

## Revista Latinoamericana de Difusión Científica



## Exploratory Data Analysis for the Classification of *Quercus Affinis* Individuals Resistant, Tolerant, and Vulnerable to the Attack of *Andricus Quercuslaurinus* (Hymenoptera: Cynipidae)

---

DOI: <https://doi.org/10.38186/difcie.611.14>

---

Austreberto Velasco-González\*

David Cibrián-Tovar\*\*

Víctor David Cibrián-Llenderal\*\*\*

Guilherme Amorim Homem de Abreu Loureiro\*\*\*\*

### ABSTRACT

The cynipid *Andricus quercuslaurinus* has affected thousands of hectares of oak forests of *Quercus laurina* and *Quercus affinis* in Hidalgo, Mexico. The incidence of this gall wasp in hosts of different ages and phenotypic conditions requires thorough characterisation. The objective of this study was to sample the *Q. affinis* population to classify the infestation levels of *A. quercuslaurinus* in branches and leaves, characterising the response of individuals to the attack in terms of resistance (R), tolerance (T), and vulnerability (V). The high, medium, low, and very low response levels were determined for the overall total length of branch galls (OTLBG) and the total number of leaves with galls (TNLG). The absence of attack within groups of individuals was verified to associate them with R, T, and V. The OTLBG was less sensitive for diagnosing these characteristics, as the absence of attack occurred in trees of all classes in the different agamic cycles (2012, 2015 and 2018). In contrast, the absence of attack was observed only in the first class of TNLG (2018 gamic cycle). Resistance by total absence of attack corresponded to 7% of the individuals in branches and 9% in leaves.

**KEYWORDS:** Forest management, Entomology, Plant ecology, Genetics.

\*Estudiante de la Maestría en Ciencias Forestales de la Universidad Autónoma Chapingo, Texcoco, Estado de México, México. ORCID: <https://orcid.org/0000-0003-4407-2092>. E-mail: [velascoaustre@gmail.com](mailto:velascoaustre@gmail.com)

\*\*Profesor e Investigador de la Universidad Autónoma Chapingo, División de Ciencias Forestales, Texcoco, Estado de México, México. ORCID: <https://orcid.org/0000-0001-6788-0178>. E-mail: [dcibrian48@gmail.com](mailto:dcibrian48@gmail.com)

\*\*\*Investigador por México (Conahcyt) comisionado al Colegio de Postgraduados (Campus Montecillo), Postgrado en Ciencias Forestales, Texcoco, Estado de México, Mexico. ORCID: <https://orcid.org/0000-0003-1999-8520>. E-mail: [vicillan@gmail.com](mailto:vicillan@gmail.com) (corresponding author)

\*\*\*\*Investigador Posdoctoral Conahcyt en el Colegio de Postgraduados, Postgrado en Ciencias Forestales, Campus Montecillo, Texcoco, Estado de México, México, ORCID: <https://orcid.org/0000-0003-1952-0551>. E-mail: [gahal.85@gmail.com](mailto:gahal.85@gmail.com)

Recibido: 20/03/2024

Aceptado: 15/05/2024

## Análisis exploratorio de datos para la clasificación de individuos de *Quercus Affinis* Scheidw resistentes, tolerantes y vulnerables al ataque de *Andricus Quercuslaurinus* (Hymenoptera: Cynipidae)

### RESUMEN

El cinípido *Andricus quercuslaurinus* ha afectado miles de hectáreas de bosques de encinos *Quercus laurina* y *Quercus affinis* en Hidalgo, México. La incidencia de esta avispa agalladora en huéspedes de diferentes edades y condiciones fenotípicas requiere una completa caracterización. El objetivo de este estudio fue muestrear la población de *Q. affinis* para clasificar los niveles de infestación de *A. quercuslaurinus* en ramas y hojas, caracterizando la respuesta de los individuos al ataque en términos de resistencia (R), tolerancia (T) y vulnerabilidad (V). Los niveles alto, medio, bajo y muy bajo de respuesta fueron determinados para la longitud total de agallas en ramas (LTAR) y el número total de hojas con agallas (NTHA). Se verificó la ausencia de ataque dentro de los agrupamientos de individuos para asociarlos a R, T y V. La LTAR fue menos sensible para diagnosticar estas características, ya que la ausencia de ataque ocurrió en árboles de todas las clases en los diferentes ciclos agámicos (2012, 2015 y 2018). En contraste, la ausencia de ataque se observó solo en la primera clase de NTHA (ciclo gámico de 2018). La resistencia por ausencia total de ataque correspondió al 7% de los individuos en ramas y al 9% en hojas.

PALABRAS CLAVE: Gestión forestal, Entomología, Fitoecología, Genética.

### Introduction

The oak gall wasp *Andricus quercuslaurinus* (Melika & Pujade-Villar) (Cynipidae: Hymenoptera) has been identified as responsible for the mortality of up to 80% of *Quercus affinis* Scheidw. trees (V. D. Cibrián Llanderal, personal communication, 7th February 2018), a predominant species in the pine-oak forests of the Municipality of Acaxochitlán, Hidalgo (Barrera-Ruiz *et al.* 2016, Melika *et al.* 2009, Pujade-Villar 2017). The pest attack began with a territorial extension of 8 hectares in 2005 on a private property and expanded from 2012, reaching an area of 2,000 hectares in 2015, with a radius of 35 km from the original point (Pujade-Villar 2017). In this affected area, different intensities of incidence of *A. quercuslaurinus* damage have been observed between trees close to each other (V. D. Cibrián Llanderal, personal communication, 7th February 2018), which allows us to assume the presence of resistance characters within the host species (Gibson *et al.* 1982, Jactel *et al.* 2021, Pike *et al.* 2021, Politowski and Browning 1978).

Plant galls are a unique and highly specialized product of insect-plant interaction (Hurley *et al.* 2023). These highly evolved herbivorous insects cause the development of these specialized plant tissues (galls) that provide them with nutrition and a certain measure of protection against physical and biotic stress (Csóka *et al.* 2017, Hurley *et al.* 2023). Their interaction is often associated with high host specificity (Csóka *et al.* 2017). Cynipid gall initiation occurs when the plant is exposed to an accessory gland secretion during oviposition (Rohfritsch 1992). The cells of the exposed tissues are physiologically modified through differentiation and hypertrophy, resulting in the formation of the inner-gall tissue and the outer gall tissue (cortical parenchyma) (Harper *et al.* 2004, Klein *et al.* 2015). The growth of the gall and internal chamber occurs at the stage of hyperplasia (cell expansion), which in turn results in the formation of nutritive tissue that the larvae can manipulate into a suitable food source (Harper *et al.* 2004, Klein *et al.* 2015, Rohfritsch 1992). Subsequently, the number of cells in galls decreases as the larva feeds and increases in size (Hurley *et al.* 2023).

Oak gall wasps belong to the tribe Cynipini, one of the six tribes of the family Cynipidae (Hymenoptera: Cynipoidea), and are characterized by their heterogonic (cyclically parthenogenic) life cycle (Cook *et al.* 2002, Pujade -Villar *et al.* 2009). There are approximately 1000 known species of oak gall wasps and each produces a sexual and an asexual generation that has morphologically distinct galls associated with the plant family Fagaceae, specifically in genus *Quercus* (oaks) (Cook *et al.* 2002). The genus *Andricus* is one of the largest and most ecologically diverse (Cook *et al.* 2002, Melika *et al.* 2009, Pujade-Villar *et al.* 2009). Species of this genus exhibit a life cycle that involves two generations per year: sexual in spring and asexual (agamic) in autumn (Stone *et al.* 2002). Each generation is specific regarding the host species and the plant organ it attacks (Cook *et al.* 2002, Melika *et al.* 2009, Pujade-Villar *et al.* 2009). Thus, cynipids such as *Andricus*, which have a heterogonic life cycle, are restricted to areas containing both of their hosts; this prerequisite is important to determine their global distribution patterns (Stone *et al.* 2002).

*A. quercuslaurinus* has a two-year biological cycle, with alternating agamic and gametic generations, forming galls on leaves and branches (Melika *et al.* 2009). The agamic generation lays its eggs on leaves, giving rise to a generation with male and female

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241 individuals (gamic). In turn, this gamic generation will lay eggs on the branches where they will form galls for a year, and at the end, the new generation made up of only females will emerge (Melika *et al.* 2009). On a leaf there may be one or several galls formed by the oviposition of the agamic generation, whose deformations almost always develop on the main vein. On the other hand, the galls on branches and trunk are quite conspicuous bulges where many development capsules usually develop and each of them will give rise to a new individual. These malformations limit the flow of sap from the roots to the aerial part, progressively causing the death of the branches or even the death of the tree (Melika *et al.* 2009).

In general, the economic importance of the damage caused by cynipids is associated with the introduction of invasive species in different forest ecosystems (Hurley *et al.* 2023). However, native species of cynipids have also had population explosions (Eliason and Potter 2000, Eliason and Potter 2001, Melika *et al.* 2009, Pujade-Villar 2017, Pujade-Villar *et al.* 2014). Management difficulties have been reported for the control of *Callirhytis cornigera* (Osten Sacken), a native wasp that extends from southern Canada to Georgia in the USA and infests *Quercus* spp. (*Q. palustris* Muenchhausen; *Q. velutina* Lamarck; *Q. nigra* L.; *Q. rubra* L.; *Q. imbricaria* Michaux; and *Q. phellos* L.) (Eliason and Potter 2000, Eliason and Potter 2001). Alternate asexual and sexual generations of cynipids are also known, both inducing galls in *Quercus laurina* Bonpl. in Mexico, Guatemala, and El Salvador (Govaerts and Frodin 1998). In Mexico, the species diversity of this tribe is closely related to the extraordinary number of *Quercus* species (Barrera-Ruiz *et al.* 2016, Pujade-Villar *et al.* 2009). The newly described species *A. quercuslaurinus* has caused the death of thousands of trees in natural stands of *Q. laurina* in the state of Hidalgo (Melika *et al.* 2009). Asexual galls are the main cause of tree decline and death (Melika *et al.* 2009). Even with the intensive silvicultural management of destruction of attacked material, the persistence of the attack could be observed between 2003 and 2009 (Melika *et al.* 2009). This is the first cynipid wasp reported in Mexico capable of killing trees of different ages and sizes, with the potential to expand its distribution and endanger the survival of *Q. laurina* in other parts of the country (Melika *et al.* 2009). Also, the wasp *A. breviramuli* Pujade-Villar, which induces galls on branches and young shoots of *Q. laeta* Liebm., has been reported as a pest with a high level of infestation in Mexico City (Pujade-Villar *et al.*

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241 (2014). Subsequently, native cynipids can become important phytosanitary problems. For this reason, it is important to determine the degree of pest incidence within the host population, in order to support prevention and management decisions, because it reflects the economic, physiological, and other thresholds within a forest production system (Adame *et al.* 2022, Heybroek *et al.* 1982, Painter 1951). Furthermore, infestation levels are attributes that allow inferring about the vulnerability, resistance or tolerance of host plants to attack by pests and diseases (Gibson *et al.* 1982, Pike *et al.* 2021, Sniezko and Koch 2017).

The genetic characteristic of resistance means that the level of affectation of a host by a pest or disease is low or null (Gibson *et al.* 1982, Politowski and Browning 1978), while tolerance means that there is some degree of vulnerability, but the host can withstand the infestation of a pest or pathogen without reducing its performance or development (Politowski and Browning 1978). From these data it is possible to estimate population dynamics and select resistance lineages of a forest species and manage its heritability for priority pests and diseases, reducing production costs and the environmental impacts of plantation management plans. forests and natural forests (Guyot *et al.* 2015, Heybroek *et al.*, 1982, Painter 1951, Sniezko and Koch 2017). Given the need to identify and characterise the phenotypic response of resistance, tolerance, and vulnerability of *Q. affinis* to the attack of *A. quercuslaurinus*, this study analysed the distribution of infestation levels in galls of branches and leaves in sampling plots.

## 1. Materials and Methods

### 1.1. Study site and environmental conditions

The three plots of the natural regeneration forest area of *Q. affinis* are located in the “La Victoria” Hacienda of the Municipality of Acaxochitlán, Hidalgo (Table 1). These plots have an average density of 134 trees per 100 m<sup>2</sup>.

### 1.2. Field sampling and statistics

The comparative observational study was carried out in 2018 through random sampling of plots previously established in 2012. The *Q. affinis* individuals were the observation units with an approximate age of 9 years, normal diameter between 2 to 4.6 cm and total height between 3.4 to 6.65 metres. *A. quercuslaurinus* infestation levels were

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241  
 measured by two main variables: overall total length of branch galls (OTLBG, cm) and total number of leaves with galls (TNLG, dimensionless).

**Table 1.** Summary of the environmental characterization of the sampled plots. La Victoria, Acaxochitlán, Hidalgo (2018).

Sampled Plot	Longitude (W)	Latitude (N)	Altitude (masl)	Climate <sup>1</sup>	Soil type <sup>2,3</sup>	Land Cover <sup>4</sup>
1	-98.19402095	20.17055	2148	C(m): humid temperate	LVcr+ANha/3:	Pine-oak Forest
2	-98.19672373	20.169626 2	2159		Luvisol (LV) Chromic (cr) +	
3	-98.19576578	20.173013 3	2150		Andosol (AN) Haplic (ha)/3	

<sup>1</sup> Köppen classification modified by García (2001). <sup>2</sup>Instituto Nacional de Estadística y Geografía (2007); <sup>3</sup>Food and Agriculture Organization of the United Nations (FAO 2015). <sup>4</sup>INEGI (2018).

The samples of branch galls and leaf galls were independent in relation to the observation units. The OTLBG was determined by the sum of the total length of branch galls (TLBG, cm) corresponding to the years 2012, 2015, and 2018. In addition to this, the proportion of trees attacked in each year could be determined. On the other hand, to complete the TNLG on the leaf gall sample, 100 leaves were removed at random from each tree to classify them into five types according to the number of galls on the leaf: zero, one, two, three, and four. In this way, it was possible to determine the proportions of healthy and attacked leaves.

The biometric variables normal diameter and total height and the main variables were subjected to tests of normality (Kolmogorov-Smirnov with Lilliefors correction) and homogeneity of variances (Bartlett) between sampling plots. The distribution of the untransformed variables and those transformed by the square root function was verified by means of probability density plots. The analysis of variance (ANOVA) F-test and the Tukey and Scheffé post hoc tests were used to determine the significant differences (at a 5% significance level) between averages of the variables of each sampling plot. On the other hand, the variables that did not meet the criteria of the classic ANOVA were subjected to non-parametric tests of Kruskal-Wallis H and Dunn with Holm correction with the same confidence interval. The resistance levels of *Q. affinis* to the attack of *A. quercuslaurinus* were determined using the technique for established class intervals of continuous variables (OTLBG and TNLG) with the method that uses the quantiles of the variables as breaks

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241 (Bivand *et al.* 2023). The precision of these intervals was determined using the goodness of variance fit (GVF) and the tabular accuracy index (TAI) indexes described by Armstrong *et al.* (2003). The Pearson correlation coefficient was used to determine the degree of association between biometric variables and OTLBG. All statistical and graphic procedures were carried out using the R programme (R Core Team 2023).

## 2. Results

The sampled biometric variables, normal diameter and total height, had a normal distribution and homoscedasticity (Table 2). By the ANOVA F-test at the 5% significance level, the null hypothesis of equality of averages between sampling plots was rejected for the normal diameter variable and accepted for total height (Table 2). According to the Tukey (Figure 1A) and Scheffé (Figure 1B) tests, the average normal diameter of plots 2 and 3 was statistically different. However, no differences could be detected between the averages of total height between plots (Figure 1C and D), confirming the result of the ANOVA F-test (Table 2).

According to the Kolmogorov-Smirnov test with Lilliefors correction, the normal distribution of the OTLBG variable was not detected (Table 2). For this reason, this variable was subjected to a transformation by the square root function (Table 2). Without the square root function transformation, the Tukey and Scheffé tests were susceptible to bias caused by outliers of OTLBG (Figure 2A and C), with the Tukey test being more sensitive to these outliers (Figure 2A). Once transformed, the hypothesis of equality of OTLBG averages between sampling plots could be accepted by the ANOVA F-test (Table 2), also confirmed by the Tukey and Scheffé tests (Figure 2B and D).

Figure 3 shows the probability density graphs of the OTLBG variable before and after the square root transformation. The average was higher than the median in the untransformed variable, causing a positive bias to the right (Figure 3A). On the other hand, the transformation brought the mean and median values closer to the centre of the distribution (Figure 3B). Both distributions of the variables with or without transformation were divided according to break values that represented their quantiles (Figure 3). Breaks were real numbers (Table 3) represented as integers in the graphs (Figure 3). The models fitted from the calculated OTLBG breaks with and without transformation corresponded to GVF and TAI values greater than 0.70 (Table 3).

**Table 2.** Summary of the analysis of variance and tests of normality and homoscedasticity. Natural regeneration forest area of *Quercus affinis* Scheidw. after the attack of *Andricus quercuslaurinus* (Melika & Pujade-Villar). Acaxochitlán, Hidalgo.

Year	Variable (Unit)	Variation Source (DF <sup>1</sup> )	Mean Square	F Test		CV (%)	KST <sup>2</sup>	BT <sup>3</sup>
				p-value	p-value			
2018	Normal Diameter (cm)	Sampled Plots (2)	1.24	0.03	19.32	0.24	0.06	
		Residuals (99)	0.35					
	Total Height (m)	Sampled Plots (2)	0.43	0.50	15.53	0.41	0.48	
		Residuals (99)	0.61					
2012+2015+2018	Overall Total Length of Branch Galls (OTLBG, cm)	Sampled Plots (2)	61273.71	0.06	92.76	<0.01	0.09	
		Residuals (99)	21705.53					
	OTLBG <sup>0.50</sup>	Sampled Plots (2)	66.58	0.19	57.00	0.11	0.15	
		Residuals (99)	38.91					
2012	Total Length Of Branch Galls (TLBG, cm)	Sampled Plots (2)	74.90	0.88	161.00	<0.01	0.88	
		Residuals (99)	586.55					
	TLBG <sup>0.50</sup>	Sampled Plots (2)	0.98	0.89	114.09	<0.01	0.87	
		Residuals (99)	8.64					
2015	TLBG (cm)	Sampled Plots (2)	35710.57	0.12	104.52	<0.01	0.05	
		Residuals (99)	16370.75					
	TLBG <sup>0.50</sup>	Sampled Plots (2)	45.28	0.33	70.18	0.02	0.34	
		Residuals (99)	40.50					
2018	TLBG (cm)	Sampled Plots (2)	3122.72	0.04	145.96	<0.01	<0.01	
		Residuals (99)	973.47					
	TLBG <sup>0.50</sup>	Sampled Plots (2)	18.11	0.14	84.55	<0.01	0.05	
		Residuals (99)	8.87					
2018	Total Number of Leaf with Galls (TNLG, dimensionless)	Sampled Plots (2)	162.1	0.26	91.79	<0.01	0.25	
		Residuals (113)	118.4					
	TNLG <sup>0.50</sup>	Sampled Plots (2)	4.292	0.26	60.44	<0.01	0.86	
		Residuals (113)	3.174					

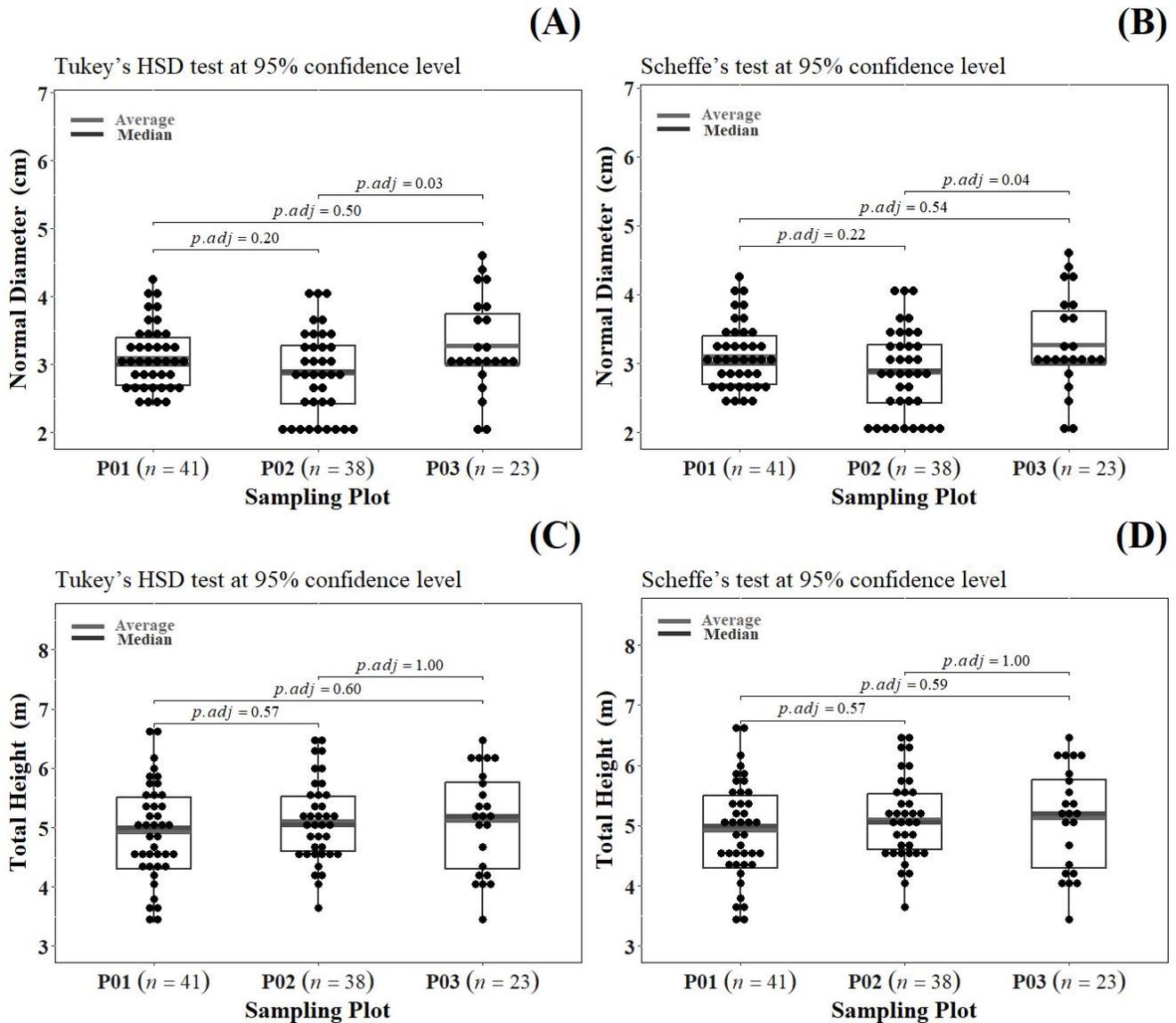
<sup>1</sup>Degree of freedom. <sup>2</sup>Kolmogorov-Smirnov normality test with Lilliefors correction. <sup>3</sup>Bartlett's test for homogeneity of variances.

**Table 3.** Class intervals for continuous variables based on quantiles and indices for the analysis of fitted models.

Variable (Unit)	Break					GVF <sup>1</sup>	TAI <sup>2</sup>
	1	2	3	4	5		
Overall Total Length of Branch Galls (OTLBG, cm)	0.00	36.18	116.80	234.93	640.30	0.85	0.71
OTLBG <sup>0.5</sup>	0.00	6.01	10.81	15.33	25.30	0.90	0.71
Total Number of Leaves with Galls (TNLG, dimensionless)	0.00	2.00	7.50	20.00	44.00	0.90	0.76
TNLG <sup>0.5</sup>	0.00	1.41	2.74	4.47	6.63	0.93	0.74

<sup>1</sup>Goodness of variance fit. <sup>2</sup>Tabular accuracy index.

**Figure 1.** Post hoc tests to detect significant statistical differences between averages of normal diameter and total height of *Quercus affinis* Scheidw. in three plots sampled in a natural regeneration forest area after the attack of *Andricus quercuslaurinus* (Melika & Pujade-Villar). Acaxochitlán, Hidalgo, 2018.

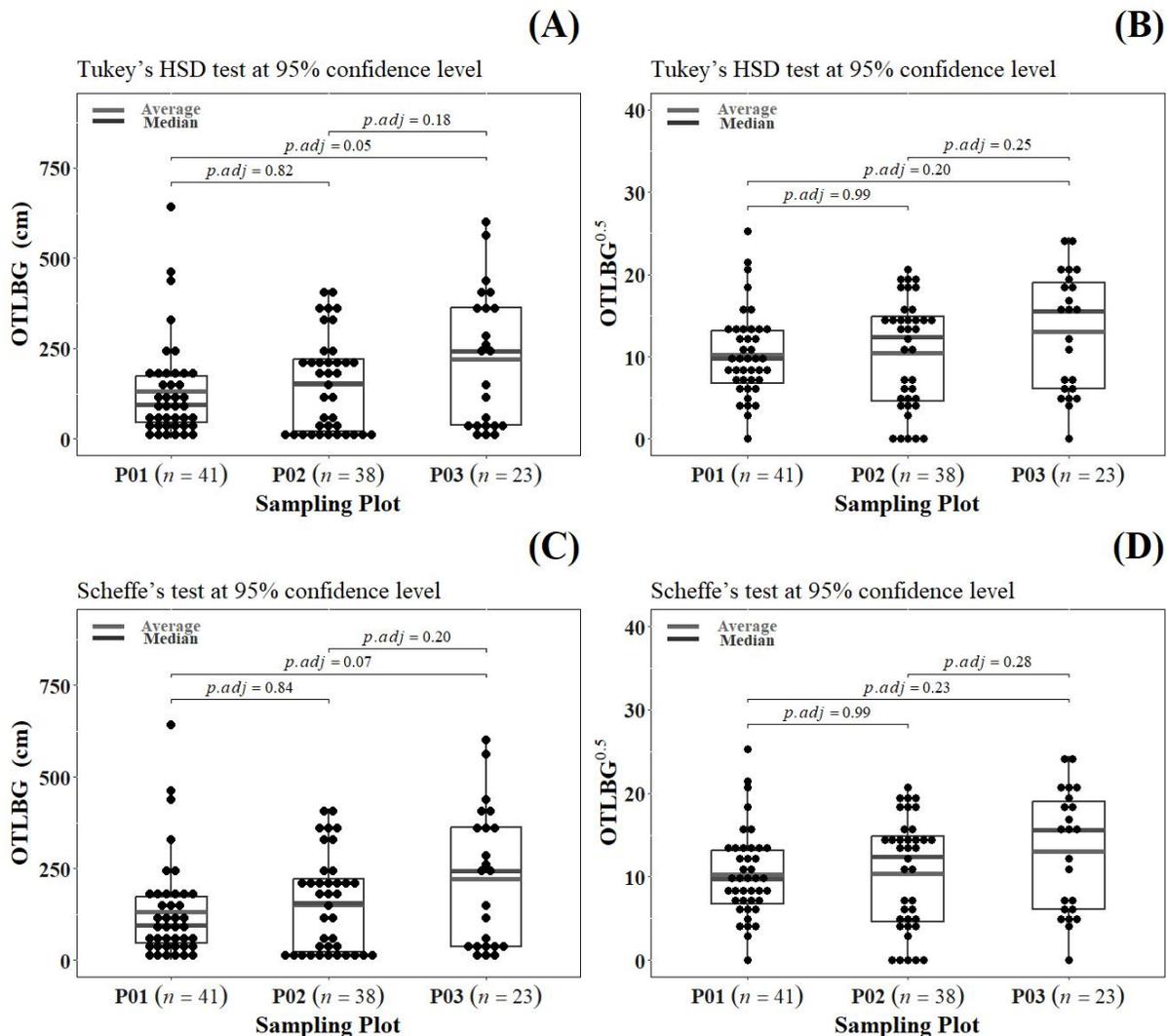


The hypothesis of normal distribution could not be verified for the TLBG variables corresponding to the years 2012, 2015, and 2018 before and after the square root transformation (Table 2). However, the square root transformation improved the condition of equality of variances between plots (Table 2). These transformed variables were subjected to post hoc non-parametric tests (Figure 4). According to the Kruskal-Wallis H

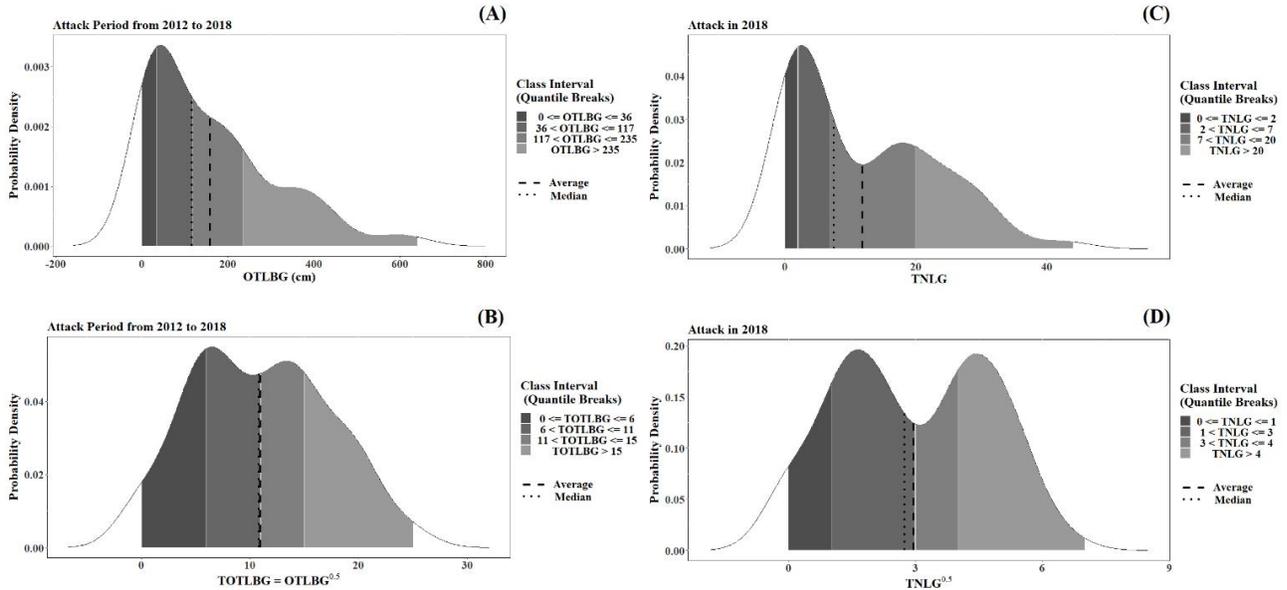
A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241

test, no differences in the stochastic domain could be detected between the samples of the sampling plots (Figure 4). The adjusted probability values of the Dunn test with Holm's correction confirmed that there were no differences between sampled plots (Figure 4). Therefore, the square root transformation was able to improve the skewed distribution with extreme right values of the TLBG variables for the years 2012, 2015, and 2018 (Figure 5). The class intervals obtained for the OTLBG variable were also applied to the TLBG variables of 2012, 2015, and 2018 (Figure 5).

**Figure 2.** Post hoc tests to detect significant statistical differences between averages of overall total length of branch galls (OTLBG) in *Quercus affinis* Scheidw. formed by *Andricus quercuslaurinus* (Melika & Pujade-Villar) infestation in three plots sampled in a natural regeneration forest area. OTLBG is the sum of the attacks from the years 2012, 2015, and 2018. Acaxochitlán, Hidalgo, 2018.



**Figure 3.** Probability density plots with the distribution of the variables overall total length of branch galls (OTLBG) not transformed (A) and transformed by the square root function (B), and total number of leaves with galls (TNLG) not transformed (C) and transformed (D). The class intervals show breaks obtained from the quantiles OTLBG and TNLG. *Quercus affinis* Scheidw. natural regeneration forest area under cyclical attack by *Andricus quercuslaurinus* (Melika & Pujade-Villar). OTLBG is the sum of the attacks from the years 2012, 2015, and 2018. Acaxochitlán, Hidalgo, 2018.



**Table 4.** Combinations between levels of occurrence of *Quercus affinis* Scheidw. leaves with different numbers of galls formed by *Andricus quercuslaurinus* (Melika & Pujade-Villar).

Number of Galls on Leaves					Combination Group	Group Description	Absolute Frequency
0	1	2	3	4			
L	-	-	-	-	1	Trees without leaf galls.	11
NA	L	-	-	-	2	Trees with leaves with a single gall.	27
NA	-	L	-	-	3	Trees with leaves with two galls.	1
NA	-	-	L	-	4	Trees with leaves with three galls.	0
NA	-	-	-	L	5	Trees with leaves with four galls.	0
NA	L	L	-	-	6	Trees with leaves with one and two galls.	58
NA	L	-	L	-	7	Trees with leaves with one and three galls.	0
NA	L	-	-	L	8	Trees with leaves with one and four galls.	0
NA	-	L	L	-	9	Trees with leaves with two and three galls.	0
NA	-	L	-	L	10	Trees with leaves with two and four galls.	0
NA	-	-	L	L	11	Trees with leaves with three and four galls.	0
NA	L	L	L	L	12	Trees with leaves with all levels of galls	2
NA	L	L	L	-	13	Trees with leaves with one, two and three galls.	17
NA	L	L	-	L	14	Trees with leaves with one, two and four galls.	0
NA	L	-	L	L	15	Trees with leaves with one, three and four galls.	0
NA	-	L	L	L	16	Trees with leaves with two, three and four galls.	0
Total (Overall Sample Size)							116

Considered level (L). Not applicable (NA).

In Figure 6, the class intervals of the OTLBG variable were established as levels of phenotypic response of *Q. affinis* to *A. quercuslaurinus* attack to classify the sample observations in the bivariate scatterplots. There were 5 levels of phenotypic response of trees to insect attack: high, medium, low, and very low (Figure 6). The hypothesis of some degree of linear association between normal diameter and total height of the sampled trees could not be rejected; however, the Pearson correlation coefficient of 0.35 was very low (Figure 6A). These biometric variables were also not shown to be associated with the evolution of the OTLBG gradient (Figure 6B and C).

The variable TNLG also did not exhibit a normal distribution (Table 2). Still, the square root transformation improved the probability of acceptance of the hypothesis of equality of variances by the Bartlett test (Table 2). Non-parametric tests for the TNLG variable before and after the square root transformation were shown in Figure 7. There were no significant differences between medians of sampling plots (Figure 7). Dunn's pairwise comparison tests confirmed that there were no significant differences between sampled plots (Figure 7). The breaks of the class intervals of the TNLG variable before and after the square root transformation (Figure 3C and D) formed class intervals with GVF and TAI values greater than 0.70 (Table 3).

The classification by the number of galls on leaves was used to generate a structure of combinations that reflected the real possibilities of finding leaves with galls on the trees (Table 4). In the sample, only six groups of combinations between levels of the number of galls on leaves were found (Table 4). Nine point four eight percent of the sampled trees did not have leaves with galls (Table 4). The relative frequency of the trees that were among the groups of possible combinations between leaves with different numbers of galls is presented below in descending order (Table 4): 50% of the trees with leaves with one and two galls; 23.28% of trees with leaves with a single gall; 14.66% of trees with leaves with one, two, and three galls; 1.72% of trees with leaves with all gall numbers; and 0.86% of trees with leaves with two galls.

Figure 8 shows the branch and leaf gall infestation statuses reflected in the condition of attacked or non-attacked trees. The levels of absence or presence of branch attack in the trees were consistent with the levels of phenotypic response (Figure 8) determined from the interval classes of the quantiles of the OTLBG variable (Table 3). One hundred

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241

percent of the trees belonging to the high phenotypic response level were not attacked in 2012 (Figure 8A). In 2015, 50% of the trees previously classified as having the high level of phenotypic response were attacked (Figure 8A). In 2018, there was an increase of 11.54% in trees attacked at this high level of phenotypic response (Figure 8A). Of the 26 trees that made up this group with a high level of phenotypic response, 10 were not attacked in 2018 and 7 were trees that were not attacked in previous years (Figure 9A). These trees that were not attacked were at the limit of 0 in the probability density graphs related to branch attack (Figures 3 and 5). The attack levels increased considerably in 2015 (Figure 8A), so that the last levels of phenotypic response, low and very low, expressed 100% of the attacked trees.

**Table 5.** Interpretation of the class intervals of the variables that characterize the phenotypic response of *Quercus affinis* Scheidw. to the attack of *Andricus quercuslaurinus* (Melika & Pujade-Villar).

Variable (Unit)	Class Interval (Quantile Breaks)			Phenotypic Response of the Tree to Insect Attack	
	Variable Condition			Levels	Hypothetical Genotypic Condition
	Not transformed <sup>1</sup>	Transformed <sup>2</sup>	Inverse Transformation <sup>3</sup>		
Overall Total	$0 \leq x \leq 36$	$0 \leq y \leq 6$	$0 \leq z \leq 36$	High	Resistance
Length of Branch	$36 < x \leq 117$	$6 < y \leq 11$	$36 < z \leq 121$	Medium	Tolerance
Galls (cm)	$117 < x \leq 235$	$11 < y \leq 15$	$121 < z \leq 225$	Low	Vulnerability
	$x > 235$	$y > 15$	$z > 225$	Very Low	Vulnerability
Total Number of	$0 \leq x \leq 2$	$0 \leq y \leq 1$	$0 \leq z \leq 1$	High	Resistance
Leaves with Galls	$2 < x \leq 7$	$1 < y \leq 3$	$1 < z \leq 9$	Medium	Tolerance
(dimensionless)	$7 < x \leq 20$	$3 < y \leq 4$	$9 < z \leq 16$	Low	Vulnerability
	$x > 20$	$y > 20$	$z > 16$	Very Low	Vulnerability

<sup>1</sup>Original value (*x*-value). <sup>2</sup>Square root transformation ( $y = \sqrt{x}$ ). <sup>3</sup> Inverse square root function ( $z = y^2$ ).

Leaf attacks were measured only in 2018 (Figures 8A and C). Unlike branch attack, where the class of non-attacked trees was distributed among all levels of phenotypic response of the trees, the class of trees not attacked by leaf gall formation only appeared at the high level of phenotypic response (Figure 8B). The trees classified with a high level of phenotypic response totalled 36, and only 30.56% of these trees did not have any type of leaf attack (Figure 8B). From the middle level onwards, all trees presented some level of

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241  
leaf attack (Figure 8B). In Figure 8C, the condition classes of attacked and non-attacked trees were separated according to the six combinations of occurrence of leaves with different numbers of galls (Table 4). Thirty-seven point five percent of the trees with a high level of phenotypic response did not have leaves with galls in the first sampled plot, followed by 25% in the second and also in the third plot (Figure 8C). Combinations with two or more kinds of leaves with different numbers of galls were more frequent as the level of phenotypic response of the trees decreased (Figure 8C).

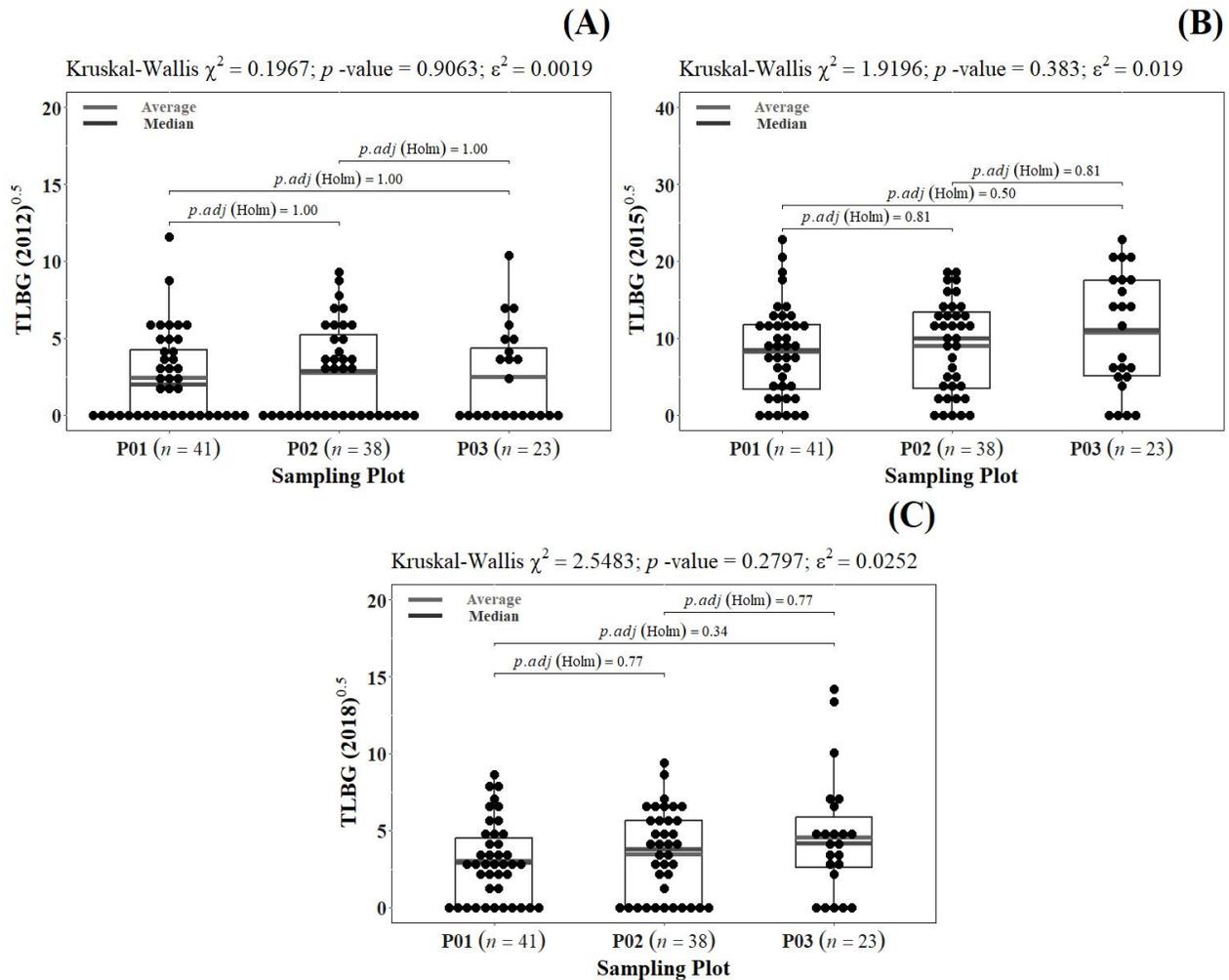
The transformations of the OTLBG and TNLG variables generated different breaks as lower and upper limits of the class intervals (Tables 3 and 5). The generated breaks could express four classes for both variables (Tables 3 and 5). The hypothetical genetic conditions of resistance, tolerance, and vulnerability were associated with the levels of phenotypic response of the trees to the attack of oak root knot wasps represented by the variables OTLBG and TNLG (Table 5). In addition to the class intervals of the OTLBG and TNLG variables before and after the square root transformation, the inverse square root transformation applied to the transformed variable is also presented (Table 5). Trees that had an OTLBG less than or equal to 36 cm and a TNLG less than or equal to 2 were considered resistant (Table 5). Tolerant trees were those that had an OTLBG greater than 36 cm and less than or equal to 117 cm (Table 5). In turn, tolerance to leaf attack was found with TNLG greater than 2 and less than or equal to 7 (Table 5). The hypothetical condition of vulnerability was associated with the groups of low and very low phenotypic response measured by the variables OTLBG and TNLG (Table 5). The lower limit of the vulnerability threshold to branch attack was determined with an OTLBG greater than 117 cm (Table 5). On the other hand, the lower limit of the vulnerability threshold to leaf attack was determined with a TNLG greater than 7 (Table 5).

The symptomatology of the attack of *A. quercuslaurinus* on *Q. affinis* is shown in Figures 9 to 11. Figures 9 and 10 shows the attack of branches of the asexual generation of *A. quercuslaurinus*. The stages of initial hypertrophy (Figure 9A), hyperplasia that increases the size of the branch gall (Figure 9B), and its final stage of maturity where adult insects emerge from their chambers (Figure 9C) were characterised. Asexual infestation is very easy to detect given the abundance of galls on oaks (Figure 10). Figure 11 shows the

leaf attack corresponding to the sexual generation of this cynipid. The galls produced deformations in the leaves from their initial stage (Figure 11A).

**Figure 4.** Kruskal-Wallis H test to detect differences in stochastic domain between the samples of the variable total length of branch galls (TLBG) obtained in sampled plots of the natural regeneration forest area of *Quercus affinis* Scheidw. under attack by *Andricus quercuslaurinus* (Melika & Pujade-Villar). Dunn's test with Holm's correction identifies significant differences between pairs of sampled plots. TLBG is transformed by the square root function. The insect attack periods are respectively 2012 (A), 2015 (B), and 2018 (C).

Acaxochitlán, Hidalgo, 2018.



### 3. Discussion

The sampling of the variables that characterised the infestation of *A. quercuslaurinus* on the population of *Q. affinis* due to the formation of galls on branches and leaves managed to capture the uniform distribution of the variability of these attacks in

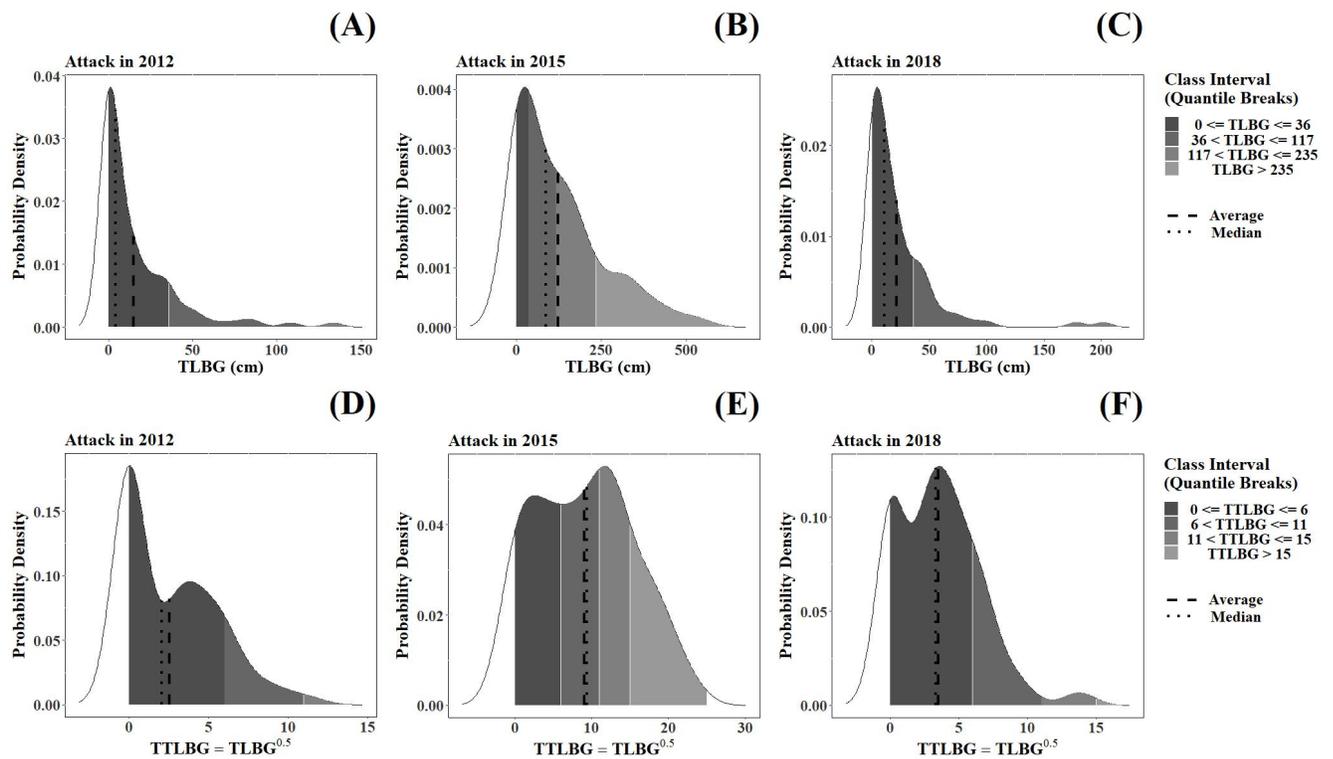
A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241

the study plots. Only the variable normal diameter showed a significant difference in averages between sampled plots 2 and 3 (Figures 1A and B), however, there was no degree of linear association that allowed inferring about the preference of the insect on a certain magnitude of the normal diameter of the trees (Figure 6). Cibrián-Tovar *et al.* (2013) reported that this cynipid had not shown preference between hosts due to age or health condition. Furthermore, the plant population in the study area originated from a natural regeneration that occurred from the collapse of infested trees in 2005 (Melika *et al.* 2009, Pujade-Villar 2017). There are two hypotheses that should be investigated in future studies: the biometric phenotypic variability was related to the species, or this variability was a product of natural hybridisation between *Q. affinis* and *Q. laurina* (González-Rodríguez *et al.* 2004a, González-Rodríguez *et al.* 2004b, Ramos-Ortiz *et al.* 2015). The environmental characteristics of the plots were very homogeneous (Table 1), therefore, it is not possible to associate these conditions with the differences found in normal diameter (Figure 1). It is important to highlight that the recurrence of the wasp in the same study area verified the hypothesis that the control of the species is complex (Barrera-Ruiz *et al.* 2016, Melika *et al.* 2009).

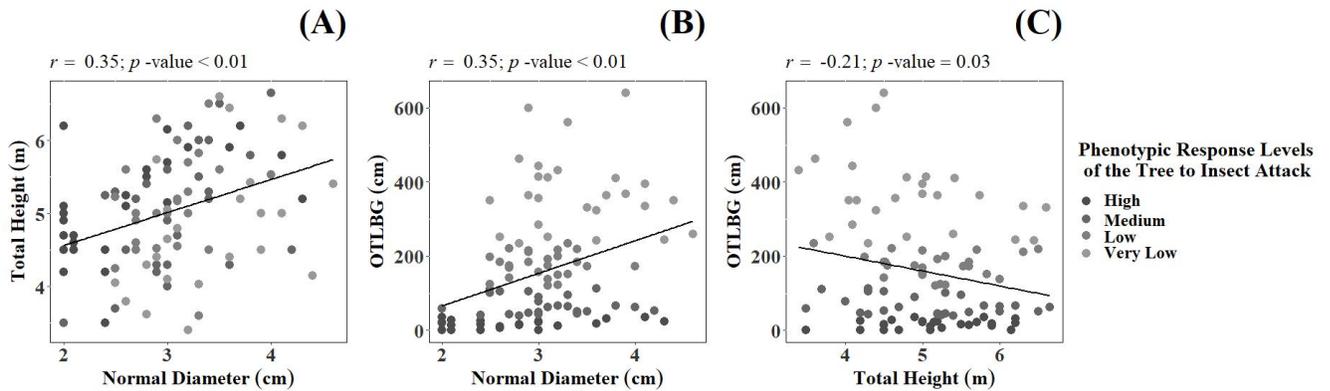
The variables used to measure the incidence of damage from branch and leaf attacks were an innovation of the present study (Table 2-5; Figures 2-8). In general, variables that represented insect attack on trees were usually averages quantifying parts affected by absolute or relative numbers. Barrera-Ruiz *et al.* (2016) quantified leaf galls and branch galls by evaluating the effect of systemic insecticides on the control of *A. quercuslaurinus* in the same study area. However, this study sought to quantify the damage of these root-knot wasps by characterising the attack on branches by the length of the gall formed on the trees. According to Cibrián-Tovar *et al.* (2013), the damage that these wasps can cause with branch galls (Figures 9 and 10) is higher than the damage caused by leaf galls (Figure 11), reducing the growth of trees, causing the dieback of branches, opening of the canopy due to the reduction of the canopy with effects on the water collection capacity and even affecting ecological succession. Once the length of the branch gall was measured, the total length was obtained through the sum of the values found in all the affected galls (Figures 2-5), thus characterising the dimension of the real damage to the tree. In turn, the quantification of the absolute number of trees with infested

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241 leaves (Figure 8B) gained a new focus with the classification of these individuals by the number of galls that were formed on each leaf (Table 4; Figure 8C). Unlike attack on leaves, attack on branches could be monitored retrospectively, because branches affected by gall formation remained on the trees (Cibrián-Tovar *et al.* 2013). Branch galls (Figures 9 and 10) went through different stages of development, initial, growth, maturation, and ageing (Stone *et al.* 2002), and each developmental stage associated with the agamic reproductive cycles of *A. quercuslaurinus* could be easily identified in the field. The galls formed in 2012 were in a height range of approximately 1.30 to 1.50 m. The galls of the successive cycles of 2015 and 2018 are also at different heights. This branch attack arrangement facilitated the quantification of the OTLBG variable.

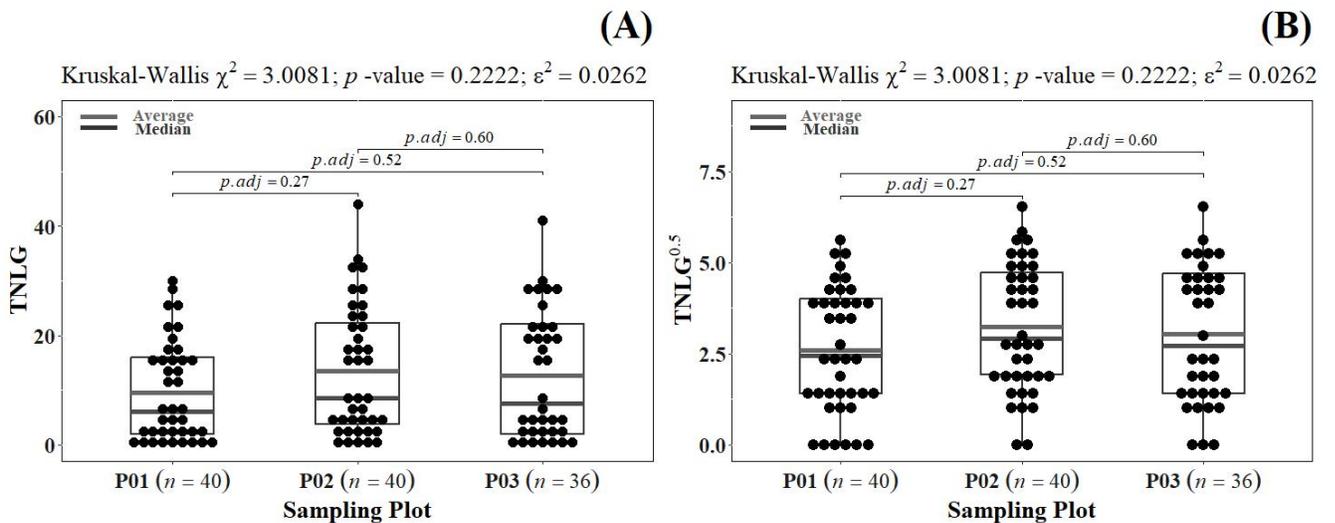
**Figure 5.** Probability density plots with the distribution of the variable total length of branch galls (TLBG) not transformed (A, B, and C) and transformed by the square root function (D, E, and F). The class intervals show breaks obtained from the quantiles of the variable overall TLBG that's correspond the sum of the attacks from the years 2012, 2015, and 2018. *Quercus affinis* Scheidw. natural regeneration forest area under cyclical attack by *Andricus quercuslaurinus* (Melika & Pujade-Villar). Acaxochitlán, Hidalgo, 2018.



**Figure 6.** Scatter plots of the relationship between biometric variables of *Quercus affinis* Scheidw. and overall total length of branched galls (OTLBG) formed by *Andricus quercuslaurinus* (Melika & Pujade-Villar). The OTLBG corresponds to the sum of the attacks from the years 2012, 2015, and 2018. Sample observations are classified by phenotypic response levels associated with OTLBG. Acaxochitlán, Hidalgo, 2018.



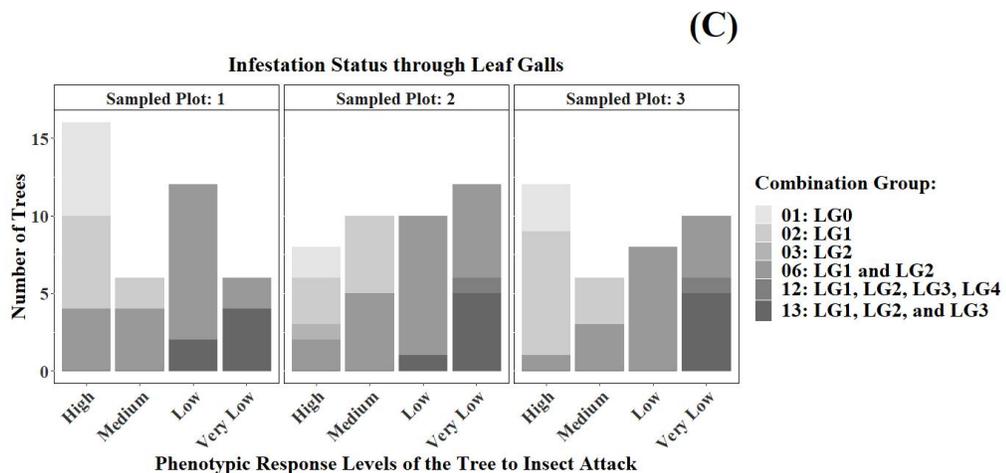
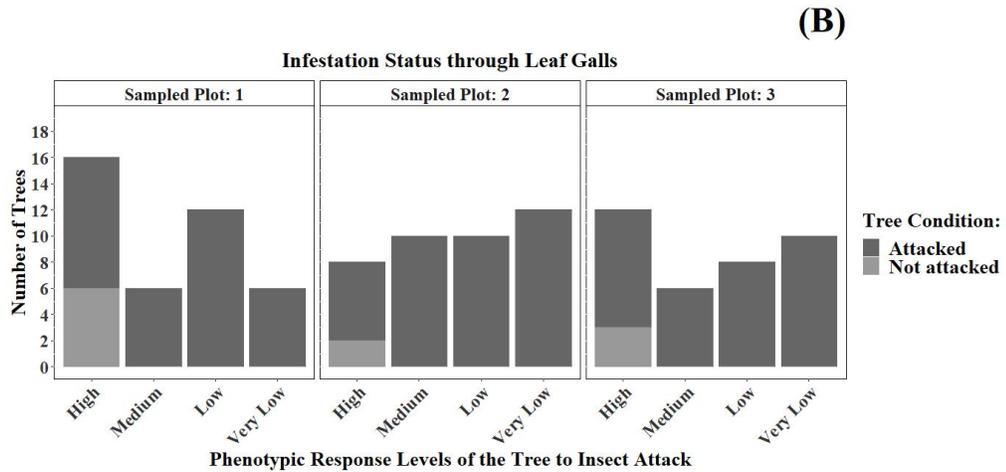
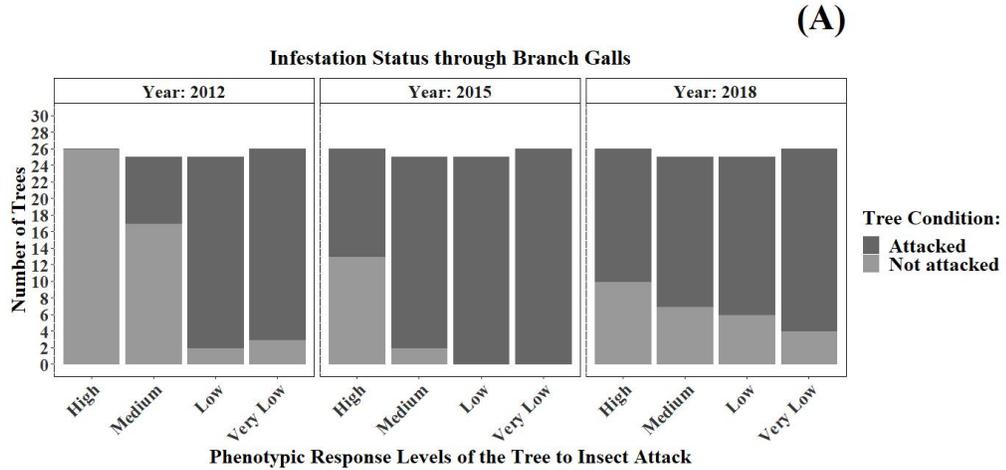
**Figure 7.** Kruskal-Wallis H test to detect differences in stochastic domain between the samples of the variable total number of leaf galls (TNLG) obtained in sampled plots of the natural regeneration forest area of *Quercus affinis* Scheidw. under attack by *Andricus quercuslaurinus* (Melika & Pujade-Villar). Dunn's test with Holm's correction identifies significant differences between pairs of sampled plots. TNLG is transformed by the square root function. Acaxochitlán, Hidalgo, 2018.



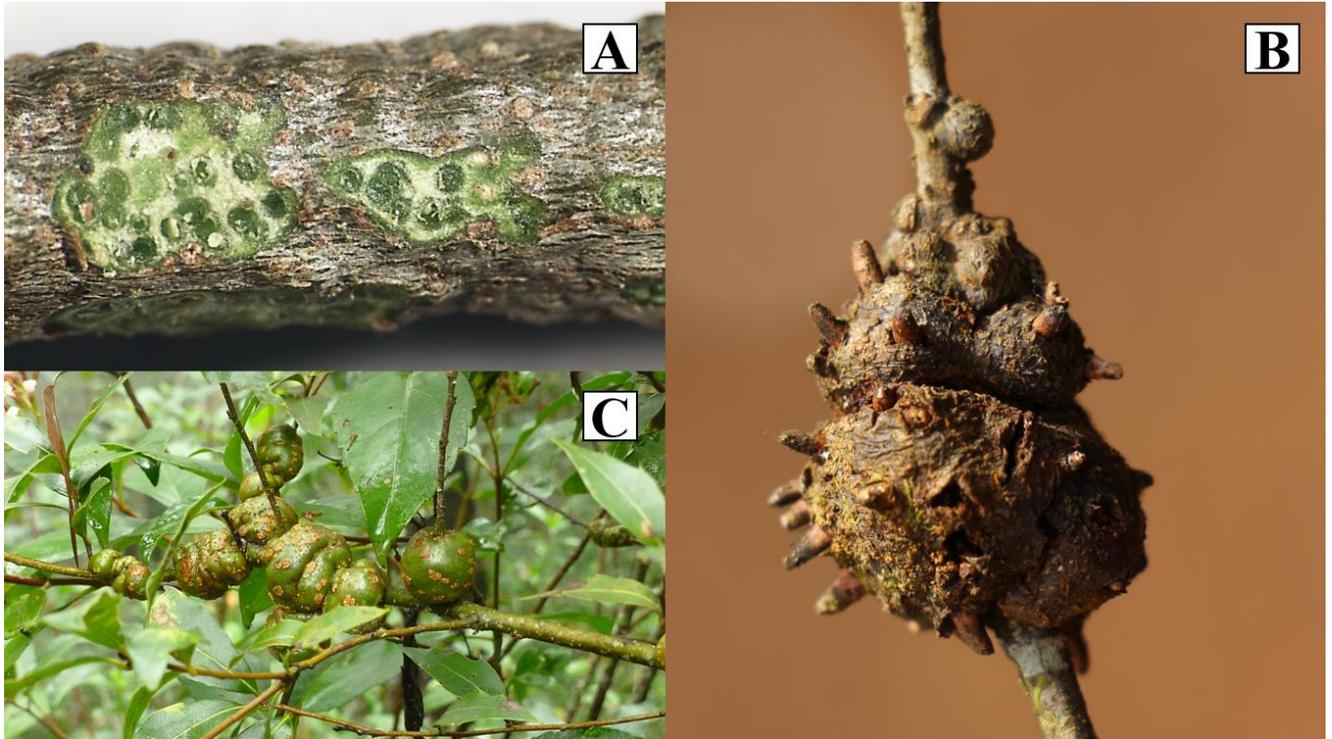
Cibrián-Tovar et al. (2013) indicated that there were differences in oak gall wasp infestation between *Q. affinis* individuals. The results of the present study with the analysis of the OTLBG and TNLG variables (Table 2-5; Figures 2-8) proved this same hypothesis of differential infestation between individuals from the same area. The phenotypic response of interest in this study was related to the different levels of occurrence of morphological modifications of branches and leaves induced by oak wasps (Figures 9-11). The technique of determining quantile class intervals of the OTLBG and TNLG variables (Table 3 and 5), associating the distribution of these variables with the levels of phenotypic response of *Q. affinis* to the attack of *A. quercuslaurinus* could have been used for pre-identification of resistant, tolerant, and vulnerable individuals (Figures 8). The breaks generated from the quantiles of these variables were able to quantitatively discriminate these levels of infestation of branches and leaves (Figure 8). From the analysis of the OTLBG and TNLG variables (Figure 8) it was possible to identify individuals with high genetic potential for resistance grouped in the first-class interval with a high level of phenotypic response of the host to the pest attack. The following classes expressed an ascending order of attack on branches and leaves, consequently an ascending order of vulnerability of these oaks to gall wasp (Figure 8). The relative frequencies associated with the combinations between classes of leaves with different numbers of galls increased the discriminant power of the TNLG variable (Table 4; Figure 8C).

It is worth mentioning that it was empirically observed that trees with fewer galls had fewer larval chambers and exit holes. However, the hypothesis of resistance or tolerance associated with the characteristics of branch galls should be analysed in future studies. The variables OTLBG and TNLG reflected the levels of damage caused by wasps in oaks. Galls are highly specialised structures of insect-plant interaction (Csóka *et al.* 2017, Harper *et al.* 2004, Hurley *et al.* 2023, Klein *et al.* 2015, Rohfritsch 1992); therefore, the measurement techniques of the variables that represent the phenomenon and the data obtained from these variables deserved a more detailed exploratory analysis in order to find the patterns that allowed quantifying the damage and characterising the condition of resistance, tolerance, and vulnerability of individuals of the host species to the attack of the gall wasp.

**Figure 8.** Bar plots show the presence and absence of *Andricus quercuslaurinus* (Melika & Pujade-Villar) infestation associated with the phenotypic response levels of *Quercus affinis* Scheidw. to the attack in branches (A) and leaves (B). Group formed from the combination between levels of leaf galls existing on the trees (C): absence of leaf galls (LG0); leaves with just one gall (LG1); leaves with two galls (LG2); leaves with three galls (LG3); and, leaves with four galls (LG4). Acaxochitlán, Hidalgo, 2018.



**Figure 9.** Stages of gall formation of the asexual generation of *Andricus quercuslaurinus* (Melika & Pujade-Villar) on branches of *Quercus affinis* Scheidw.. Initial stage of hypertrophy (A). Hyperplasia stage (B). Mature gall (C). Acaxochitlán, Hidalgo. Photography: Cibrián-Tovar, D. (13th March 2015).



Exploratory analysis of data associated with cynipid attack was essential to understand the spatiotemporal distribution patterns of these species, characterise and quantify the damage they cause to host species. The populations of invasive (Hurley *et al.* 2023) and native cynipids had a high potential for population explosion (Eliason and Potter 2000, Eliason and Potter 2001, Melika *et al.* 2009, Pujade-Villar 2017, Pujade-Villar *et al.* 2014). Even though the environmental and/or anthropic factors related to the phenomenon of the growth of local cynipid populations had not been completely studied, it was extremely necessary to anticipate possible contingencies with studies that allowed establishing appropriate forest management strategies. For this reason, the present study added to previous studies with *A. quercuslaurinus* (Barrera-Ruiz *et al.* 2016, Melika *et al.* 2009) to fill in some knowledge gaps.

Prevention and management depend on accurate information about the incidence of a pest within the host population (Adame *et al.* 2022, Heybroek *et al.* 1982, Painter 1951).

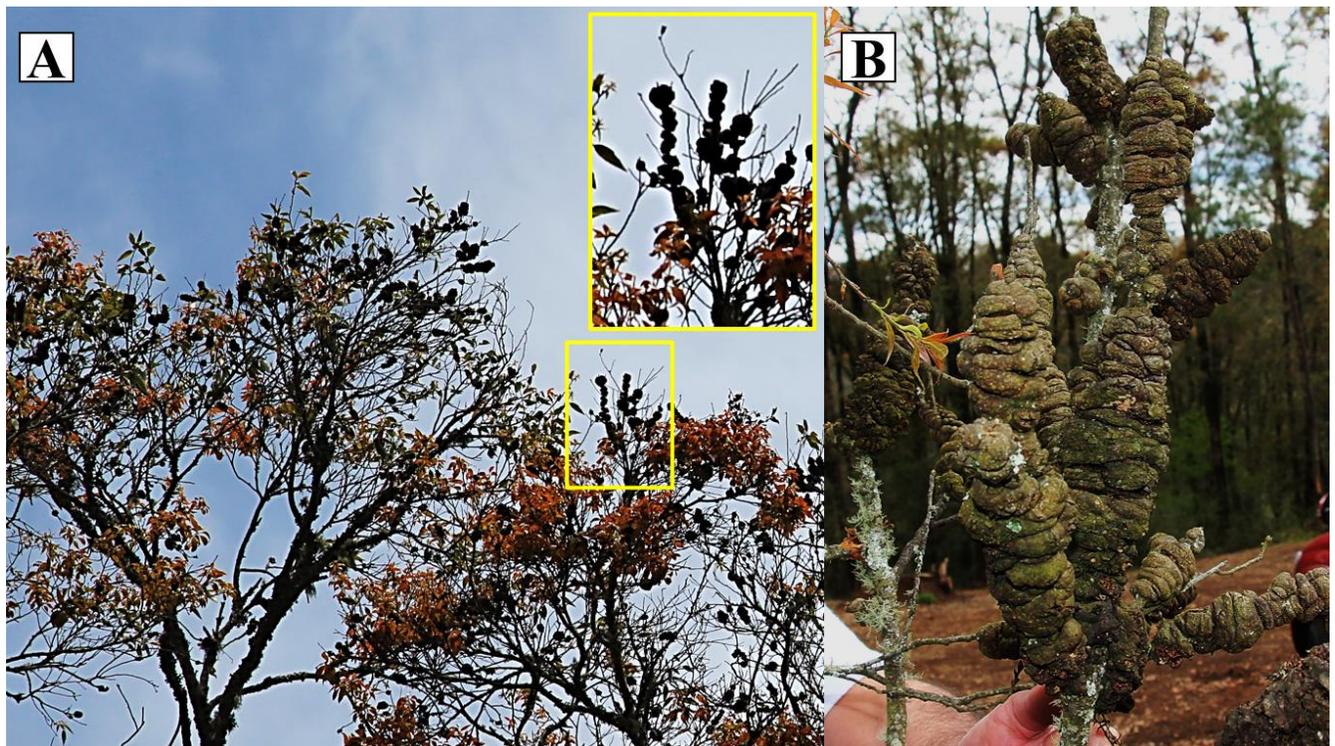
The detailed exploratory analysis of the variables that represent the incidence of a pest allows inferring about the vulnerability, resistance, or tolerance of host plants to attacks (Gibson *et al.* 1982, Pike *et al.* 2021, Sniezko and Koch 2017). The present analysis of the incidence of *A. quercuslaurinus* on *Q. affinis* forests served as a basis for future studies related to conservation and forest management programmes of the genus *Quercus* in Mexico. The identification of resistant, tolerant and vulnerable individuals through the execution of sampling and statistical analysis methods and techniques such as those carried out in the present study constitutes a powerful tool for forest genetic improvement, because it allows reducing the uncertainty related to the selection of individuals that truly had some degree of heritability of resistance or tolerance characteristics to priority pests and diseases, a topic supported in specialised literature (Gibson *et al.* 1982, Guyot *et al.* 2015, Heybroek *et al.* 1982, Painter 1951, Politowski and Browning 1978, Sniezko and Koch 2017).

The asymmetry in the distribution of the original data of the variables (Figures 3 and) was of the type of right bias and directly influenced the results of the parametric (Figure 2) and non-parametric inferential tests (Figures 4 and 7). In this context, it could be said that the right bias was a characteristic of the probability distribution of the variables related to the incidence of the oak gall wasp.

Inferential tests are not always appropriate for the available sampled data set. For this reason, many insect pest incidence studies usually omitted the previous stages of verifying the distribution of data associated with the main variables. For example, when the requirements of normality and homoscedasticity are violated, the results of parametric tests can mainly induce type I and II errors. Depending on the distribution of variances between independent samples, non-parametric tests can also induce type I and II errors. The shape of the probability distribution is very important in both cases. In this study, the square root transformation was performed to bring the distribution of the variables of interest closer to the hypotheses of normality and homoscedasticity, that is, it sought to stabilise the variance, reduce the asymmetry, and attenuate the effects of outliers. Additionally, square root transformations can preserve relative magnitudes by maintaining a data scale equivalent to the original absolute magnitude. Table 5 shows how the quantile breaks of the OTLBG and TNLG variables transformed by the square root, when they undergo the inverse

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241 transformation, present values very close to the quantile breaks of their original values. Therefore, this study highlights the methodological congruence between inferential and exploratory techniques that are carried out in the same context of analysis. Initial inferential stages performed with variable transformations that greatly change the scale of values can generate incorrect interpretations that in turn will affect subsequent exploratory analysis stages that depend on these interpretations.

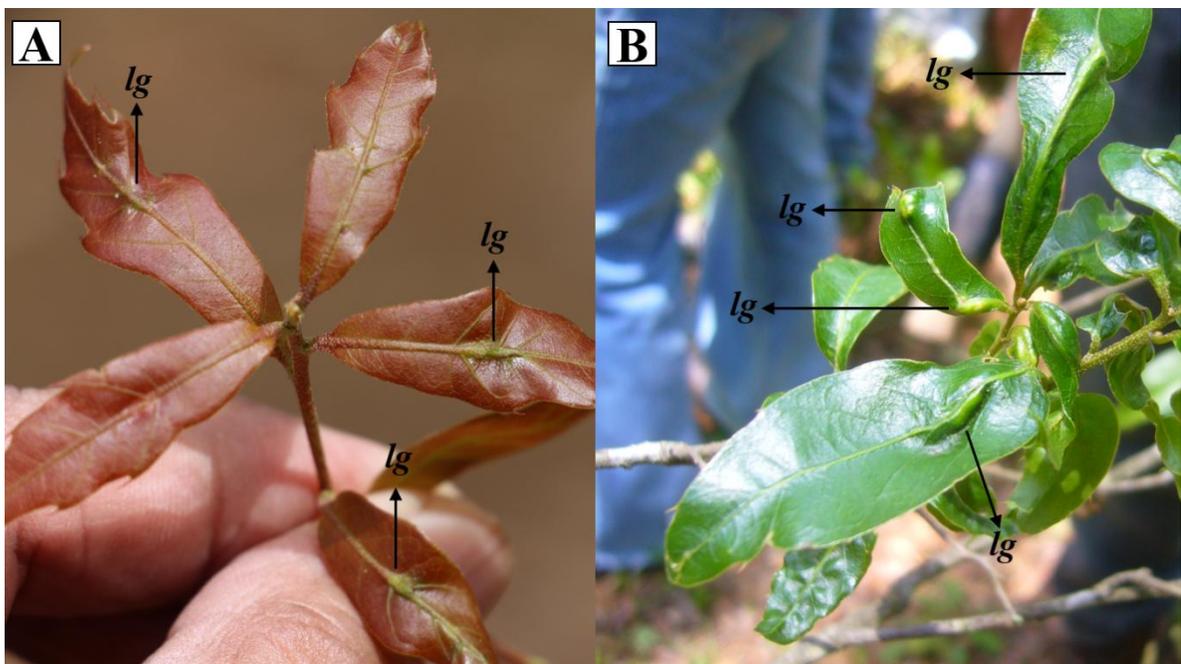
**Figure 10.** Infestation of *Andricus quercuslaurinus* (Melika & Pujade-Villar) in individuals of *Quercus affinis* Scheidw.. Branch galls in the crowns of trees between 6 to 7 meters in total height (A). Branch galls of the asexual generation of the cynipid (B). Acaxochitlán, Hidalgo. Photography: Cibrián-Tovar, D. (13th March 2015).



As definitive class intervals (hypothetically associated with the conditions of resistance, tolerance, and vulnerability of *Q. affinis* to the attack of *A. quercuslaurinus*), only the quantile breaks of the original variables were used (Table 5) in order to represent the distribution of actual data that based the final exploratory analysis (Figure 8). The subdivision of the values of continuous variables by breaks allowed the data to be grouped into significant intervals that modelled the behaviour of this variable (Armstrong *et al.* 2003,

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241 Bivand *et al.* 2023). The GVF index compared the variance predicted by the model based on class intervals with the actual variance of the data, and values closer to one indicated better fits. Likewise, TAI values close to one indicated greater accuracy of the model in terms of adjusting the width of the predicted class intervals. From the point of view of advantages, the quantile method maintained fairness in the distribution of the data and allows the class intervals to adapt to changes in the distribution. However, quantiles may not be as representative when asymmetric distributions have many extreme or outlier values. It should be noted that the prediction must undergo the scrutiny of experts. The class intervals should not only represent the variance gradient of the variable but also a gradient that allowed direct interpretation by users of the method. These associations between levels of phenotypic response through branch and leaf gall formation and genetic conditions or responses (Table 5) are hypothetical because they need to be verified by post hoc genotypic analysis (Kersten *et al.* 2013, Naidoo *et al.* 2019, Pearse *et al.* 2018). The phenotypic selection of individuals by incidence of insect pest damage to the host plant population is the main contribution of this exploratory data analysis.

**Figure 11.** Galls formed by sexual generation of *Andricus quercuslaurinus* (Melika & Pujade-Villar) on immature (A) and mature (B) leaves of *Quercus affinis* Scheidw.. Acaxochitlán, Hidalgo. Photography: Cibrián-Tovar, D. (13th March 2015).



## Conclusions

The high, medium, low, and very low response classes of *Q. affinis* individuals to the attack of *A. quercuslaurinus* of the OTLBG were less sensitive for the diagnosis of resistance, tolerance, and vulnerability characteristics given the evidence of absence of attack occurring in all quantile intervals when these were separated by the agamic cycles of the insect (2012, 2015, and 2018). However, the absence of gall attack decreases as the vulnerability of the trees increases in each of these agamic cycles. This ratio between unattacked and attacked trees with branch gall formation decreased in 2015, where favourable conditions probably existed for the agamic cycle of *A. quercuslaurinus*. On the other hand, the same relationship decreased in the 2018 agamic cycle. Seven percent of the individuals in the branch gall sample had a high probability of association with the hypothesis of genetic resistance to the attack of the oak gall wasp.

The TNLG was more accurate in diagnosing the resistance, tolerance, and vulnerability characteristics of *Q. affinis* individuals in the 2018 gamic cycle. The distribution of the attack of the gall wasp on leaves was better represented by this variable. The absence of attack was an exclusive condition of the high class of response against leaf attack. As the response of the trees to the attack decreases, the occurrence of leaves with a greater number of galls increases. Nine percent of the individuals in the leaf gall sample have a high probability of association with the hypothesis of genetic resistance against the attack of *A. quercuslaurinus*.

## References

- Adame, P., Alberdi, I., Cañellas, I., Hernández, L., Aguirre, A., Ruano, A., Moreno-Fernández, D., Isabel González, A., Torres, M. B., Montes, F. (2022). Drivers and spread of non-native pests in forests: The case of *Gonipterus platensis* in Spanish Eucalyptus plantations. *Forest Ecology and Management*, 510:120104. <https://doi.org/10.1016/j.foreco.2022.120104>
- Armstrong, M.P., Xiao, N., Bennett, D.A. (2003). Using Genetic Algorithms to Create Multicriteria Class Intervals for Choropleth Maps. *Annals of the Association of American Geographers*, 93(3): 595–623. <https://doi.org/10.1111/1467-8306.9303005>
- Barrera-Ruiz, U.M., Cibrián-Tovar, D., Llanderal-Cázares, M. C. M., Cibrián-Llanderal, V. D., Lagunes-Tejeda, Á. (2016). Chemical combat of gall wasps *Andricus quercuslaurinus*

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241

Melika & Pujade-Villar (Cynipidae) in *Quercus affinis* Scheidw. *Revista Chapingo Serie Ciencias Forestales y Del Ambiente*, 22(2): 115–123.  
<https://doi.org/10.5154/r.rchscfa.2015.05.020>

Bivand, R., Ono, H., Dunlap, R., Stigler, M., Denney, B., Hernangómez, D. (2023). Package 'classInt.' R Package Version 0.4-10. <https://cran.r-project.org/web/packages/classInt/index.html>

Cibrián Tovar, D. (2013). Manejo del barrenador de las meliáceas, *Hypsipyla grandella*. In D. Cibrián Tovar ed. Manual para la identificación y manejo de plagas en plantaciones forestales comerciales Texcoco, México. Universidad Autónoma Chapingo, CONAFOR. p. 144–153.

<http://www.conafor.gob.mx:8080/documentos/docs/43/6043Manual%20para%20la%20identificaci%C3%B3n%20y%20manejo%20de%20plagas%20en%20Plantaciones%20Forestales%20Comerciales.pdf>

Cook, J. M., Rokas, A., Pagel, M., Stone, G. N. (2002). Evolutionary shifts between host oak sections and host-plant organs in *Andricus gallwasps*. *Evolution* 56(9): 1821–1830.  
<https://doi.org/10.1111/j.0014-3820.2002.tb00196.x>

Csóka, G., N. Stone, G., Melika, G. (2017). Non-native gall-inducing insects on forest trees: a global review. *Biological Invasions* 19(11): 3161–3181. <https://doi.org/10.1007/s10530-017-1466-5>

Eliason, E.A., Potter, D. A. (2000). Impact of whole-canopy and systemic insecticidal treatments on *Callirhytis cornigera* (Hymenoptera: Cynipidae) and associated parasitoids on pin oak. *Journal of Economic Entomology* 93(1): 165–171. <https://doi.org/10.1603/0022-0493-93.1.165>

Eliason, E., Potter, D. (2001). Biology and management of the horned oak gall wasp on pin oak. *Arboriculture & Urban Forestry* 27(2): 92–101. <https://doi.org/10.48044/jauf.2001.012>

Food and Agriculture Organization of the United Nations [FAO]. (2015). World reference base for soil resources 2014. International soil classification system for naming soils and creating legends for soil maps. 192 p. (World Soil Resources Reports 106).  
<https://www.fao.org/3/i3794en/i3794en.pdf>

García, E. (2001). Climas. Clasificación de Köppen. Modificado por García. Escala 1:1000000. CONABIO.  
<http://geoportal.conabio.gob.mx/metadatos/doc/html/clima1mgw.html>

Gibson, I. A. S., Burley, J., Spiegth, M. R. (1982). The adoption of agricultural practices for the development of heritable resistance to pest and pathogens in forest crops. In H. M. Heybroek, B. R. Stephan, K. von Weissenberg (Eds), Resistance to diseases and pest in forest trees (pp. 9–21). Centre for Agricultural Publishing and Documentation

González-Rodríguez, A., Arias, D.M., Valencia, S., Oyama, K. (2004). Morphological and RAPD analysis of hybridization between *Quercus affinis* and *Q. laurina* (fagaceae), two

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241

Mexican red oaks. *American Journal of Botany*, 91(3): 401–409.  
<https://doi.org/10.3732/ajb.91.3.401>

González-Rodríguez, A., Bain, J. F., Golden, J. L., Oyama, K. (2004). Chloroplast DNA variation in the *Quercus affinis*-*Q. laurina* complex in Mexico: geographical structure and associations with nuclear and morphological variation. *Molecular Ecology*, 13(11): 3467–3476. <https://doi.org/10.1111/j.1365-294X.2004.02344.x>

Govaerts, R., Frodin, D. G. (1998). World checklist and bibliography of fagales (Betulaceae, Corylaceae, Fagaceae and Ticodendraceae). Kew. Royal Botanic Gardens. 407 p.

Guyot, V., Castagneyrol, B., Vialatte, A., Deconchat, M., Selvi, F., Bussotti, F., Jactel, H. (2015). Tree diversity limits the impact of an invasive forest pest. *PLOS ONE*, 10(9): e0136469. <https://doi.org/10.1371/journal.pone.0136469>

Harper, L. J., Schönrogge, K., Lim, K. Y., Francis, P., Lichtenstein, C. P. (2004). Cynipid galls: insect-induced modifications of plant development create novel plant organs. *Plant, Cell & Environment*, 27(3): 327–335. <https://doi.org/10.1046/j.1365-3040.2004.01145.x>

Heybroek, H. M., Stephanand, B. R., von Weissenberg K. (1982). Resistance to diseases and pests in forest trees. In Third International Workshop on the Genetics of Host-Parasite Interactions in Forestry. 14-21 September 1980. 503 p.

Hurley, B. P., Dittrich-Schröder, G., Gevers, C. R. (2023). Gall formers. In J. D. Allison, T. D. Paine, B. Slippers, M. J. Wingfield (Eds), *Forest Entomology and Pathology* (pp. 457–493) Springer International Publishing. [https://doi.org/10.1007/978-3-031-11553-0\\_14](https://doi.org/10.1007/978-3-031-11553-0_14)

Instituto Nacional de Estadística y Geografía [INEGI] (2007). Conjunto de Datos Vectorial Edafológico. Escala 1:250 000 Serie II Continuo Nacional. Carta Edafológica F14-11, Pachuca de Soto. <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=702825235635>

Instituto Nacional de Estadística y Geografía [INEGI] (2018). Conjunto de datos vectoriales de uso del suelo y vegetación. Escala 1:250 000. Serie VII. Conjunto Nacional. <https://www.inegi.org.mx/app/biblioteca/ficha.html?upc=889463842781>

Jactel, H., Moreira, X., Castagneyrol, B. (2021). Tree diversity and forest resistance to insect pests: patterns, mechanisms, and prospects. *Annual Review of Entomology* 66(1): 277–296. <https://doi.org/10.1146/annurev-ento-041720-075234>

Kersten, B., Ghirardo, A., Schnitzler, J. P., Kanawati, B., Schmitt-Kopplin, P., Fladung, M., Schroeder, H. (2013). Integrated transcriptomics and metabolomics decipher differences in the resistance of pedunculate oak to the herbivore *Tortrix viridana* L. *BMC Genomics*, 14(1): 737. <https://doi.org/10.1186/1471-2164-14-737>

Klein, H., Naser, S., Hoffmann, J. H., Dittrich-Schroder, G. (2015). Evidence that *Quadrastichodella nova* (Hymenoptera: Eulophidae) is the only gall inducer among four hymenopteran species associated with seed capsules of *Eucalyptus camaldulensis* (Myrtaceae) in South Africa. *African Entomology*, 23(1): 207–223.

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241

Melika, G., Cibrián-Tovar, D., Cibrián-Llenderal, V. D., Tormos, J., Pujade-Villar, J. (2009). HYMENOPTERA: New species of oak gallwasp from Mexico (Hymenoptera: Cynipidae: Cynipini) – a serious pest of *Quercus laurina* (Fagaceae). *Dugesiana. Revista de Entomología*, 16(2): 67–73. <https://doi.org/10.32870/dugesiana.v16i2.3932>

Naidoo, S., Slippers, B., Plett, J. M., Coles, D., Oates, C. N. (2019). The road to resistance in forest trees. *Frontiers in Plant Science*, 10(273): 1–8. <https://doi.org/10.3389/fpls.2019.00273>

Painter, R. H. (1951). Insect resistance in crop plants. *Soil Science*, 72(6): 481. <https://doi.org/10.1097/00010694-195112000-00015>

Pearse, I. S., Paul, R., O. de P. J. (2018). Variation in plant defense suppresses herbivore performance. *Current Biology*, 28(12): 1981-1986.e2. <https://doi.org/10.1016/j.cub.2018.04.070>

Pike, C.C., Koch, J., Nelson, C. D. (2021). Breeding for resistance to tree pests: successes, challenges, and a guide to the future. *Journal of Forestry*, 119(1): 96–105. <https://doi.org/10.1093/jofore/fvaa049>

Politowski, K., Browning, A. (1978). Tolerance and resistance to plant disease: an epidemiological study. *Phytopathology*, 68(8): 1177. <https://doi.org/10.1094/Phyto-68-1177>

Pujade-Villar, J. (2017). Familia Cynipidae. In D. Cibrián-Tovar (Ed.), *Fundamentos de entomología forestal* (pp. 318–321). Universidad Autónoma Chapingo.

Pujade-Villar, J., Cibrián-Tovar, D., Barrera-Ruiz, U. M., Melika, G. (2014). *Andricus breviramuli* Pujade-Villar n. sp. (Hymenoptera: Cynipidae: Cynipini): New pest of oaks in Mexico. *Southwestern Entomologist*, 39(1): 97–106. <https://doi.org/10.3958/059.039.0110>

Pujade-Villar, J., Equihua-Martínez, A., Estrada-Venegas, E. G., Chagoyán-García, C. (2009). Estado del conocimiento de los Cynipini (Hymenoptera: Cynipidae) en México: perspectivas de estudio. *Neotropical Entomology*, 38(6): 809–821. <https://doi.org/10.1590/S1519-566X2009000600015>

R Core Team (2023). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Ramos-Ortiz, S., Oyama, K., Rodríguez-Correa, H., González-Rodríguez, A. (2015). Geographic structure of genetic and phenotypic variation in the hybrid zone between *Quercus affinis* and *Q. laurina* in Mexico. *Plant Species Biology*, 31(3): 219–232. <https://doi.org/10.1111/1442-1984.12109>

Rohfritsch, O. (1992). Patterns in gall development. In J. D. Shorthouse, O. Rohfritsch (Eds.). *Biology of insect-induced galls* (pp. 60–86). Oxford University Press.

A. Velasco-González //Exploratory Data Analysis for the Classification of *Quercus Affinis*... 213-241

Snieszko, R. A. A., Koch J. (2017). Breeding trees resistant to insects and diseases: putting theory into application. *Biological Invasions*, 19(11): 3377–3400. <https://doi.org/10.1007/s10530-017-1482-5>

Stone, G. N., Schönrogge, K., Atkinson, R. J., Bellido, D., Pujade-Villar, J. (2002). The population biology of oak gall wasps (Hymenoptera: Cynipidae). *Annual Review of Entomology* 47(1): 633–668. <https://doi.org/10.1146/annurev.ento.47.091201.145247>

## Acknowledgements

The authors thank the CONAFOR (National and State Phytosanitary Environmental Contingency Project for the Oak Gall Wasp) and the Foresters Association of the Pachuca-Tulancingo region for supporting this research. The first author thanks Conahcyt for the master's level Postgraduate Scholarship.

## Conflicto de interés

Los autores de este manuscrito declaran no tener ningún conflicto de interés.

## Copyright

La *Revista Latinoamericana de Difusión Científica* declara que reconoce los derechos de los autores de los trabajos originales que en ella se publican; dichos trabajos son propiedad intelectual de sus autores. Los autores preservan sus derechos de autoría y comparten sin propósitos comerciales, según la licencia adoptada por la revista.

## Licencia Creative Commons

Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-Compartir Igual 4.0 Internacional

