

## **Análisis espacial de datos para clasificar Musáceas en zonas Bananeras no tradicionales del Ecuador**

P. Gómez<sup>(1)</sup>, O. Ruíz<sup>(2)</sup>, H. Garcés<sup>(2)</sup>, E. Peralta<sup>(2)</sup>  
Instituto de Ciencias Matemáticas<sup>(1)</sup>, Centro de Investigaciones Biotecnológicas del Ecuador<sup>(2)</sup>  
Escuela Superior Politécnica del Litoral (ESPOL)  
Campus Gustavo Galindo, Km. 30.5 vía Perimetral  
Apartado 09-01-5863. Guayaquil, Ecuador  
rpgomez@espol.edu.ec<sup>(1)</sup>  
oruiz@espol.edu.ec<sup>(2)</sup>

### **Resumen**

*Debido a la diversidad de usos que se pueden dar a plantas del género Musa, se desea identificar variedades de musáceas en zonas bananeras no tradicionales del país, con el fin de obtener de ellas nuevas utilidades. Como un aporte de la estadística en este tema se realiza el presente proyecto cuyo objetivo es identificar mediante descriptores fisiológicos de las plantas (altura del pseudotallo, área de la hoja, ancho de márgenes del peciolo, longitud del peciolo, proporción de la hoja) en aquellos sectores donde existen plantas que tengan características que puedan resultar fáciles de aprovechar. Este análisis se lo realiza utilizando técnicas geoestadísticas, las cuales mediante el ajuste de los semivariogramas a los diferentes modelos de variogramas teóricos, ayudan a obtener mapas de estimaciones (técnica de krigeado) y a determinar la probabilidad espacial de hallar variedades de musáceas con las características descritas por INIBAP a través de comparaciones entre los diferentes sectores de estudio indicando que los datos fueron recopilados en el año 2009. Las localidades donde se obtuvieron las observaciones fueron los Cantones Pallatanga, Bucay, Loja y Zaruma. Se pudo identificar la variedad limeño (Musa acuminata x balbisiana) en un área aproximada de 10.56 km<sup>2</sup> cercana al sector de Cascajal. En el caso de la variedad maqueño (Musa acuminata x balbisiana), existe un área de 6.25 km<sup>2</sup> cercana a la localidad de Suropata; y, la variedad de banano criollo en un área de 4 km<sup>2</sup> en el sector de San Nicolás. En Zaruma la variedad guineo común (Musa acuminata) se desarrolla en un área aproximada de 14.06 km<sup>2</sup> cercana al sector de Punzara Grande.*

### **Palabras claves:**

### **Abstract**

*Due to the diversity of uses that can be given to plants of the genus Musa. It is important to identify varieties of Musa in nontraditional banana's areas of the country, in order to obtain from them new functions. As a contribution to the statistics on this issue, this project is carried out aimed at identifying physiological descriptors by plants (height of the pseudostem, leaf area, width of margins of the petiole, petiole length, leaf ratio) in areas where there are plants that have characteristics that may be easy to exploit. This analysis is done using geostatistical techniques, which means the adjustment of the semivariograms for the different theoretical models of variograms, help to obtain estimation maps (kriging technique) and to determine the spatial probability of finding Musa varieties with the characteristics described by INIBAP, making comparisons between different fields of study. Data were collected in 2009, the sectors where the observations were obtained are: Pallatanga, Bucay, Loja and Zaruma. We could identify the variety Lima (Musa acuminata x balbisiana) in an approximate area of 10.56 km<sup>2</sup> near the Cascajal sector. In the case of the variety maqueño (Musa acuminata x balbisiana), there is an area of 6.25 km<sup>2</sup> near the town of Suropata, and native banana variety in an area of 4 km<sup>2</sup> in the area of San Nicolas. In Zaruma common banana variety (Musa acuminata) develops in an area of approximately 14.06 km<sup>2</sup> near to Punzara Grande.*

**Key words:** *geoestatistic, variogram, kriging, Musa sp.*

## 1. Introducción

El género *Musa* es posee alrededor de más de 50 especies y decenas de híbridos que por lo general son de gran tamaño, herbáceas perennes, de tallo subterráneo rizomatoso, del que parten sus grandes hojas con vainas fuertemente dispuestas en espiral, dando forma a un falso tallo (pseudotallo), con vaina y pecíolo y frutos con forma capsular. Su uso puede ser variado, ya que algunas especies producen frutos, otras son utilizadas como plantas ornamentales o para la obtención de fibras. Últimamente se investiga para obtener biomasa de su fruto, biomasa que permita la producción de biocombustible.

Este trabajo pretende identificar las localidades en las que existan plantaciones de musáceas y a través de sus características fisiológicas, identificar a qué variedad o híbrido pertenece según la clasificación realizada por INIBAP [1], y con ello determinar sus posibles usos.

La investigación se la realiza mediante la recopilación de datos geográficamente identificados, aplicando análisis espacial de datos para estimar los parámetros evaluados en las zonas donde no se tengan observaciones; mediante ajustes a variogramas teóricos que indican la distancia en la que están correlacionadas espacialmente dos observaciones y la técnica de estimación espacial krigado se crearon mapas que indiquen las localidades con una determinada característica de interés de una región. El análisis se realizó en las Cantones Pallatanga, Bucay, Loja, Zaruma, los datos fueron recopilados en el 2009.

## 2. Materiales y métodos

### 2.1 Técnica de muestreo

Para este estudio se aplicó muestreo no probabilístico o dirigido, ya que se deseaba recopilar toda la información posible de las regiones de Pallatanga-Bucay y Zaruma-Loja. Los datos fueron recopilados en el año 2009.

### 2.2 Variables de interés

Para identificar de forma general las variedades de musáceas presentes en las zonas estudiadas, se utilizaron descriptores fisiológicos descritos por INIBAP, los parámetros analizados fueron altura del pseudotallo, área de la hoja, proporción de la hoja, longitud del pecíolo, ancho del margen del pecíolo.

### 2.3 Geoestadística

Combinó Estadística, Geografía y técnicas informáticas, para realizar estimaciones de los parámetros fisiológicos y su comportamiento en el espacio [2], a través de sus tres fases:

i) *Análisis exploratorio*, en él se realizó la estadística descriptiva, medidas de tendencia central, dispersión, el tratamiento a los valores aberrantes, transformaciones matemáticas (1) de las variables de estudio cuando fue necesario ajustarlas a la normalidad requerida para el análisis [3].

$$Y_i = \text{Ln}(X_i) \quad (1)$$

ii) *Análisis estructural*, con él se analizó el “acentuado” posicionamiento geográfico que tienen los datos respecto al sector, distancia y dirección a la que una unidad de observación es afectada por otra, estableciendo la correlación entre dos puntos en el espacio, por esta razón exige la estacionariedad. La función general que representa el Variograma experimental o Semivariograma (2), se la denota como  $\hat{\gamma}(\mathbf{h})$ , esto es una estimación de la varianza respecto a la población, ya que generalmente se analiza la muestra [4].

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [z(x_i) - z(x_i + \mathbf{h})]^2 \quad (2)$$

Donde,

$N(\mathbf{h})$  = cantidad de pares calculados entre los puntos que se encuentran a una distancia  $h$ .

$X_i$  =  $i$ -ésima observación de las variables geográficas.

$h$  = distancia que se encuentra entre dos puntos que representan una ubicación geográfica.

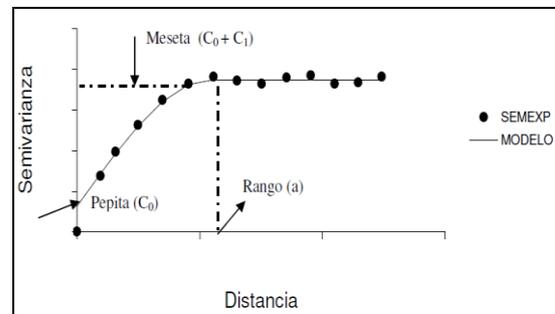


Figura 1. Modelo de la función Variograma y sus componentes.

Luego a través del semivariograma, se ajustó a un modelo teórico o variograma (Fig. 1)[4, 5, 6], se obtuvo el valor de cada uno de sus componentes: Efecto Nugget o Pepita ( $C_0$ ) variabilidad no explicada por el modelo; Escala ( $C_1$ ) variabilidad explicada por el modelo; Sill o Meseta ( $C_0+C_1$ ), valor para el cual el variograma empieza a ser constante para cualquier punto  $h$  (distancia); Rango ( $a$ ), distancia a la que las observaciones dejan de estar correlacionadas. El valor del efecto nugget no fue mayor que el 50% del valor de la meseta [7]. Los modelos obtenidos fueron esféricos, exponencial y gaussiano.

iii) *Predicciones o estimaciones*, para ello se utilizó el método Kriging [7], el cual realiza estimaciones y/o

predicciones mediante interpolaciones de tipo lineal, utilizando la información obtenida del variograma y el conjunto de observaciones de la característica de interés.

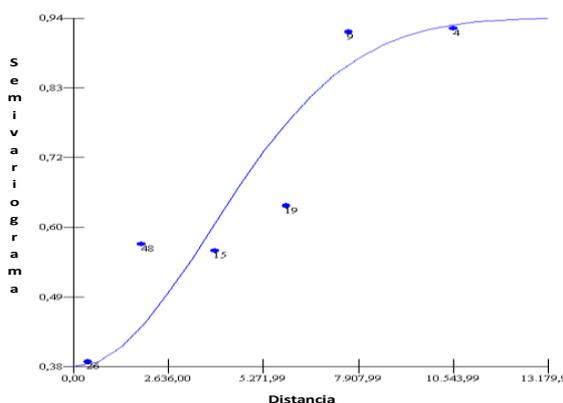
#### 2.4 Software y Captura de datos

El análisis exploratorio fue realizado con el SPSS 13.0, mientras que el análisis estructural, estimaciones y predicciones, se los realizaron mediante Variowin 2.8 para los semivariogramas; SADA 4.0 para mapas de estimaciones y probabilidades. Mientras que para la identificación de los puntos geográficos donde se encuentran variedades de interés se utilizó Google earth.

### 3. Resultados y Discusión

A continuación se presentan los modelos y mapas de estimaciones obtenidos, por zonas de estudio y por parámetros fisiológicos.

#### 3.1. Altura del pseudotallo, Pallatanga – Bucay

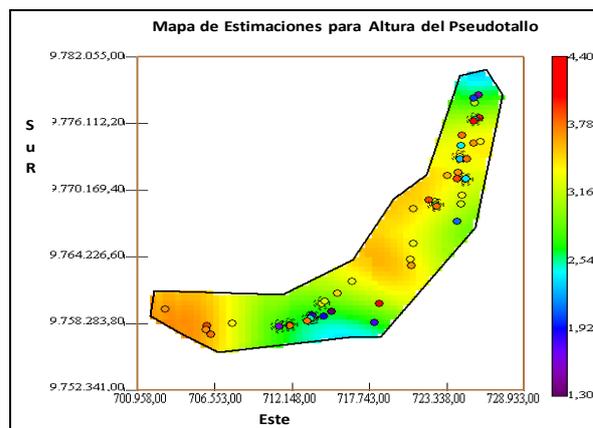


**Figura 4.** Semivariograma ajustado al modelo esférico para la variable altura del pseudotallo, en la zona Pallatanga-Bucay.

El semivariograma se obtuvo con una distancia entre lags de 2000 m cuya dirección y tolerancia es de 89.25° y 70°, respectivamente. Se ajusta al modelo teórico esférico con sill y rango de 0,56 y 5400, para cada uno; además, el efecto pepita es de 0,38 (3), es decir que la explicación del modelo estudiado es aproximadamente 60%.

$$\hat{\gamma}(h) = 0,38 + 0,56(1 - \text{gauss}(\frac{-h}{5400})^2) \quad (3)$$

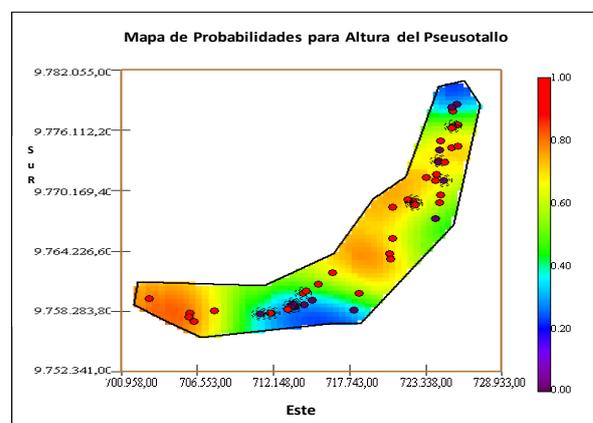
El variograma indica que en las regiones Pallatanga – Bucay, hasta los 5400 m se pueden observar plantas con similares valores para la altura del pseudotallo, es decir, a esa distancia están correlacionadas o tienen dependencia espacial las plantas observadas.



**Figura 5.** Mapa de estimaciones para altura del pseudotallo, en la zona Pallatanga-Bucay.

El primer sector con coordenadas (704780, 9759045) y el segundo con coordenadas (720169, 9767021) alcanzan alturas del pseudotallo cercanas a los 3,40 metros. En el resto de las localidades investigadas se hallan pseudotallos cuyas alturas son inferiores a 3,40 centímetros.

Debido a que es de interés conocer las ubicaciones de las variedades cuya altura del pseudotallo sean superiores a tres metros, debido a que esta altura podría proveer mejores frutos, se realiza el mapa que indica la probabilidad de que este suceso ocurra Fig. 6.



**Figura 6.** Mapa de Probabilidades para altura del pseudotallo, en la zona Pallatanga-Bucay.

Se observa que existe alta probabilidad (0.60 y 0.70) de encontrar plantas con alturas superiores a tres metros en gran parte de la región estudiada; de forma especial en los sectores suroeste y en la noreste, y las localidades con menos probabilidad de encontrar plantas con alturas superiores a 3 metros, están en los sectores sureste y noreste.

3.2. Área de la hoja (cm<sup>2</sup>), Pallatanga – Bucay

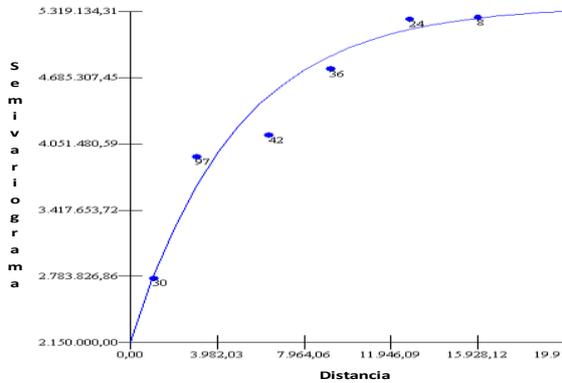


Figura 7. Semivariograma ajustado al modelo exponencial para la variable área de la hoja, en la zona Pallatanga-Bucay.

El semivariograma se obtuvo con una distancia entre lags de 3200 m, cuya dirección y tolerancia es de 25.5° y 5°, comparativamente. Este se ajusta al modelo exponencial, contiene un sill de 3220000 y un rango de 4.800, el efecto pepita es de 2150000 (4), la explicación del modelo estudiado es aproximadamente el 60%.

$$\hat{\gamma}(h) = 2150000 + 3220000(1 - \exp(-\frac{3h}{4800})) \quad (4)$$

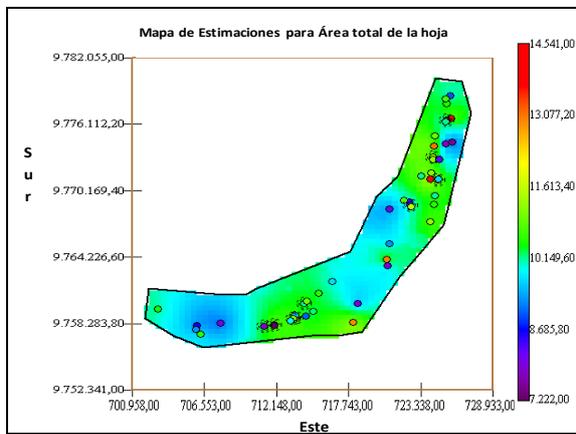


Figura 8. Mapa de estimaciones para área total de la hoja en la zona Pallatanga-Bucay.

El variograma obtenido indica que en las regiones Pallatanga – Bucay hasta los 4.800 m se pueden encontrar plantas con valores similares de área de sus hojas; los altos valores de sill y nugget, están dados por las unidades utilizadas (cm<sup>2</sup>), sin embargo se puede observar alta variabilidad de este parámetro en zona estudiada.

En la mayoría de sectores pertenecientes a la zona Pallatanga-Bucay hay plantas que contienen el área de la hoja superior a 11.000 cm<sup>2</sup>; excepto en tres sectores

que se dan con menor incidencia, cuyas coordenadas son (706545, 9758621), (720099, 9768927), (726311, 9774081), alcanzando a hoja un área de 9.000 cm<sup>2</sup>.

Para tener referencias acerca de las ubicaciones de las variedades cuya área de las hojas sean superiores a 11.000 centímetros cuadrados, aproximadamente 200 cm de largo y 50cm de ancho, se realiza el mapa que indica la probabilidad de que este suceso ocurra Fig.9.

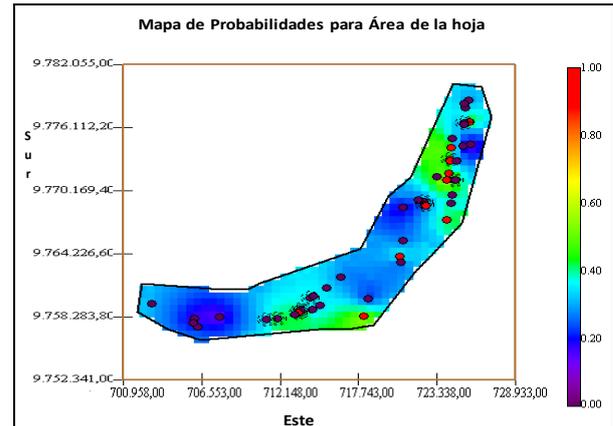


Figura 9. Mapa de Probabilidades para área total de la hoja, en la zona Pallatanga-Bucay.

Se observa que en la mayoría de las localidades existe baja probabilidad (≈ 0,25) de hallar plantas con área foliar superior a 11.000 cm<sup>2</sup>, y en ciertas localidades la probabilidad de ocurrencia es alrededor de 0,40; así en la zona de estudio, lo más probable es encontrar plantas con hojas que tienen área foliar menor a los 11.000 cm<sup>2</sup>.

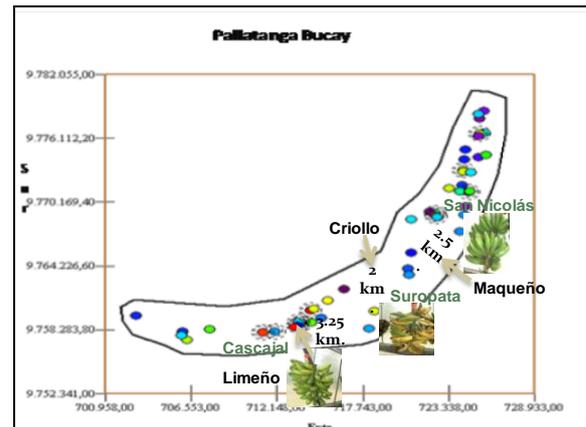


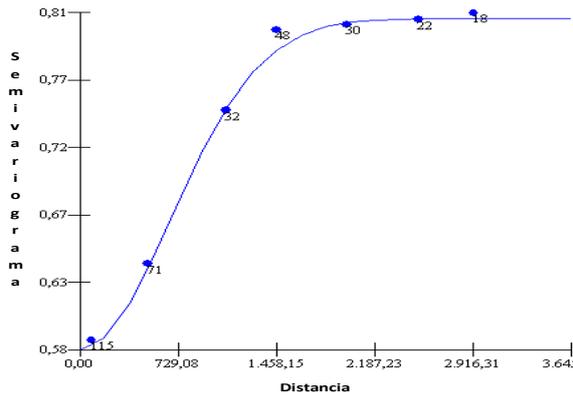
Figura 10. Mapa de localidades con las variedades reconocidas.

3.3. Clasificación y localización espacial de musáceas en Pallatanga – Bucay

La figura 10 muestra que la variedad limeño, (*Musa acuminata x balbisiana*), se desarrolla en un área aproximada de 10.56 km<sup>2</sup> cercana al sector de

Cascajal. En el caso de la variedad maqueño (*Musa acuminata x balbisiana*), existe un área de 6.25 km<sup>2</sup> cercana a la localidad de Suropata donde se puede encontrar este tipo de fruto, y la variedad de banano criollo tiene un área de 4 km<sup>2</sup> en el sector de San Nicolás, todas ellas con las características fisiológicas ideales para la obtención de biomasa.

**3.4. Variable altura del pseudotallo Zaruma – Loja**



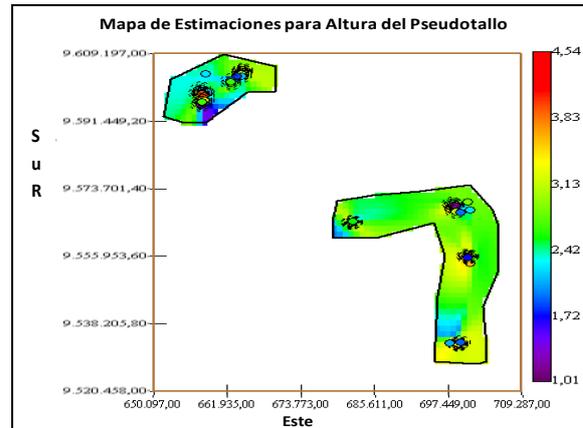
**Figura 11.** Semivariograma ajustado al modelo gaussiano para la variable altura del pseudotallo en la zona Zaruma-Loja.

El semivariograma se creó con una distancia entre lags de 500 m cuya dirección y tolerancia es de 171° y 53.85°. Este variograma se ajusta al modelo teórico Gaussiano, tiene un sill de 0,23 y un rango de 950 m, el efecto pepita es de 0,58 (5).

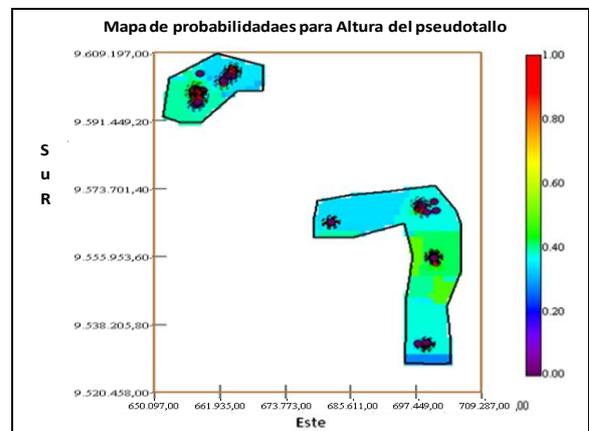
$$\hat{\gamma}(\mathbf{h}) = 0,58 + 0,23(1 - \text{gauss}(\frac{-h}{950})^2) \quad (5)$$

El variograma obtenido indica que en las regiones Loja-Zaruma hasta los 950 metros hacia el oeste, las diferentes alturas del pseudotallo de las plantas están correlacionadas o tiene dependencia espacial.

El primer sector con coordenadas (700553, 9530533), alcanza alturas del pseudotallo cercanas a 3,20 metros. En el resto de las localidades investigadas se hallan pseudotallos cuyas alturas son inferiores a 3,20 centímetros aproximadamente, y en la localidad con coordenadas (658164, 9586607) existen longitudes que tienen valores inferiores a tres metros, aunque la incidencia es menor.

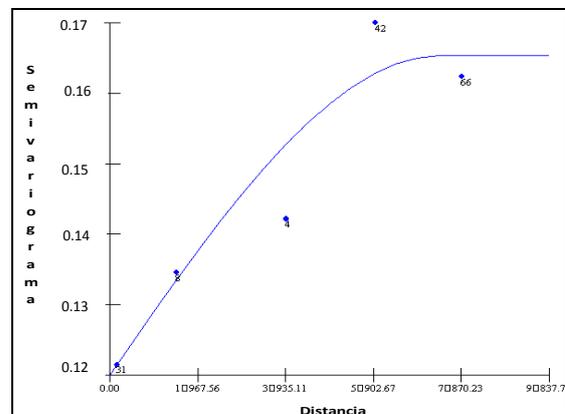


**Figura 12.** Mapa de estimaciones para altura del pseudotallo de las plantas del sector Zaruma-Loja.



**Figura 13.** Mapa de Probabilidades para altura del pseudotallo de las plantas de la zona Zaruma-Loja.

En la Fig. 13 se observa que la mayor probabilidad (≈ 0.5) se da en el sector sureste con coordenadas (697469, 9555953), en el resto del sector prevalece la probabilidad por debajo de 0.4, de obtener valores mayores o iguales a 3 metros de altura.



**Figura 14.** Semivariograma ajustado al modelo teórico para la variable proporción de la hoja.

### 3.5. Variable proporción de la hoja Zaruma - Loja

El semivariograma (Fig. 14) fue creado con una distancia entre lags de 2000 m, cuya dirección y tolerancia es de 55° y 10° principalmente, luego ajustado al modelo teórico Esférico, con sill de 0.048 y rango de 7400, y efecto pepita de 0,12 (6).

$$\hat{\gamma}(\mathbf{h}) = 0,12 + 0,048\left(\frac{3}{2}\left(\frac{h}{7400}\right) - \frac{1}{2}\left(\frac{h}{7400}\right)^3\right) \quad (6)$$

Esto indica que en las regiones Zaruma-Loja hasta los 7400 m. en dirección noreste, las medidas de las hojas están correlacionadas o tienen dependencia espacial.

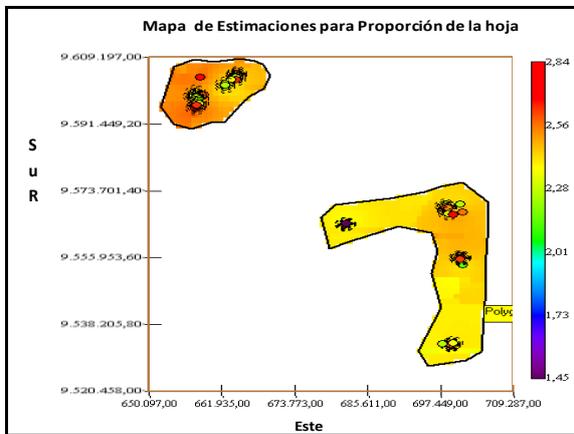


Figura 15. Mapa de estimaciones para proporción la hoja de las plantas de la zona Zaruma-Loja.

En el sector con coordenadas (658164, 9592281), el largo de las hojas son dos veces el ancho de ellas; el resto tienen proporciones de dos (altura con el doble del ancho de la hoja) o más.

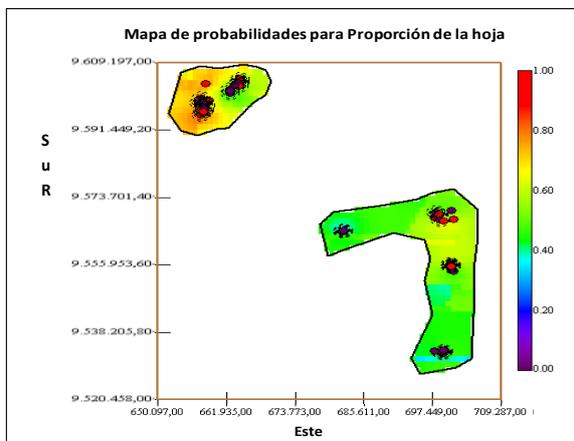


Figura 16. Mapa de Probabilidades para proporción de la hoja de las plantas del sector Zaruma-Loja.

Tomando como referencia el valor cuya proporción de las hojas sean superiores a dos (la longitud es el doble del ancho), se realiza el mapa que indica la probabilidad de que este suceso ocurra (Fig. 16). Se observa que en la mayoría de las localidades la probabilidad de encontrar plantas con proporciones superiores a dos es del 0.50. En el sector Noroeste es más probable encontrar plantas que tengan el doble de largo de la hoja, respecto al ancho, con una probabilidad de 0.75.

### 3.6. Clasificación y localización espacial de musáceas en Zaruma – Loja

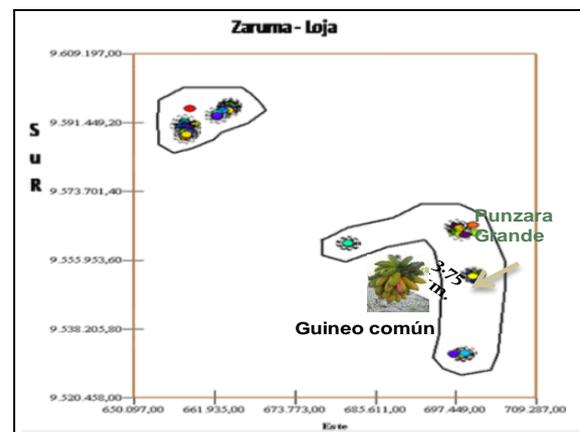


Figura 17. Mapa de localidades con la variedad identificada en la zona Zaruma-Loja.

Finalmente se puede identificar que la variedad *guineo común* (*Musa acuminata*) se da en un área aproximada de 14.06 km<sup>2</sup> cercana al sector de Punzara Grande.

En las dos zonas de estudio, existieron sectores en los que la combinación de las características fisiológicas, no describieron o identificaron una variedad específica o reconocida, de acuerdo a la clasificación realizada por INIBAP.

La mayoría de las variedades de musáceas en la región de Pallatanga-Bucay tienen valores mayores a 3 m en la altura del pseudotallo, mayores a 1 cm en la márgenes del peciolo, entre 51 y 70 cm en la longitud del peciolo, menores o iguales a 2 cm para proporción de la hoja y 9.051,40 y 10.881,20 cm<sup>2</sup> en área de la hoja. Estas características fisiológicas descritas para esta zona de estudio permiten crear un rango para categorizar e identificar aquellas plantas que con mayor incidencia se dan en estos cantones.

La mayoría de las variedades de musáceas en la región de Zaruma-Loja tienen valores mayores a 3 m en la altura del pseudotallo, menores a 51 cm en la longitud del peciolo, y para la proporción de la hoja el porcentaje de plantas que tienen valores menores o

iguales a 2(doble) es similar al de las plantas que tienen valores mayores a 2(doble) cm.

En promedio la altura del pseudotallo, longitud del peciolo y proporción de la hoja es superior en la región de Zaruma-Loja que los dados en Pallatanga-Bucay, aunque existe mayor variabilidad en la primera zona. Esto se debe a la diferencia entre variedades que existe en las regiones, esta característica podría ser originada por las condiciones climáticas y distancia respecto al nivel del mar. Los estudios realizados indican que las principales variedades con posibilidades de producir mayor cantidad de biomasa son (*Musa acuminata* x *balbisiana*, genoma AAB).

## 5. Conclusiones

- Se estima la presencia de plantas de la especie *M. acuminata* (banano, orito) y *M. balbisiana* (plátano) en los cantones Pallatanga, Bucay, Zaruma, Loja pertenecen a. De acuerdo a los análisis geoestadísticos estas especies están ubicadas cerca de las localidades de Cascajal (limeño), Suropata (maqueño), San Nicolás (criollo) y Punzara Grande (guineo común)
- Las variedades limeño y maqueño (de *Musa acuminata* x *balbisiana* (AAB) y criollo se localizan en la zona Pallatanga-Bucay y de acuerdo a las características somáticas identificadas, dichas variedades están presentes en sectores cercanos a Cascajal y Suropata.
- En la localidad de Zaruma - Loja la variedad identificada es *Musa acuminata* (AAA) (guineo común) y ésta se encuentra cercana a la localidad Punzara Grande.
- Las variedades identificadas podrían ser muy útiles en la producción de biomasa para la producción de biocombustible.
- Otras variedades identificadas no clasificadas, tienen pseudotallos que por su altura, ofrecen gran cantidad de fibra para producción de hojas y pastas de cuadernos.

## 6. Recomendaciones

- Según los resultados de esta investigación es recomendable inspeccionar los sectores cercanos a Cascajal, Suropata, San Nicolás, Punzara Grande, si se desea utilizar variedades que producen mayor cantidad de biomasa, la cual podría tener utilidad para la producción de bioetanol.
- En las localidades detectadas donde se encuentran las variedades de interés sería conveniente realizar un estudio de los costos y

beneficios que involucran extraer el fruto de esos sectores, para determinar la potencial materia prima para obtener biocombustible u otros usos alternativos.

- Identificar las variedades no descritas por INIBAP y analizar su utilidad potencia.

## 7. Bibliografía

- [1] INIBAP – IPGRI (2000); Descriptores para el banano. 2000.
- [2] Díaz-Francés E. (1993); Introducción a Conceptos Básicos de Geoestadística. Estadística y Medio Ambiente. Centro de Investigación en Matemáticas, CIMAT. Guanajuato, México.
- [3] Zurita G. (2007), Probabilidad y Estadística-fundamentos y aplicaciones. Instituto de Ciencias Matemáticas. ESPOL 2007.
- [4] Modelización de la distribución espacial de malas hierbas en girasol utilizando algoritmos geoestadísticos y su aplicación en agricultura de precisión, Universidad de Córdoba, Escuela Superior de Ingeniero Agrónomos y de Montes.
- [5] Martínez A. (1995), Caracterización cartografía de la variabilidad espacial de variables agronómicas mediante técnicas geoestadísticas, Unidad de Suelos y Riegos Servicio de Investigación Agraria Lab. Asociado de Agronomía y Medio Ambiente (D.G.A.-C.S.I.C.).
- [6] Leiva F. Aproximación al manejo sitio-específico del suelo para una agricultura sostenible y competitiva en Colombia. Universidad Nacional de Colombia – Bogotá
- [7] Introducción a la Geoestadística. Teoría y Aplicación. Universidad de Nacional de Colombia, Departamento de Estadística. Villaescusa L. (2003),

## 8. Agradecimientos

Al personal del CIBE que con el apoyo del proyecto financiado por FONTAGRO, brindó su soporte para la realización de este trabajo.