of rotations for these variables were at least 6- to 10-fold larger than those of dieback, canopy colour, and defoliation. Such results aligned with the biology of the pest observed in Singapore and as described by Matsumoto and Irianto (1998). Based on internal pest survey and observational data, *X. globosa* was the only borer species found infesting rain trees. Appearance of holes in rain trees is unique to *X. globosa* infestation because the adult borers create these holes when they emerge.

In this study, the high absolute value of rotation of lowest position of bark peeling and/or exit holes along PC1 provided the empirical evidence to downward-infestation behaviour of *X. globosa* in rain trees. It could be hypothesized that younger branches located at higher positions in the canopy are more frequently pruned as part of routine maintenance operation, creating open wounds and bark cracks that allow easier oviposition by females. As larvae feed and mature, they would progress downward to find more food in bigger and more mature branches lower in the canopy.

The developed BIS formula gave values that were positively correlated with *X. globosa* infestation. Noninfested trees had typical BIS values of –1.42 or lower, while infested trees had BIS values of –1.01 or higher. Despite their low level of associations with *X. globosa* infestation in rain trees, symptoms such as dieback, canopy colour, and defoliation still contained valuable information which can help to obtain a more accurate assessment of infestation level. Therefore, weights for these variables were balanced in such a way that these variables still contributed to overall calculation but did not cause noninfested trees to have high BIS values. As shown from the confusion matrix for infestation status of the 98 trees in the test data set (Table 5), there was only one noninfested tree with a BIS value that was within the range of those infested. This may be due to an assessment error in recording bark peeling. Nonetheless, the error rate was found to be low (1.02%). Thus, the BIS formula provided a clear separation between infested and noninfested trees. Furthermore, when applied to calculate BIS for all trees within an area, average BIS for each area reflected the infestation level of said area based on feedback from managers. In terms of temporal progression of infestation, we demonstrated that by tracking infestation symptoms and using BIS calculation, we could see significant increase in infestation level if trees were left untreated.

From an operational perspective, although managers knew that symptoms such as exit holes, bark peeling, defoliation, and dieback were associated with borer infestation, the lack of standardized and systematic approaches in recording and computing these variables hindered their ability to accurately assess the progress of infestation and its severity. Consequently, managers could not ascertain whether their attempts at treatment including pruning of infested branches, chemical soil drenching, or chemical tree injection were effective due to lack of reference points for infestation severity. Assessment using BIS calculation filled the gap by providing a quantitative score to estimate infestation severity based on visual symptoms which could be quickly observed and recorded on site. A nonpositive change in BIS values over time after treatment would indicate effectiveness of the treatment in controlling *X. globosa* infestation. Since symptoms such as exit holes, bark peeling, and their proximity to trunk base have very high weights in BIS formula, pruning to remove infested branches would cause a temporary decrease in BIS values. Therefore, for accurate interpretation, it is important to take into consideration the pruning schedule of trees and multiple BIS values of trees over an extended period. A temporary drop in BIS values after pruning followed by an increase in BIS value indicates that the larvae inside tree trunks and branches are still alive and have emerged to create new exit holes and bark peeling.

Our survey took approximately 3 to 10 minutes to complete the assessment of one tree to calculate BIS values. Therefore, BIS assessment system can be readily integrated into surveillance routines without major additional daily workload.

CONCLUSIONS

We demonstrated that the newly developed BIS system provides a nondestructive method of estimating the infestation severity of *X. globosa* in individual rain trees and trees in an area. Future research can investigate the capacity of BIS in assessing the efficacy of *X. globosa* control measures such as chemical treatments via soil drenching or trunk injection. Since the BIS assessment method generates quantitative data, further study can be done to impose appropriate action thresholds based on BIS values to devise effective *X. globosa* management programs following the principles of an integrated pest management framework.

LITERATURE CITED

Anulewicz AC, McCullough DG, Cappaert DL. 2007. Emerald ash borer (*Agrilus planipennis*) density and canopy dieback in three North American ash species. *Arboriculture & Urban Forestry*. 33(5):338-349. https://doi.org/10.48044/jauf.2007.039

Auer P, Gervini D. 2008. Choosing principal components: A new graphical method based on Bayesian model selection. *Communications in Statistics—Simulation and Computation*. 37(5):962-977. https://doi.org/10.1080/03610910701855005

Barton DE, David FN. 1956. Some notes on ordered random intervals. *Journal of the Royal Statistical Society: Series B (Methodological)*. 18(1):79-94. https://doi.org/10.1111/j.2517-6161.1956.tb00213.x

Beeson CFC. 1941. *The ecology and control of the forest insects of India and the neighbouring countries*. Dehradun (Uttarakhand, India): Bishen Singh Mahendra Pal Singh. 1007 p.

Boa E, Lenné JM. 1994. *Diseases of nitrogen fixing trees in developing countries: An annotated list*. Chatham Maritime (Kent, United Kingdom): Natural Resources Institute. 82 p.

Bro R, Smilde AK. 2014. Principal component analysis. *Analytical Methods*. 6(9):2812-2831. https://doi.org/10.1039/C3AY41907J

Chareprasert S, Piapukiew J, Thienhirun S, Whalley AJ, Sihanonth P. 2006. Endophytic fungi of teak leaves *Tectona grandis* L. and rain tree leaves *Samanea saman* Merr. *World Journal of Microbiology and Biotechnology*. 22(5):481-486. https://doi.org/10.1007/s11274-005-9060-x

Hauer RJ, Miller RW, Ouimet DM. 1994. Street tree decline and construction damage. *Journal of Arboriculture*. 20(2):94-97. https://doi.org/10.48044/jauf.1994.017

Hossain MR. 2004. Study of the dieback of rain tree (*Albizia saman*) in the North Bengal [master's thesis]. Khulna (Bangladesh): Forestry and Wood Technology Discipline, Khulna University. 58 p.

Jolliffe IT, Cadima J. 2016. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 374(2065):20150202. https://doi.org/10.1098/rsta.2015.0202

Kim D, Kim S-K. 2012. Comparing patterns of component loadings: Principal component analysis (PCA) versus independent component analysis (ICA) in analyzing multivariate non-normal data. *Behavior Research Methods*. 44(4):1239-1243. https://doi.org/10.3758/s13428-012-0193-1

King FY. 2008. Defoliation of *Samanea saman* due to severe water stress, twig dieback disease and caterpillar infestation. Singapore Botanic Gardens (Singapore): National Parks Board. Research Technical Note 01/2008. 1 p. https://www.nparks.gov.sg/-/media/cuge/pdf/rtn-01-2008---defoliation-of-samanea-saman-due-to-severe-water-stress,-twig-dieback-disease-and-caterpillar-infestation.pdf

Matsumoto K, Irianto RSB. 1998. Adult biology of the albizia borer, *Xystrocera festiva* Thomson (Coleoptera: Cerambycidae), based on laboratory breeding, with particular reference to its oviposition schedule. *Journal of Tropical Forest Science*. 10(3):367-378. https://www.jstor.org/stable/23616014

Matsumoto K, Irianto RSB, Kitajima H. 2000. Biology of the Japanese green-lined albizzia longicorn, *Xystrocera globosa* (Coleoptera: Cerambycidae). *Entomological Science*. 3(1):33-42. https://dl.ndl.go.jp/info:ndljp/pid/10656289

McCullough DG, Poland TM, Cappaert DL, Lewis P, Molongowski J. 2005. Evaluation of trunk injections for control of emerald ash borer. In: Mastro V, Reardon R, comps. *Proceedings of the emerald ash borer research and development meeting*. FHTET 2004-15; 2004 October 5–6; Romulus, MI, USA. Morgantown (WV, USA): US Forest Service, Forest Health Technology Enterprise Team. p. 38-39.

McCullough DG, Siegert NW. 2007. Estimating potential emerald ash borer (Coleoptera: Buprestidae) populations using ash inventory data. *Journal of Economic Entomology*. 100(5):1577-1586. https://doi.org/10.1603/0022-0493(2007)100[1577:epeabc]2.0.co;2

Mercader RJ, McCullough DG, Bedford JM. 2013. A comparison of girdled ash detection trees and baited artificial traps for *Agrilus planipennis* (Coleoptera: Buprestidae) detection. *Environmental Entomology*. 42(5):1027-1039. https://doi.org/10.1603/EN12334

Morin RS, Liebhold AM, Pugh SA, Crocker SJ. 2017. Regional assessment of emerald ash borer, *Agrilus planipennis*, impacts in forests of the Eastern United States. *Biological Invasions*. 19:703-711. https://doi.org/10.1007/s10530-016-1296-x

Murfitt J, He Y, Yang J, Mui A, De Mille K. 2016. Ash decline assessment in emerald ash borer infested natural forests using high spatial resolution images. *Remote Sensing*. 8(3):256. https://doi.org/10.3390/rs8030256

NParks. 2019. *Samanea saman* (Jacq.) Merr. Flora & Fauna Web. Singapore Botanic Gardens (Singapore): National Parks Board. [Updated 2021 July 9; Accessed 2021 October 13]. https://www.nparks.gov.sg/florafaunaweb/flora/3/1/3106

Pontius J, Martin M, Plourde L, Hallett R. 2008. Ash decline assessment in emerald ash borer-infested regions: A test of tree-level, hyperspectral technologies. *Remote Sensing of Environment*. 112(5):2665-2676. https://doi.org/10.1016/j.rse.2007.12.011

Smitley D, Davis T, Rebek E. 2008. Progression of ash canopy thinning and dieback outward from the initial infestation of emerald ash borer (Coleoptera: Buprestidae) in Southeastern Michigan. *Journal of Economic Entomology*. 101(5):1643-1650. https://doi.org/10.1093/jee/101.5.1643

Suharti M, Irianto RSB, Santosa S. 1994. Behaviour of the stem borer *Xystrocera festiva* Pascoe on *Paraserianthes falcataria* and integrated control. *Buletin Penelitian Hutan*. 558:39-53.

Ter Braak CJF. 1988. CANOCO—a FORTRAN program for canonical community ordination by [partial][detrended][canonical] correspondence analysis, principal components analysis and redundancy analysis (version 2.1). Wageningen (The Netherlands): Agricultural Mathematics Group. LWA-88-02. 126 p. https://edepot.wur.nl/248698

Ter Braak CJF. 1990. Update notes: CANOCO version 3.10. Wageningen (The Netherlands): Agricultural Mathematics Group. 36 p. https://edepot.wur.nl/250652

Wang FK, Du TCT. 2000. Using principal component analysis in process performance for multivariate data. *Omega*. 28(2):185-194. https://doi.org/10.1016/S0305-0483(99)00036-5

Wang M, Kornblau SM, Coombes KR. 2018. Decomposing the apoptosis pathway into biologically interpretable principal components. *Cancer Informatics*. 17:1176935118771082. https://doi.org/10.1177/1176935118771082

ACKNOWLEDGMENTS

The study was conducted using internal funding from National Parks Board, Singapore. The authors are grateful for the support received from management and staff of NParks.

Derrick Nguyen Hoang Danh (corresponding author)
Plant Science & Health Branch
National Parks Board
Singapore
derrick_nguyen@nparks.gov.sg

Muhammad Izaan bin Istijab
Plant Science & Health Branch
National Parks Board
Singapore

Muhammad Azhari bin Mohammad Zain
Plant Science & Health Branch
National Parks Board
Singapore

Xue Huiyu
Arboriculture Branch
National Parks Board
Gardens by the Bay
Singapore

Diana Chiu
Streetscape Southeast Branch
Skyrise Greenery Branch
National Parks Board
Singapore

Varughese Philip
Plant Science & Health Branch
National Parks Board
Singapore

Mei Lai Yap
Plant Science & Health Branch
National Parks Board
Singapore

Abdul Hamid bin Mohammad
Arboriculture Branch
National Parks Board
Singapore

Conflicts of Interest:
The authors reported no conflicts of interest.

Résumé. Contexte: Déterminer le niveau d'infestation du perceur de lebbeck (*Xystrocera globosa*) dans les arbres à pluie de Singapour (*Albizia saman* [Jacq.] Merr.) est la première étape cruciale de la lutte contre ce ravageur. Les méthodes d'estimation actuelles reposent sur des symptômes tels que la coloration de la canopée, la défoliation, le dépérissement ainsi que la mesure réelle de la population de perceurs par le comptage des larves ou des trous de sortie créés par les adultes. Actuellement, on constate un manque d'approches systématiques permettant d'intégrer les différents indicateurs et symptômes de santé des arbres afin de quantifier le niveau d'infestation. Cette lacune présente des enjeux pour l'évaluation de l'efficacité des traitements puisque les gestionnaires ne peuvent déterminer quantitativement si le niveau d'infestation a changé suite au traitement. Par conséquent, cette étude visait à développer une méthode d'évaluation visuelle qui puisse intégrer tous les symptômes mentionnés afin de quantifier le niveau d'infestation. Méthodes: Nous avons examiné un total de 388 arbres à pluie et utilisé l'analyse en composantes principales (ACP) pour étudier la corrélation entre l'infestation par *X. globosa* et les différents symptômes d'infestation par les perceurs. Une formule de pointage d'infestation des perceurs (PIP) a été développée sur la base des combinaisons linéaires de composantes principales statistiquement significatives. Résultats: Le niveau d'infestation était fortement associé au décollement de l'écorce, aux trous de sortie et à la proximité de ces derniers avec la base du tronc et faiblement associé à la défoliation, au dépérissement et à la coloration de la canopée. La formule PIP développée a généré des valeurs numériques permettant de distinguer les arbres non infestés des arbres infestés, de refléter le niveau d'infestation dans les zones étudiées et de montrer la progression temporelle de l'infestation. Conclusions: La méthode d'évaluation visuelle intégrée décrite peut être exécutée rapidement sur le terrain. La formule PIP génère des scores quantitatifs qui peuvent être facilement interprétés, suivis et comparés.

Zusammenfassung. Hintergrund: Die Bestimmung des Befallsgrades des Lebbek-Bohrers (*Xystrocera globosa*) in Singapur-Regenbäumen (*Albizia saman* [Jacq.] Merr.) ist der entscheidende erste Schritt zur Bekämpfung dieses Schädlings. Die derzeitigen Bewertungsmethoden beruhen auf Symptomen wie Kronenverfärbung, Entlaubung, Absterben und der tatsächlichen Schätzung der Zünslerpopulation durch das Zählen der Larven oder der von den erwachsenen Tieren geschaffenen Austrittslöcher. Derzeit fehlt es an systematischen Ansätzen, um verschiedene Indikatoren für die Baumgesundheit und Symptome zu integrieren und den Befallsgrad zu quantifizieren. Diese Lücke erschwert die Bewertung der Wirksamkeit der Behandlung, da die Verantwortlichen nicht quantitativ feststellen können, ob sich der Befallsgrad nach der Behandlung verändert hat. Ziel dieser Studie war es daher, eine visuelle Beurteilungsmethode zu entwickeln, die alle genannten Symptome zur Quantifizierung des Befallsgrades integriert. Methoden: Wir untersuchten insgesamt 388 Regenbäume und nutzten die Hauptkomponentenanalyse (PCA), um die Korrelation zwischen dem Befall mit *X. globosa* und verschiedenen Befallssymptomen zu untersuchen. Auf der Grundlage der linearen Kombinationen der statistisch signifikanten Hauptkomponenten wurde eine Formel für den Borer Infestation Score (BIS) entwickelt. Ergebnisse: Der Befallsgrad stand in engem Zusammenhang mit Rindenablösung, Austrittslöchern und der Nähe der

Rindenablösung oder der Austrittslöcher zur Stammbasis und in schwachem Zusammenhang mit Entlaubung, Absterben und Kronenfärbung. Die entwickelte BIS-Formel lieferte numerische Werte, die zwischen nicht befallenen und befallenen Bäumen unterschieden, den Befallsgrad in den untersuchten Gebieten widerspiegelten und den zeitlichen Verlauf des Befalls zeigten. Schlussfolgerungen: Die beschriebene integrierte visuelle Bewertungsmethode kann im Feld schnell durchgeführt werden. Die BIS-Formel erzeugt quantitative Werte, die leicht zu interpretieren, zu verfolgen und zu vergleichen sind.

Resumen. Antecedentes: La determinación del nivel de infestación del barrenador de lebbek (*Xystrocera globosa*) en los árboles de acacia en Singapur (*Albizia saman* [Jacq.] Merr.) es el primer paso crucial en el control de esta plaga. Los métodos de evaluación actuales se basan en síntomas como el color del dosel, la defoliación, la muerte y la estimación real de la población de barrenadores a través del recuento de larvas o agujeros de salida creados por adultos. Actualmente, faltan enfoques sistemáticos para integrar diferentes indicadores y síntomas de salud de los árboles para cuantificar el nivel de infestación. Esta brecha plantea desafíos en la evaluación de la eficacia del tratamiento, ya que los manejadores no pueden determinar cuantitativamente si el nivel de infestación ha cambiado después del tratamiento. Por lo tanto, este estudio tuvo como objetivo desarrollar un método de evaluación visual que pueda integrar todos los síntomas mencionados para cuantificar el nivel de infestación. Métodos: se encuestó un total de 388 árboles de acacia y se utilizó el Análisis de Componentes Principales (PCA) para investigar la correlación entre la infestación por *X. globosa* y los diferentes síntomas de infestación de barrenadores. Se desarrolló una fórmula de Puntuación de Infestación del Barrenador (BIS) basada en las combinaciones lineales de los componentes principales estadísticamente significativos. Resultados: El nivel de infestación se asoció fuertemente con la afectación de la corteza, los orificios de salida y la proximidad del daño de la corteza o los orificios de salida a la base del tronco, y se asoció débilmente con la defoliación, la muerte y el color del dosel. La fórmula BIS desarrollada generó valores numéricos que distinguieron entre árboles no infestados e infestados, reflejando el nivel de infestación en las áreas estudiadas y mostraron la progresión temporal de la infestación. Conclusiones: El método de evaluación visual integrado descrito puede ejecutarse rápidamente en campo. La fórmula BIS genera puntajes cuantitativos que se interpretan, rastrean y comparan fácilmente.

Appendix.

Table S1. List of 33 trees identified as outliers based on Mahalanobis distance. Total average refers to average of the entire data set of 388 trees for respective variable.

Serial number	Colour (Total average = 1.04)	Defoliation (Total average = 1.06)	Dieback (Total average = 0.05)	Bark peel (Total average = 0.17)	Holes (Total average = 0.15)	Position (Total average = 5.78)	Mahalanobis distance
1	2	1	0.00	0.00	0.00	7	27.869
2	2	1	0.00	0.40	0.40	5	27.949
3	1	2	0.00	0.67	0.67	2	29.816
4	1	2	0.40	0.00	0.40	4	25.671
5	1	2	0.14	0.43	0.43	2	23.174
6	1	2	0.00	0.20	0.40	4	28.076
7	1	1	0.00	0.00	0.50	2	28.447
8	2	1	0.00	0.17	0.33	4	30.460
9	1	1	0.00	0.00	0.75	4	33.969
10	1	1	0.00	0.00	0.60	4	22.843
11	2	1	0.00	0.00	0.00	7	27.869
12	2	1	0.00	0.50	0.50	4	28.373
13	2	1	0.00	0.00	0.00	7	27.869
14	2	1	0.00	0.33	0.33	4	28.020
15	1	2	0.57	0.71	0.14	5	34.665
16	1	2	0.83	0.00	0.00	7	30.862
17	1	1	0.50	0.75	0.00	5	40.254
18	1	2	0.75	0.00	0.00	7	25.782
19	1	2	1.00	0.00	0.00	7	44.284
20	2	1	0.17	0.00	0.00	7	28.227
21	1	3	0.43	0.14	0.14	4	71.571
22	1	1	0.75	0.75	0.75	5	49.529
23	2	2	0.60	0.00	0.00	7	44.882
24	1	1	0.33	1.00	1.00	5	33.516
25	2	1	0.40	0.20	0.20	5	37.062
26	2	1	0.00	0.17	0.17	4	30.570
27	1	1	0.20	1.00	0.40	5	25.404
28	2	1	0.00	0.67	0.50	5	33.209
29	1	1	0.00	0.86	0.00	1	46.740
30	1	1	0.00	0.71	0.00	4	24.363
31	2	1	0.00	0.40	0.40	5	27.949
32	1	2	0.00	0.00	0.00	7	29.003
33	2	1	0.00	0.00	0.00	7	27.869