

## CÁLCULO DEL TAMAÑO DE LA MUESTRA FOLIAR EN PLÁTANO HARTÓN\*

### CALCULATION OF LEAF SAMPLE SIZE IN PLANTAIN HARTÓN

**Samuel Gustavo Ceballos Pérez<sup>§</sup>, Vianel de Jesús Rodríguez Pérez<sup>1</sup>, Orlando Antonio Rodríguez Rodríguez<sup>1</sup>, Carlos Gómez Cárdenas<sup>2</sup> y Zulime Fatima Rodríguez Guzmán<sup>3</sup>**

<sup>1</sup>Universidad Centroccidental Lisandro Alvarado. Barquisimeto 3001. Venezuela. A. P. 400. (vianelr@ucla.edu.ve), (orlandorodriguez@ucla.edu.ve). <sup>2</sup>INIA-Zulia-Estación Local Chama. El Vigía 5145. Mérida, Venezuela. A. P. 11. (cgomez@inia.gob.ve). <sup>3</sup>Facultad de Agronomía. Maracaibo 5001. Venezuela. (zuliro@fa.luz.edu.ve). <sup>§</sup>Autor para correspondencia: samuelceballos@ucla.edu.ve.

#### RESUMEN

Para realizar el seguimiento o auditoría nutricional al cultivo del plátano Hartón, fueron evaluadas dos alternativas de muestreo de su tejido foliar mediante: a) la muestra preliminar completamente al azar; y b) la muestra estratificada aleatoria, con el objetivo de determinar el marco muestral a fin de obtener el menor tamaño de la muestra en 2008. La unidad de análisis o experimental estuvo, constituida por dos plantas, la planta “madre” al momento de la emisión de la inflorescencia, y su brote lateral o “hijo” en pleno desarrollo. La muestra foliar fue colectada, acorde a las condiciones establecidas en el muestreo internacional de referencia (MEIR). Esa unidad fue identificada con pintura brillante, de manera tal que permitiera identificarla 10 a 14 semanas después y fuese cosechada y pesado su racimo. Los resultados demostraron que el marco muestral generado para muestreo aleatorio estratificado, permite determinar el menor tamaño de la muestra foliar en el ámbito del Sur del Lago de Maracaibo.

**Palabras clave:** azar, muestreo estratificado, racimo de plátano.

#### ABSTRACT

In order to perform the monitoring or nutritional audit to cultivation Hartón plantain, two alternative sampling of leaf tissue were evaluated by: a) the preliminary completely random sample and b) stratified random sample, in order to determine the sampling frame and get the smaller sample in 2008. The analysis or experimental unit consisted of two plants, the “mother” plant at the time of inflorescence emergence and its lateral bud or “son” in full development. The leaf sample was collected according to the conditions laid down in the international reference sample (IRS). That unit was tagged with bright paint, in order to identify it at 10 to 14 weeks later and its cluster were harvested and weighed. Results showed that the sampling frame generated for stratified random sampling, allow us to determine the lowest leaf sample size in the South of Maracaibo Lake area.

**Key words:** banana cluster, random, stratified sampling.

Al realizar un estudio por muestreo, uno de los aspectos más importantes en su diseño, es determinar el tamaño de muestra que se usará cuando se quiere estimar algunos de los parámetros de la población bajo estudio (Álvarez, 1988).

A pesar que se disponen de diversos procedimientos, para la obtención del tamaño de la muestra, desafortunadamente y con frecuencia, el tamaño es decidido arbitrariamente, tomando en consideración sólo el factor económico (Tan, 1996), sin cuantificar la normalidad y la precisión, generalmente bajo el supuesto hecho que la distribución de la población es simétrica o próxima a la normalidad.

De manera general, una muestra de treinta (30) elementos es lo suficientemente grande para asegurar una adecuada aproximación a la distribución normal (sin embargo, puede ser necesario más de 30 elementos para lograr la precisión deseada). El basamento teórico para justificar estos 30 elementos, es el teorema central de límite (TCL), dado que a medida que aumenta la muestra, esta converge en la probabilidad de la distribución normal (Lindgren, 2005).

Sin embargo, cuando priva cuantificar la confiabilidad y la precisión de la media de la población, debe calcularse estrictamente el tamaño de la muestra. Ante esta situación, se sabe que la precisión de la media de la muestra aumenta a medida que incrementa su tamaño  $\left(\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\right)$  y esa precisión dependerá de cuanto el investigador o el asistente técnico, deseen o permitan que el estimador y el parámetro no se alejen entre sí (Nelson, 1999). Esto quiere decir, en términos estadísticos, que el estimador sería no sesgado.

$$E(\theta - \hat{\theta}) = 0 \rightarrow E(\theta) - E(\hat{\theta}) = 0 \rightarrow E(\theta) - \theta = 0 \rightarrow E(\theta) = \theta$$

En este orden de ideas, Jones (1981) señaló que cuando se desea evaluar el estado nutricional de una plantación y establecer valores de referencias locales, cultivos e híbridos, la media aritmética permite evaluar nutricionalmente los cultivos, por los argumentos expresados por Beaufils (2003) sobre las plantaciones con los registros de rendimientos más altos, las cuales nutricionalmente presentan menor variabilidad y distribución normal.

Para realizar el seguimiento o auditoría nutricional al cultivo del plátano Hartón, fueron evaluadas dos alternativas de muestreo de su tejido foliar mediante: a) la muestra preliminar

When making a sample survey, one of the most important aspects in its design is to determine the sample size that will be used to estimate some parameters of the population under study (Álvarez, 1988).

Although there are very different procedures available to obtain the sample size, unfortunately, the size is often determined arbitrarily, taking into consideration only the economic factor (Tan, 1996) without quantifying normality and accuracy, usually under the assumption that the population distribution is symmetric or close to normal.

In general, a sample of thirty (30) elements is large enough to ensure a fair approximation of the normal distribution (however, it may take more than 30 elements to achieve the desired accuracy). The theoretical justification for these 30 elements, is the central limit theorem (CLT), since as the sample increases, this converges in normal distribution probability (Lindgren, 2005).

However, when is necessary to quantify the reliability and accuracy of the population mean, it should be calculated strictly the sample size. In this situation, it is known that the accuracy of the sample mean increases when its size augments  $\left(\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}\right)$  and that accuracy depends on how much do the researcher or technical assistant, want or allow that the estimator and parameter are not away from each other (Nelson, 1999). This means, in statistical terms, the estimator would be unbiased.

$$E(\theta - \hat{\theta}) = 0 \rightarrow E(\theta) - E(\hat{\theta}) = 0 \rightarrow E(\theta) - \theta = 0 \rightarrow E(\theta) = \theta$$

In this same vein, Jones (1981), noted that for evaluating the nutritional status of a plantation and establish local reference values, cultures and hybrids, the arithmetic mean allows to nutritionally evaluate crops, by the arguments expressed by Beaufils (2003), on plantation with records of higher yields, which nutritionally have less variability and normal distribution.

In order to carry out monitoring or nutritional audit to cultivation Hartón plantain, two sampling alternatives of leaf tissue were evaluated by: a) the preliminary sample completely random; and b) stratified random sampling, in order to determine the sampling frame and to get the smaller sample size.

completamente al azar; y b) la muestra estratificada aleatoria, con el objetivo de determinar el marco muestral a fin de obtener el menor tamaño de la muestra.

### **Fecha y localización del estudio**

La muestra preliminar completamente al azar se colectó entre los meses de mayo 2007 y mayo 2008, mientras que la muestra estratificada aleatoria se obtuvo entre los meses de julio a diciembre de 2009. Para ambos casos, el área muestreada se encuentra ubicada en la región de la planicie sur del Lago de Maracaibo, Venezuela, entre el río Mucujepe y el río Chama.

### **Población muestreada**

Para ambos muestreos, la población objeto estuvo constituida por fincas dedicadas al comercio de la exportación, distribuidas aleatoriamente por toda la región. Esas fincas se caracterizan por su homogeneidad en la aplicación de las prácticas agronómicas y contabilización continua de la producción.

### **Definición de la unidad de análisis o experimental**

La unidad de análisis o experimental estuvo constituida por dos plantas, la planta “madre” al momento de la emisión de la inflorescencia, y su brote lateral o “hijo” en pleno desarrollo, con las características arquitecturales de plantas de alto rendimiento descritas por Rodríguez *et al.* (1999).

La muestra foliar en la planta madre fue colectada, acorde a las condiciones establecidas en el muestreo internacional de referencia (MEIR) (Martín-Prevel, 1980a; 1980b). Esa planta fue identificada con pintura brillante, de manera tal que permitiera identificarla 10 a 14 semanas después, y fuese cosechada y pesado su racimo.

### **Definición del marco muestral y de las muestras**

El marco muestral fue definido de la siguiente manera; para el muestreo completamente al azar, se establecieron tres fases: a) elaboración de listado de fincas dedicadas a la exportación; b) selección de los mejores lotes productivos dentro de las fincas listadas; y c) selección dentro de los mejores lotes, las unidades de análisis experimentales usando la tabla de números aleatorios.

### **Date and location of study**

The preliminary and completely random sample, were collected between May 2007 and May 2008, while the stratified random sample was obtained from July to December 2009. In both cases, the sampled area is located in the south plateau region of Lake Maracaibo, Venezuela, between the rivers Mucujepe and Chama.

### **Sampled population**

For both samples, the target population consisted of export-oriented farms, randomly distributed throughout the region. These farms are characterized by their homogeneity in the application of agronomic practices and continuous production accounting.

### **Defining the analysis or experimental unit**

The analysis or experimental unit consisted of two plants, the “mother” plant at the time of inflorescence emergence and its lateral bud or “son” in full swing, with the architectural characteristics of high-yield plants described by Rodríguez *et al.* (1999).

The leaf sample on the mother plant was collected according to conditions laid down in the international reference sample (IRS) (Martín-Prevel, 1980a; 1980b). The plant was tagged with bright paint, so that it could be identified 10 to 14 weeks later and its cluster were harvested and weighed.

### **Defining the sampling frame and samples**

Sampling frame was defined as follows: for completely random sampling, three phases were established: a) lists preparation of export-oriented farms; b) selection of the best production lots on the listed farms; and c) selection among the best lots, the analysis experimental units using the table of random numbers.

For stratified random sample an additional step was included, after the first stage of completely randomized design, which consisted on stratified or grouped farms by soil series, so that identified their production lots over similar ground units (Kijeswski *et al.*, 2001).

Por su parte, para la muestra estratificada aleatoria se incluyó una etapa adicional, posterior a la primera etapa de muestreo completamente al azar, la cual consistió en estratificar o agrupar las fincas por series de suelo, de manera tal que se identificaron sus lotes de producción, sobre semejantes unidades de suelo (Kijeswski *et al.*, 2001).

### Tamaño de la muestra

Para la muestra preliminar completamente al azar, dada la homogeneidad arquitectural de las plantaciones y sus correspondientes unidades experimentales, se propuso muestrear un mínimo de 1% de la superficie cultivada, de la cual el total se aproximaba a 100 000 hectáreas.

En cambio, en la muestra estratificada aleatoria, para asegurar la homogeneidad de los estratos, dada la heterogeneidad del suelo bajo estudio y basándose en la teoría del teorema del límite central, se colectó al azar, un mínimo de 30 muestras en cada unidad de suelos. Sin embargo, en virtud que la cosecha se efectuaría 10 semanas después, se colectaban y marcaban hasta 25% más de unidades experimentales por lote, por razones de seguridad.

### Procesamiento de las muestras

Al momento de colectar cada muestra foliar, se le colocaba en una bolsa de papel, previamente identificada. Una vez en el laboratorio, la muestra era lavada con una solución jabonosa al 1%, enjuagada con agua de chorro y luego con agua destilada, recolocada en bolsas de papel perforadas y limpias y de inmediato, secadas en estufa con ventilación forzada a 70 °C.

Acada muestra se les determinaron las variables nutricionales potasio, calcio, magnesio, cobre, hierro, manganeso y cinc por espectrofotometría de absorción atómica; el nitrógeno y el azufre se determinaron con un analizador de N-S; el fósforo y el boro fueron cuantificados por colorimetría ultravioleta en solución vanado-molibdica y colorimetría con la Azometina-H, respectivamente (Malavolta *et al.*, 1997).

### Análisis de los datos

La prueba de Shapiro y Wilk (2000), se utilizó para determinar la normalidad de las variables nutricionales. Posteriormente por cada variable nutricional se procedió a calcular el nuevo número de muestras (n), con la fórmula de muestreo aleatorio simple (1), para que con un grado de

### Sample size

For preliminary completely random sample, given the architectural homogeneity of the plantations and their corresponding experimental units, a minimum of 1% of the cultivated area was proposed, which consists of approximately 100 000 hectares.

However, to ensure the strata homogeneity in the stratified random sample, given the heterogeneity of studied soil under and based on the theory of the central limit theorem, at least 30 samples were collected at random in each stratum or soil unit. As the harvest would take place 10 weeks later, up to 25% of experimental units per lot were collected and marked, for security reasons.

### Sample processing

At the time of collecting each leaf sample, it was placed in a paper bag previously identified. Once in the laboratory, the sample was washed with a 1% soap solution, rinsed with tap water and then with distilled water, then sample was replaced in perforated and clean paper bags and immediately dried in an oven with forced ventilation at 70 °C.

Nutritional variables were measured in each sample: potassium, calcium, magnesium, copper, iron, manganese and zinc, by atomic absorption spectrophotometry; nitrogen and sulfur were determined with an N-S analyzer; phosphorus and boron were quantified by UV colorimetry in vanado-molybdic solution and colorimetry with azomethine-H, respectively (Malavolta *et al.*, 1997).

### Data analysis

The test of Shapiro and Wilk (2000), was used to determine the normality of nutritional variables. Then for every nutritional variable, the new number of samples (n) was calculated, with the simple random sampling formula (1), so that with a confidence of 95%, the estimate of the farms proportion would not differ more than 0.1 of the true value. It follows that:

$$P(|\hat{\theta} - \theta| \leq 0.1) = 0.95 \leftrightarrow P(-0.1 \leq \hat{\theta} - \theta \leq 0.1) = 0.95 \leftrightarrow$$

$$P\left(\frac{-0.10}{\sigma(\hat{\theta})} \leq \frac{\hat{\theta} - \theta}{\sigma(\hat{\theta})} \leq \frac{0.1}{\sigma(\hat{\theta})}\right) = 0.95$$

confianza de 95% la estimación de la proporción de fincas no difiriera de más del 0,1 del valor verdadero. De lo anterior se deduce que:

$$P(|\hat{\theta} - \theta| \leq 0.1) = 0.95 \leftrightarrow P(-0.1 \leq \hat{\theta} - \theta \leq 0.1) = 0.95 \leftrightarrow$$

$$P\left(\frac{-0.10}{\sigma(\hat{\theta})} \leq \frac{\hat{\theta} - \theta}{\sigma(\hat{\theta})} \leq \frac{0.1}{\sigma(\hat{\theta})}\right) = 0.95$$

$$P\left(\frac{-0.1}{\sigma(\hat{\theta})} \leq N(0,1) \leq \frac{0.1}{\sigma(\hat{\theta})}\right) = 0.95$$

Donde:  $\frac{0.1}{\sigma(\hat{\theta})} = \pi_{\alpha} = 1.96 \rightarrow \sigma(\hat{\theta}) = \frac{0.10}{1.96} = 0.051$

Así tenemos que el problema se traduce en calcular el tamaño de muestra necesario para cometer un error de muestreo de 0.051 al estimar la proporción de fincas.

Para que el tamaño de muestra sea máximo hacemos  $\theta(1 - \theta) = \frac{1}{2}$

$$n = \frac{N\theta(1 - \theta)}{\theta(1 - \theta) + (N - 1)e^2}$$

Donde:  $e = 0.051$  La precisión utilizada fue de 5, 10 y 20% de la media ( $\bar{X}$ ) ( $d_1 = 0.05\bar{X}$ ,  $d_2 = 0.1\bar{X}$ ,  $d_3 = 0.2\bar{X}$ ; por lo tanto tenemos que:

$$n_0 = \frac{t^2 S^2}{d^2} \quad (1)$$

Donde:  $n_0$  = tamaño de la muestra calculada;  $t_n$  = distribución t-estudent;  $S$  = desviación estándar de la muestra aleatoria;  $d$  = precisión.

Después el tamaño obtenido, es corregido ( $n$ ) aplicando la fórmula (2), considerando la dimensión de la subpoblación finita ( $N$ ):

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad (2)$$

Donde:  $n$  = tamaño de la muestra corregida;  $n_0$  = tamaño de la muestra calculada;  $N$  = tamaño de la población finita.

$$P\left(\frac{-0.1}{\sigma(\hat{\theta})} \leq N(0,1) \leq \frac{0.1}{\sigma(\hat{\theta})}\right) = 0.95$$

Is as follows:  $\frac{0.1}{\sigma(\hat{\theta})} = \pi_{\alpha} = 1.96 \rightarrow \sigma(\hat{\theta}) = \frac{0.10}{1.96} = 0.051$

Thus we have that the problem is to calculate the sample size needed to commit a sampling error of 0.051 when estimating the farms proportion.

For maximum sample size:  $\theta(1 - \theta) = \frac{1}{2}$

$$n = \frac{N\theta(1 - \theta)}{\theta(1 - \theta) + (N - 1)e^2}$$

Where: Accuracy used was 5, 10 and 20% of the mean ( $\bar{X}$ ) ( $d_1 = 0.05\bar{X}$ ,  $d_2 = 0.1\bar{X}$ ,  $d_3 = 0.2\bar{X}$ ; therefore we have:

$$n_0 = \frac{t^2 S^2}{d^2} \quad (1)$$

Where:  $n_0$  = calculated sample size;  $t_n$  = Student's t distribution;  $S$  = standard deviation of random sample;  $d$  = precision.

Then, the size obtained is corrected ( $n$ ) using the formula (2), considering the dimension of finite subpopulation ( $N$ ):

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad (2)$$

Where:  $n$  = corrected sample size;  $n_0$  = calculated sample size;  $N$  = finite population size.

### Completely random preliminary sample

1185 experimental units were collected (finite subpopulation ( $N$ )), representing an area of 1 800 ha; however, out of this subpopulation only 410 random samples were analyzed, when we returned 10 to 14 weeks later for weighing the clusters, they had disappeared for several casualties, such as floods, high winds or vandalism.

The results of cluster weights, allow to qualify the sampling unit with a high yield, as reported by Haddad *et al.* (2008) (Table 1). However, regardless of such high

### Muestra preliminar completamente al azar

Se colectaron 1185 unidades experimentales (subpoblación finita (N)), representando una extensión de 1 800 ha; sin embargo, de esa subpoblación se analizaron sólo 410 muestras aleatorias, dado que al regresar 10 a 14 semanas después, para proceder al pesaje del racimo, estos habían desaparecido por diversas causas fortuitas, tales como inundaciones, vientos fuertes o vandalismo.

Los resultados de los pesos del racimo, permiten calificar a la unidad de muestreo con un alto rendimiento, de acuerdo a lo reportado por Haddad *et al.* (2008) (Cuadro 1). Sin embargo, a pesar de esos altos pesos promedios y además, que toda la plantación es visualmente homogénea y vigorosa, se contraponen contra las variables nutricionales, las cuales por presentar altos coeficientes de variación, indican gran heterogeneidad en la población objetivo (en los nutrientes cobre, manganeso y zinc). Esta situación se puede explicar de acuerdo con Beaufils (2003), quien señaló que existe gran heterogeneidad de las variables nutricionales, mientras el cultivo no expresa aún los más altos rendimientos, lo cual se puede confirmar por los rendimientos en subpoblaciones de hasta 22 kg racimo<sup>-1</sup>, reportados por Rodríguez y Rodríguez (1997), con bajos coeficientes de variación en todos sus nutrientes; sin embargo, no especificó los tipos de suelos.

average weights and also that architecturally the entire plantation is visually very homogeneous and vigorous, it contrasts with the nutritional variables, which have high coefficients of variation, indicating substantial heterogeneity in the target population (markedly in copper, manganese and zinc nutrients). This can be explained according to Beaufils (2003), who noted that there is great heterogeneity of nutritional variables, while the crop does not express higher yield, which can be confirmed by yields in subpopulations up to 22 kg cluster<sup>-1</sup>, reported by Rodríguez and Rodríguez (1997), with low coefficients of variation in all their nutrients, but types of soils were not specified.

The Table 2 shows the results obtained by calculating the corrected sample size (n), with  $t_n$  to 1.96 (degrees of freedom at 5%) at different levels of precision (d). The leaf sample sizes at each level of precision are dissenting when they are discriminated by nutrients, thus they range from units to hundreds of samples. For example, with an accuracy of 0.05x, 11 samples are sufficient to evaluate the foliar nitrogen, while for the same but with zinc, 707 samples are required.

This divergence is too large, unwieldy and impractical for a next sampling with purposes of monitoring or nutritional audit of macro and micro nutrients, that is why it is

**Cuadro 1. Medias, desviación estándar y coeficiente de variación del peso del racimo y de las variables nutricionales.**  
**Table 1. Means, standard deviation and coefficient of variation of cluster weight and nutritional variables.**

Estimadores	Racimo (kg)	Nutrientos									
		N	P	K	Ca	Mg	B	Cu	Fe	Mn	Zn
Normalidad	NS <sup>1</sup>	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
$\bar{X}$	17.4	27.5	2.4	45.8	10.8	3.2	14.1	15.8	96.2	187.1	30.9
S	2.97	2.34	0.73	8.9	4.27	0.86	4.8	10.3	25.43	148.1	33.06
CV (%)	17	8.5	30.9	19.4	39.4	27.1	34.1	65.2	26.4	79.1	106.8

<sup>1</sup>= normalidad de distribución de los datos mediante la prueba Shapiro y Wilk (2000); N, P, K, Ca y Mg en (g kg<sup>-1</sup>) y B, Cu, Fe, Mn y Zn en (mg kg<sup>-1</sup>);  $\bar{X}$ = media aritmética; S= desviación estándar; CV= coeficiente de variación.

En el Cuadro 2, se muestran los resultados obtenidos al calcular el tamaño de la muestra corregida (n), con  $t_n$  a 1.96 (con grados de libertad al 5%) a los diferentes niveles de precisión (d). Los tamaños de muestra foliar, en cada nivel de precisión son divergentes cuando se discriminan por nutrientes, así de esta manera, van desde unidades a centenas de muestras. Por ejemplo, con una precisión del 0.05x, con 11 muestras es suficiente para evaluar el nitrógeno foliar, mientras que para realizar lo mismo pero con el cinc, se requieren 707 muestras.

discarded this procedure as an alternative to calculate taken samples, in order to reconcile with the assumptions that the sample shows the lowest possible error and economically acceptable.

However, it is not completely ruled out the random sampling in extensive sub-populations, because with less accuracy (10 and 20%), useful nutritional information can be obtained for macronutrients, because N, P, K, Ca and Mg are the most demanded for the crop,

**Cuadro 2. Tamaño de la muestra corregida (n), con t= 1.96 (grados de libertad al 5%) y diferentes niveles de precisión (d) de la media ( $\bar{X}$ ).**

**Table 2. Corrected sample size (n), with t= 1.96 (degrees of freedom at 5%) and different levels of precision (d) of the mean ( $\bar{X}$ ).**

Parámetros	Nutrimentos									
	N	P	K	Ca	Mg	B	Cu	Fe	Mn	Zn
$d_1^{(1)}$	0.14	0.01	0.23	0.05	0.02	0.7	0.79	4.81	9.36	1.55
$n_1^{(2)}$	11	131	55	199	103	156	421	98	531	707
$d_2$	0.28	0.024	0.458	0.108	0.032	1.4	1.58	9.63	18.68	3.09
$n_2$	3	36	14	57	28	44	144	27	202	322
$d_3$	0.55	0.048	0.92	0.21	0.06	2.8	3.16	19.3	37.36	6.18
$n_3$	1	9	4	15	7	11	40	7	58	101

<sup>(1)</sup>= precisión:  $d_1= 5\%$ ;  $d_2= 10\%$  y  $d_3= 20\%$  de la media. <sup>(2)</sup>= Tamaño de la muestra corregida:  $n_1= 5\%$ ;  $n_2= 10\%$ ; y  $n_3= 20\%$  de la precisión.

Esa divergencia por ser excesivamente grande, inmanejable y poco práctica para un siguiente muestreo, con fines de seguimiento o auditoría nutricional de macro y micro nutrientes, motivó a descartar este procedimiento como alternativa para calcular las muestras a ser tomadas, para conciliar con las premisas que la muestra presente el menor error posible y económicamente aceptable.

Sin embargo, no es totalmente descartable la muestra aleatoria al azar, en subpoblaciones tan extensas, dado que con menor precisión (10 y 20%) se puede obtener información nutricional de utilidad para los macronutrientes, por el hecho que el N, P, K, Ca y Mg son demandados en mayor cantidad por el cultivo, lo cual implica grandes volúmenes de fertilizantes y enmiendas a ser aplicados, que deben ser constantemente auditados nutricionalmente.

### Muestra estratificada aleatoria

El marco muestral permitió agrupar las unidades de producción más rendidoras, en tres fincas ubicadas en suelos de las unidades Chama # 1 (unidad con 41.114 ha) y padre # 37 (unidad con 29.563 ha), por tener ambas, predominancia del tipo textural franco con alta presencia de limos (Kijeswski *et al.*, 2001).

En cada finca se muestrearon 38 unidades experimentales, para un total de 114 [subpoblación finita (N)], provenientes de 160 ha; sin embargo, fue menor el número de muestras analizadas, por las mismas razones adversas, ya expuestas en el muestreo aleatorio (Cuadro 3) y en el Cuadro 4, se presentan los resultados obtenidos del análisis foliar.

implying large amounts of fertilizers and amendments to be applied, which must be constantly nutritionally audited.

### Stratified random sample

The sampling frame allowed grouping the most yielding production units, on three farms in soils from units Chama #1 (unit with 41,114 ha) and parent #37 (unit with 29,563 ha), they both have predominance of loam textural type with high presence of silt (Kijeswski *et al.*, 2001).

In each farm 38 experimental units were sampled, for a total of 114 [finite subpopulation (N)], from 160 ha, however, the number of analyzed samples was lower, for the same adverse reasons already discussed in the random sampling (Table 3) and Table 4 presents results of foliar analysis.

### Cuadro 3. Fincas, tipos texturales y número de muestras.

**Table 3. Farms, textural types and number of samples.**

Fincas	Tipos texturales	Muestras	
		Colectadas	Analizadas
Janeiro	F, Fa, FL, L	38	25
Bella Vista	F, Fa, L	38	31
La Fortuna G	F, Fa, FL	38	11
Total		114	67

When comparing foliar analysis results in Table 1, highlights the decrease in the coefficient of variation in nutrients: copper, manganese and zinc, which was a condition sought through stratification. The comparison of these results, showed the high influence of soils on the variability of nutritional composition of the plant.

Al comparar los análisis foliares, con resultados del Cuadro 1, se resalta la disminución del coeficiente de variación en los nutrientes cobre, manganeso y zinc, la cual fue una condición deseable o buscada a través de la estratificación. La comparación de estos resultados, demostraron la alta influencia que tienen los suelos sobre la variabilidad de la composición nutricional de la planta. Kijeswski *et al.* (2001) identificó unidades de suelos desde la fracciones gruesas hasta las más finas, a pesar que la plantación presentó condiciones muy semejantes.

En esta oportunidad, los pesos del racimo (Cuadro 4), calificarían a la unidad de muestreo estratificada, como de bajo rendimiento, de acuerdo a lo reportado por Haddad *et al.* (2008), pero esta vez, toda la plantación es visualmente muy homogénea, vigorosa y nutricionalmente el coeficiente de variación indica menor heterogeneidad. Estos resultados, no coinciden con lo expuesto por Beaufils (2003), por lo tanto se ratifica que la variabilidad nutricional no depende completamente de los pesos del racimo; por lo tanto, el suelo tiene una fuerte contribución y para este cultivo debe utilizarse este marco muestral con fines de diagnóstico nutricional.

**Cuadro 4. Normalidad, medias ( $\bar{X}$ ), desviación estándar (S) y coeficiente de variación (CV) del peso del racimo y los nutrientes estudiados.**

**Table 4. Normality, mean ( $\bar{X}$ ), standard deviation (S) and coefficient of variation (CV) of the cluster weight and studied nutrients.**

Parámetros	Racimo (kg)	Nutrientes									
		N	P	K	Ca	Mg	B	Cu	Fe	Mn	Zn
Normalidad	NS <sup>(1)</sup>	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS
$\bar{X}$ <sup>(2)</sup>	14.3	25.7	2	42.3	6.8	2.8	13.9	8.5	63.1	88.8	19.2
S	3.13	2.41	0.21	5.9	2.24	0.55	3.8	1.44	25.76	26.41	7.35
CV	21.9	9.4	10.8	13.9	32.7	19.3	27.3	17.1	40.8	29.7	38.3

<sup>(1)</sup>= normalidad de distribución de los datos mediante la prueba Shapiro y Wilk (2000); <sup>(2)</sup>= Media del N, P, K, Ca y Mg en (g kg<sup>-1</sup>) y B, Cu, Fe, Mn y Zn en (mg kg<sup>-1</sup>).

En el Cuadro 5 se observan los resultados obtenidos al calcular el tamaño de la muestra corregida (n), con t= 1.96 (grados de libertad al 5%) a diferentes niveles de precisión (d). En la misma, se analiza en la muestra corregida (n), el número de muestras a recolectar posteriores y necesarios para estimar la media con 95% de probabilidad en cada variable.

Los resultados del Cuadro 5, en cada nivel de precisión van desde un número de muestras por el orden de unidades hasta decenas, los cuales son números más manejables a la hora

Kijeswski *et al.* (2001), identified soil units from the coarse fractions to the finest, although planting showed very similar architectural conditions.

This time the cluster weights (Table 4), would qualify the stratified sampling unit as low yield, as reported by Haddad *et al.* (2008), but this time, architecturally the entire plantation is visually homogeneous, vigorous and nutritionally the coefficient of variation indicates less heterogeneity. These results do not coincide with those of Beaufils (2003), thus it is confirmed that nutritional variability does not depend entirely on the clusters weights; therefore, the soil has a strong contribution and for this crop this sampling frame should be used for nutritional diagnosis.

The Table 5 shows results obtained by calculating the corrected sample size (n), with t= 1.96 (degrees of freedom at 5%) at different levels of precision (d). In the same is observed the corrected sample (n), the number of samples needed to be collected in subsequent samples to estimate the mean with 95% probability for each variable.

The results in Table 5 at each level of precision, ranging from a number of samples in the order of units up to ten, which are more manageable numbers when deciding

the sample size. However, if sampling was for research purposes, the accuracy level of 0.05x, would be limited by the sample size of boron and iron. However, if the statistical rigor is less important, as when carrying out a monitoring of nutritional status for purposes of recommending fertilizer, the representativeness of all nutrients would be satisfied with 15 units, with an acceptable precision of 0.2 $\bar{X}$ . This low number of subsamples, motivated to select and recommend this procedure as an alternative to calculate samples to be taken, maintaining the lowest or acceptable possible error in bananas crop.



**Cuadro 5. Tamaño de la muestra corregida (n), con  $t_n = 1.96$  (grados de libertad al 5%) y diferentes niveles de precisión (d) de la media ( $\bar{X}$ ).****Table 5. Corrected sample size (n), with  $t_n = 1.96$  (degrees of freedom at 5%) and different levels of precision (d) of the mean ( $\bar{X}$ ).**

Parámetros	Nutrimentos									
	N	P	K	Ca	Mg	B	Cu	Fe	Mn	Zn
$d_1^{(1)}$	0.12	0.01	0.21	0.034	0.014	0.69	0.42	3.15	4.44	0.95
$n_1^{(2)}$	12	16	24	68	39	80	33	80	63	77
$d_2$	0.26	0.02	0.423	0.068	0.028	1.39	0.84	6.31	8.88	1.92
$n_2$	3	4	7	31	13	44	11	42	27	39
$d_3$	0.51	0.04	0.85	0.13	0.06	2.7	1.69	12.6	17.8	3.83
$n_3$	1	1	2	10	4	15	3	15	8	13

<sup>(1)</sup>= precisión:  $d_1 = 5\%$ ;  $d_2 = 10\%$ ;  $d_3 = 20\%$  de la media; <sup>(2)</sup>= tamaño de la muestra corregida:  $n_1 = 5\%$ ;  $n_2 = 10\%$ ;  $n_3 = 20\%$  de la precisión.

de decidir formular el tamaño de la muestra. Sin embargo, si el muestreo fuere con fines de investigación, el nivel de precisión de  $0.05x$ , estaría limitado por el tamaño de muestra del boro y el hierro. En cambio, si la rigurosidad estadística es menos importante, como ocurre cuando se realiza un seguimiento del estado nutricional con fines de recomendar fertilizantes, la representatividad de todos los nutrimentos, estaría satisfecha con 15 unidades, con una precisión aceptable de  $0.2\bar{X}$ . Ese bajo número de submuestras, motivó a seleccionar y recomendar este procedimiento, como alternativa para calcular las muestras a ser tomadas, manteniendo el menor error posible o aceptable en el cultivo de plátano.

## CONCLUSIONES

La muestra estratificada aleatoria, permite determinar el menor tamaño de la muestra, la cual facilita el seguimiento o auditoría nutricional al cultivo del plátano Hartón en el ámbito del Sur del Lago de Maracaibo.

## LITERATURA CITADA

Álvarez, C. V. M. 1988. Tamaño de muestra: procedimientos usuales para su determinación. Tesis de Maestro en Ciencias. Colegio de Posgraduados, Chapingo, México. 161p.

Beaufils, E. R. 2003. Diagnosis and recommendation integrated system (DRIS): a general scheme for experimentation and calibration based on principles developed from research in plant nutrition. *Soil Sci.* 1:1-132.

## CONCLUSIONS

Stratified random sample, allows determining the smallest sample size; facilitating nutritional monitoring or auditing of banana Hartón crop in the area South of Lake Maracaibo.

*End of the English version*



Haddad, O. W. y Machado, del V. R. 2008. Un índice para evaluar el vigor en las musaceas comestibles para el bosque seco tropical. *Fruits*. 49(1):47-60.

Jones, C. 1981. Proposed modifications o the diagnosis and recommendation integrated system (DRIS) for interpreting plant analyse. *Communication in soil science and plant analysis*. 12(8):785-794.

Kijeswski, J. J.; Colina, P.; Slegmager, J. M. y Bojanowski, Z. 2001. Estudio de suelos semidetallado, sector río Mucujepe-río Escalante. Zona Sur del Lago de Maracaibo. Serie informes técnicos. Zona 5-IT-156. MARNR. Maracaibo. 278 p.

Lindgren, B. 2005. *Statistical theory*. 2<sup>da</sup>. edition. The Macmillan Company. New York, USA. 259 p.

Malavolta, E. G.; Vitti, G. y De Oliveira, S. 1997. Avaliação do estado nutricional das plantas. Princípios e aplicações. 2<sup>da</sup>. edição. POTAFOS. Piracicaba, SP, Brasil. 73 p.

Martin-Prével, P. 1980a. La nutrition minerale du bananier dans le monde. Première partie. *Fruits* 35(9):503-518.

Martin-Prével, P. 1980b. La nutrition minerale du bananier dans le monde. Deuxième partie. *Fruits*. 35(10):583-593.

- Martin-Prével, P.; Lacoëuilhe, J. J. and Marchal, J. 2006. Orientations du diagnostic foliaire de bananier. *Fruits*. 24(3):153-161.
- Nelson, L. 1999. Estadística en la investigación del uso de fertilizantes. INPOFOS. Quito, Ecuador. 66 p.
- Rodríguez, V.; Bautista, D.; Rodríguez, O. y Díaz, L. 1999. Relación entre el balance nutricional y la biometría del plátano (*Musa* AAB subgrupo plátano cv. Hartón) y su efecto sobre el rendimiento. *Revista de la Facultad de Agronomía de la Universidad del Zulia*. 16(4):425-432.
- Rodríguez, V. y Rodríguez, O. 1997. Normas foliares DRIS para el diagnóstico nutricional del plátano (*Musa* AAB subgrupo plátano cv. Hartón). *Revista de la Facultad de Agronomía de la Universidad del Zulia*. 14(6):285-296.
- Shapiro, S. S. and Wilk, M. B. 2000. An analysis of variance test for normality (complete samples). *Biométrica*. 52:591-611.
- Tan, K. 1996. Soil sampling, preparation, and analysis. 10<sup>th</sup> edition. Marcel Dekker Inc. New York, USA. 407 p.