

**DISEÑO DE MUESTREO MULTI-ETÁPICO PARA ESTIMACION DE PRODUCCIÓN DEL CULTIVO DE CAÑA DE AZUCAR EN TONELADAS POR HECTÁREA A NIVEL DE SUERTE.**

**MULTI-STAGE SAMPLING DESIGN FOR ESTIMATING SUGARCANE CROP PRODUCTION (tons per hectare) AT PLOT SCALE.**

Yohana Patricia Melo Hernandez.

Héctor A. Chica

Alvaro Gomez Gonzalez

Jefe de Agronomía-Ingenio Pichichi S.A. [ypmelo@ingeniopichichi.com](mailto:ypmelo@ingeniopichichi.com)

Biometrista-Cenicaña. [hachica@cenicana.org](mailto:hachica@cenicana.org)

Gerente de campo-Ingenio Pichichi S.A. [agomez@ingeniopichichi.com](mailto:agomez@ingeniopichichi.com)

**Resumen.**

El sector azucarero colombiano constituye un renglón importante en el ámbito socioeconómico del país, lo que hace necesario garantizar el abastecimiento permanente de caña de azúcar. Sin embargo, existe una brecha entre la producción proyectada antes de cosecha y lo que ingresa a fábrica, situación que puede ocasionar pérdidas que se estiman para un Ingenio Azucarero de 14.000 ha en USD 3.959 por cada hora. Es por esto, que el Ingenio Pichichi S.A en cooperación con Cenicaña realizó, un estudio de muestreo multi-etápico de 3 etapas así: 1. Surco 2. Sitio de 10 m dentro de surco y 3. Tallos dentro de sitio dentro de surco; se determinó el tonelaje de caña por hectárea (TCH) al momento del muestreo. El error estándar asociado al TCH se calculó considerando la incertidumbre del peso medio del tallo y de la población media en cada sitio. Con la estimación temprana del TCH se ajustó un modelo lineal para estimar el TCH al momento de la cosecha de la forma  $TCH = -73.34 + 0.89(TCH\text{-muestreo}) + 31.9\text{Raíz}(\text{Edad Cosecha}) - 14.34I$ , en donde I toma el valor uno si en la suerte se presentó algún incumplimiento de oportunidad de realización en las labores de riego, fertilización y roturación, o cero si no se presentaron. Se encontró que al muestrear en cada suerte 12 surcos, 2 sitios por surco y 5 tallos por sitio, es posible estimar el TCH a los 8 meses de edad con una incertidumbre de  $\pm 5$  TCH. Además, que al predecir el TCH a la cosecha, la incertidumbre es de  $\pm 7.8$  TCH. Se concluye que la metodología es

generalizable debido a que los errores de estimación de la población media de tallos, peso medio de un tallo y TCH al momento del muestreo son independientes del área de la suerte.

### **Abstract.**

The Colombian sugar industry is an important sector in the socio-economic sphere of the country, which is necessary to ensure the continuous supply of sugar cane. However, there is a gap between the projected production before harvest and the actual raw materials that are entering the factory, a situation that can cause losses estimated at about USD 3,959 per hour for a sugar mill of 14.000 hectares. That is why, the Ingenio Pichichi S.A in cooperation with Cenicaña conducted a multi-stage sample study, the 3 stages are as follows: 1. Groove 2. Site 10 m inside groove and 3. Stems in place within groove; the tonnage of cane per hectare (TCH) at the time of sampling was determined. The standard error associated with the TCH was calculated considering the uncertainty of the average weight of the stem and the average population at each site. With early TCH estimation a linear model was fitted to estimate the TCH at the time of harvest form  $TCH = -73.34 + 0.89 (TCH\text{-sampling}) + 31.9Raíz (Age Harvest) - 14.34I$ , where  $I$  takes value one if the plot presented a breach of opportunity during irrigation, fertilization and tillage activities, or value zero if no such breach took place. It was found that by sampling 12 grooves in each Luck, 2 sites in each groove and 5 stems per site, it is possible to estimate the TCH at 8 months old with an uncertainty of  $\pm 5$  TCH. In addition, predicting that the TCH harvest, uncertainty is  $\pm 7.8$  TCH. It is concluded that the methodology is generalizable because the errors of estimation of the average population of stems, average weight of a stem and TCH at the time of sampling are independent of the area of the plot

**Palabras clave:** Predicción, modelo, estimación, variabilidad, muestreo.

**Key Words:** Prediction, model, estimate, variability, sampling

### **Introducción.**

Según el Ministerio de Agricultura y Desarrollo Rural, para el año 2014 el área cultivada en caña de azúcar fue de 230.303 hectáreas (ha) de las cuales 176.244 se encuentran en el departamento del Valle del Cauca, por lo tanto, este cultivo es un renglón importante para la agroindustria y fuente de empleo

del departamento. En el Valle, existen 13 Ingenios productores de azúcar y miel de diferentes calidades, adicionalmente, cinco de estos ingenios cuentan con destilerías para la producción de alcohol carburante.

En Colombia el cultivo de la caña constituye la principal materia prima de la industria azucarera y es necesario que ésta llegue a la fábrica en óptimas condiciones de calidad y en la fecha programada para su molienda, por esta razón, predecir a nivel de parcela (suerte) las toneladas de caña por hectárea (TCH) entre dos a cuatro meses antes de la cosecha garantiza el abastecimiento permanente de la materia prima y se evita la variación en la edad de cosecha con el impacto negativo que esto genera.

En el sector se han desarrollado modelos de predicción de la producción a nivel de finca que incluyen variables de desarrollo y crecimiento y a nivel sectorial que incluyen el comportamiento histórico y variables climatológicas. En el Ingenio Pichichi S.A. esta estimación es realizada por el personal técnico de forma visual, lo que ha representado poca exactitud con respecto a la producción definitiva del cultivo.

Teniendo en cuenta lo anterior, existe la necesidad de generar una herramienta que permita predecir la producción en suertes con una incertidumbre conocida y controlada, con el fin de que el programa de cosecha se articule con las necesidades de fábrica.

En este sentido, el muestreo en campo constituye una alternativa para la determinación de la biomasa existente al momento del muestreo, al respecto Carmona, 2004 evaluó la combinación entre un tamaño de muestra y una forma de conteo de la población de tallos molinables y su correlación con la producción real, obteniendo un coeficiente de correlación de 0,77. Además, es necesario que la estimación de la producción sea sensible a la calidad y la oportunidad de las labores realizadas durante el levante del cultivo, variables que influyen en la producción (Sanchez & Gómez, 2014).

Para el sector azucarero, la estimación del TCH al momento del muestreo y su proyección a cosecha, mediante la implementación de un modelo a nivel de suerte con una incertidumbre menor o igual a 10 TCH y una confiabilidad del 95%, permitirá definir la disponibilidad de la materia prima en el campo con meses de antelación a la cosecha.

## **Materiales y Métodos.**

Este trabajo de investigación se realizó en el Ingenio Pichichi S.A. ubicado en el Municipio de Guacari, Departamento del valle del Cauca - Colombia. Se seleccionaron 62 suertes o parcelas, con áreas entre 1.9 y 18.05 hectáreas y edades entre 7.8 y 10.3 meses después de corte y/o siembra.

La generación del modelo se realizó en dos fases, la primera correspondió al diseño de un método para determinar el TCH al momento del muestreo y la segunda la herramienta para la predicción de TCH a cosecha.

Durante la fase 1, en 29 suertes seleccionadas se realizó un muestreo multietapico piloto de la siguiente manera:

**Etapa 1:** Se determinó el número total de surcos de la suerte y se dividió este valor por 12, el valor obtenido correspondió al número de surcos a contar desde la orilla de la suerte y se continuó consecutivamente el mismo número hasta completar los 12 surcos a muestrear.

**Etapa 2:** En cada uno de los surcos se seleccionaron 2 sitios de 10 metros lineales cada uno, escogidos de forma aleatoria dividiendo la longitud del surco por 10 y sorteando dentro de las 10 posibilidades los sitios aleatoriamente.

**Etapa 3:** Dentro de cada sitio, se tomaron 5 tallos molinables al azar, cortados desde la base hasta el punto natural de quiebre.

Para cada suerte se registró el número total de tallos molinables por cada sitio y el peso de 120 tallos, adicionalmente, se realizó la determinación de la eficiencia (horas/muestreo) en tiempo de la labor. Para estimar el TCH se utilizó la siguiente fórmula:

$$TCH = PT \times PS \times SH$$

Dónde:

TCH: Toneladas de caña por hectárea al momento del muestreo.

PT: Peso medio de un tallo.

PS: Número medio de tallos en 10 metros lineales.

SH: Número de sitios de 10m que hay en una hectárea de acuerdo a la distancia de siembra.

Para la determinación de la incertidumbre a partir de los errores estándar, se definió que la incertidumbre de la estimación del TCH se compone de: las incertidumbres asociadas a la estimación del peso medio del tallo y del número de tallos en sitios de 10m. A su vez, la incertidumbre asociada al peso medio del tallo se compone de la variabilidad entre surcos, de sitios dentro de surco y de tallos dentro de sitio dentro de surco. De igual manera, la incertidumbre asociada a la estimación del número de tallos por 10 m de surco se compone de la variabilidad entre surcos y entre sitios dentro de surco. A cada variabilidad se le conoce con el nombre de componente de varianza y su suma conformó la variabilidad total de cada uno de los indicadores utilizados para estimar el TCH. Para los componentes de varianza se utilizaron modelos anidados y el procedimiento NESTED de SAS System Versión 9.3.

Posterior al muestreo piloto, se colectaron datos en 33 suertes más, para un total de 62 suertes, lo que permitió validar la metodología.

Para calcular el tamaño de la muestra en cada etapa (surcos, sitios/surco y tallos/sitio/surco) se utilizaron los componentes de varianza medianos en cada una de ellas y se creó una hoja de Excel en donde se registró la distancia entre surcos, confiabilidad deseada y número de individuos por etapa a muestrear. El error deseado es el doble del error estándar debido a que la incertidumbre se expresa “más o menos” el error, en este caso más o menos 5 TCH da un rango de 10 TCH.

Para el desarrollo de la fase dos, se colectaron los datos de cosecha de cada una de las suertes muestreadas y se determinó si durante el levante se había presentado algún incumplimiento en los rangos de oportunidad de las labores de roturación, despaje y semanas de déficit hídrico, definidos por el Ingenio Pichichi S.A., los cuales influyen en la productividad del cultivo (Sanchez & Gómez, 2014).

La estimación a cosecha se hizo con el siguiente modelo lineal:

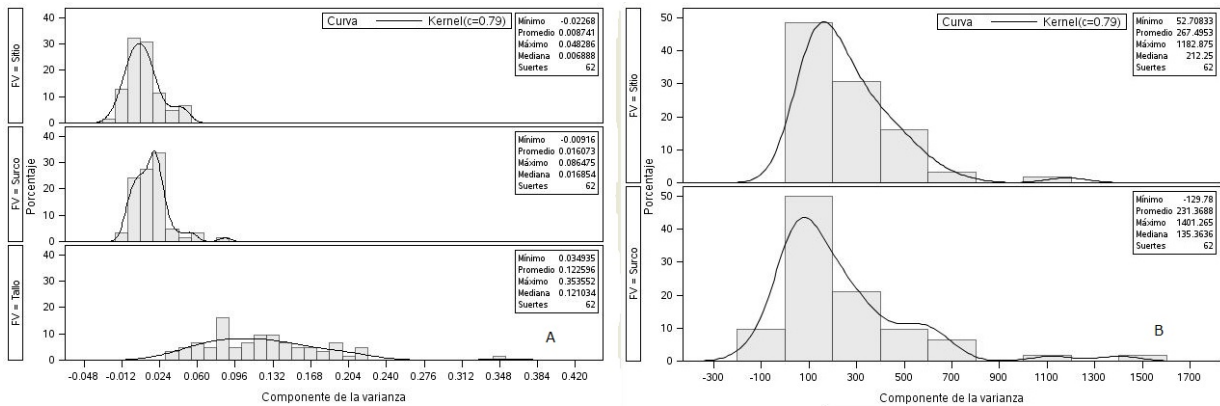
$$\text{TCH(Cosecha)} = \beta_0 + \beta_1 \text{TCH(muestreo)} + \beta_2 \sqrt{2 \times \text{Edad de cosecha}} + \beta_3 I$$

Donde I toma el valor uno cuando en la suerte se presentaron problemas con la ejecución de alguna labor y cero cuando las labores se hicieron oportunamente.

## Resultados y Discusión.

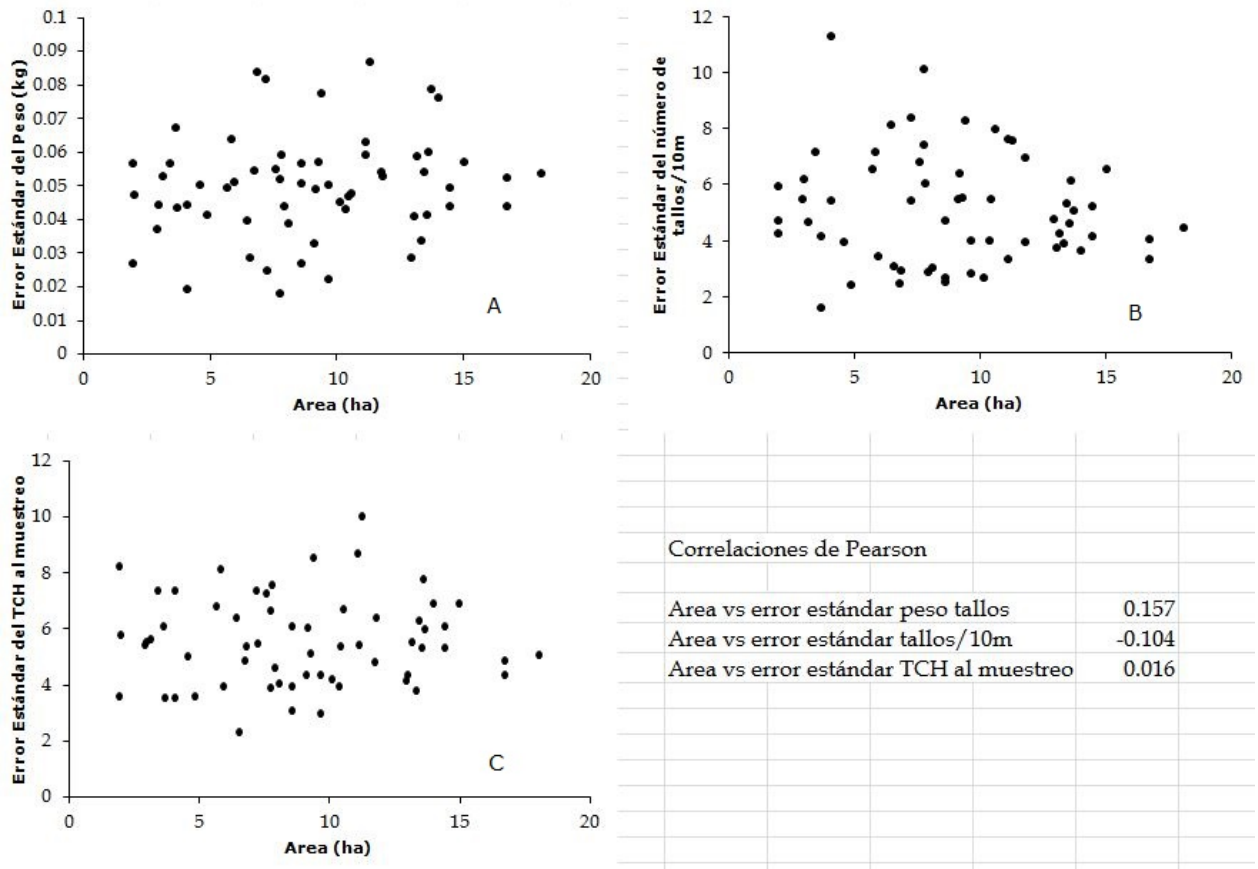
En la figura 1.A se observa que la mayor variabilidad entre los pesos de los tallos se dio a nivel de tallo dentro de sitio (0.12 kg<sup>2</sup>), seguida por la variabilidad entre surcos (0.016 kg<sup>2</sup>) y finalmente entre sitios dentro de surco (0.008 kg<sup>2</sup>). Se encontró además que el desarrollo de los tallos consecutivos de un cultivo puede ser muy homogéneo en unas suertes y muy heterogéneo en otras, mientras que en suertes de diferentes características (edad, área, variedad etc.) el comportamiento entre surcos y entre sitios es similar.

Se observa que la población de tallos en 10 metros varió de una manera similar entre surcos y entre sitios, siendo un poco mayor la variabilidad entre sitios dentro de un mismo surco (figura 1B).



**Figura 1. A.** Distribución porcentual de las componentes de varianza del peso del tallo por cada etapa de muestreo. **B.** Distribución porcentual de las componentes de varianza del número de tallos en 10 metros lineales por cada etapa de muestreo.

Los diagramas de dispersión entre el área y los errores estándar mencionados demuestran que la incertidumbre de las estimaciones del peso medio del tallo, de la población por 10 metros y el TCH calculado a partir de éstas no depende del tamaño de la suerte (Figura 2), por lo que la metodología se puede adoptar independiente del tamaño de la suerte.



**Figura 2.** Dispersión y correlaciones de Pearson entre el área y los errores estándar de: **A.** El peso del tallo **B.** La población en 10 m **C.** El TCH al muestreo

Con la metodología establecida en el muestreo, se encontró que, si se muestrean 12 surcos, 2 sitios por surco y 5 tallos por sitio dentro de surco es posible estimar TCH a los 8 meses de edad con una incertidumbre de  $\pm 5$  TCH. Según Cochran, 1977 citado por Badii et al; 2011, si existe una heterogeneidad en la población muy compleja, es decir, compuesta de varios niveles o etapas, entonces, el diseño óptimo de la muestra será el muestreo multietápico.

Mediante el registro de tiempo de ejecución de la labor de muestreo, se encontró que dos obreros tardan en promedio 3:43 minutos en cortar y pesar 10 tallos, 6:33 minutos en cada sitio y 13:07 min en cada surco. En total desde la llegada al sitio, preparación y muestreo tardan en promedio 4 horas, es decir que por día se pueden muestrear dos suertes, para un total de 44 por mes en una jornada de lunes a viernes de 8 horas y sábado de 4 horas.

El rango de suertes cosechadas mensualmente en el Ingenio Pichichi fluctúa entre 100 y 120. El costo total del personal operativo que realiza esta labor es de \$ 5.431.618 pesos colombianos (USD 1.810) teniendo en cuenta que devengan un salario mínimo legal vigente año 2016 (\$ 689.454 pesos colombianos (USD 229.8).

La implementación de esta estrategia de muestreo y uso del modelo permite detectar con cuatro meses de antelación paros de fábrica por falta de materia prima, que se estiman para el Ingenio Pichichi en \$ 11.877.000 millones de pesos por hora para el año 2016 (USD 3.959). Esto indica, que la inversión en el muestreo mensual es inferior a las pérdidas ocasionadas por detener el funcionamiento de la fábrica en una hora.

Adicionalmente, la herramienta permite pronosticar aumento en la producción, lo que ocasionaría un incremento en la edad de corte por la capacidad de molienda diaria, esto influye negativamente en el rendimiento del cultivo y por consiguiente en la recuperación de sacarosa.

Actualmente en el cultivo de caña de azúcar se está explorando la utilización de fotografías aéreas con la tecnología percepción remota para la estimación de producción, la cual disminuiría el tiempo de muestreo en campo. Según Bappel et al; 2015; Fortes, 2003; descrito por Sandoval & Carbonell, 2012 con la percepción remota se puede estimar producción de diferentes formas, con el uso de bandas espectrales, el cálculo de índices de vegetación o con la estimación de variables biofísicas, combinando cada una de las anteriores con modelos agro meteorológicos o de regresión.

Con el banco de datos obtenidos en las 62 suertes muestreadas se obtuvo el modelo ajustado para estimar el TCH a cosecha con un error estándar de 7.8 TCH:

$$\text{TCH(Cosecha)} = -73.34 + 0.889 \text{ TCH(muestreo)} + 31.93 \sqrt{2 \times \text{Edad de cosecha}} - 14.4I$$

El análisis de los resultados mostró que en promedio el TCH a cosecha se pueda reducir en 14.4 toneladas si en la suerte se presentó algún problema en la oportunidad de ejecución de alguna labor.

Los resultados del coeficiente de determinación (87%), indican que del 13% de la variabilidad del TCH entre los 8 meses y la cosecha no pudieron ser explicados por el premuestreo, la edad y la oportunidad de la labor.



Otros avances en estimación de producción para las condiciones del Valle del Cauca-Colombia, fueron descritos por Bello et al; 2013, donde el autor utilizó un modelo de desarrollo de follaje, enfocado principalmente a la simulación del desarrollo de la biomasa potencial del cultivo y la producción cosechable en respuesta al agua disponible (AquaCrop) concluyendo que es una buena herramienta, pero requiere de ajustes para su implementación.

La evaluación e implementación de diferentes metodologías de pronóstico utilizando herramientas estadísticas han permitido obtener resultados acertados, al respecto Ramirez et al; 2011, concluyo que el análisis de series de tiempo en el pronóstico de la producción de caña de azúcar con el modelo autoregresivo integrado de medias móviles (ARIMA) fue el apropiado en el pronóstico de la zafra 2006/2007 del ingenio Independencia de Martínez de la Torre, Veracruz, México, y este resulta de gran utilidad en la toma de decisiones para el mismo.

### **Conclusiones**

- Con la implementación del muestreo multi-etápico en 12 surcos, 2 sitios por surco y 5 tallos por sitio dentro de surco, es posible reducir el error de estimación del TCH al momento del muestreo por debajo de 5TCH.
- El modelo para estimar el TCH a la cosecha tuvo un error estándar de 7.8 TCH.
- El método de muestreo es independiente del área de la suerte.
- Independientemente del tamaño de la suerte dos muestreadores tardan en promedio 4 horas realizando el muestreo.
- El modelo de predicción, es una herramienta económicamente viable que permite tomar decisiones cuatro meses antes de cosecha con el fin de evitar pérdidas económicas por falta y/o exceso de caña, según la capacidad de molienda de la fábrica.
- Se requiere continuar con la colecta de datos de muestreo, para disminuir el error estándar de TCH del modelo.

## Referencias Bibliográficas.

Bello, C. A., Patiño, J., Almanza, E., Morroy, J., (2003). Uso del modelo Aquacrop para estimar rendimientos para el cultivo de caña de azúcar en el departamento del valle del cauca. ISBN I3427S/1/08.13 en: <http://www.fao.org/docrep/field/009/i3427s/i3427s.pdf>

Badii, M.H., Guillen, A., Landeros J., Cerna E. (2011). Análisis y Aplicación de Muestreo Multietápico, Estimación de Submuestreo y Muestreo de Respuesta Aleatoria. Daena: International Journal of Good Conscience. ISSN 1870-557X. 6(2) 88-95.

Carmona, Juan. (2004). evaluación de veinte métodos de estimación de la producción de caña en pie de unidades experimentales en ensayos de caña de azúcar (*Saccharum sp.*); Chicacao, Suchitepequez. Universidad de San Carlos de Guatemala.

Ramírez, J., Hernandez, G., Rodriguez, R. (2011). Análisis de series de tiempo en el pronóstico de la producción de caña de azúcar. Terra Latinoamericana, vol. 29, núm. 1, enero-marzo, 2011, pp. 103-109

Sanchez, L. Fernando., Gómez, Alvaro. (2014). Impacto de la oportunidad en la realización de las labores de campo en la productividad y el costo en el Ingenio Pichichi S.A. 9 congreso de técnicos azucareros de Latinoamérica y el caribe. San José de Costa Rica.

Sandoval, J. P., Carbonell, J. (2012). Principios y aplicaciones de la percepción remota en el cultivo de la caña de azúcar en Colombia. Centro de Investigación de la caña de azúcar-Cenicaña. 184 p.