

VALIDACIÓN EN EL CAMPO DE UN MÉTODO DE MUESTREO PARA EVALUAR LA CALIDAD EN LA RECOLECCIÓN DE CAFÉ

Rubén Darío Medina Rivera*; Esther Cecilia Montoya Restrepo*; Hernando García Osorio*

MEDINA R., R.D.; MONTOYA R., E.C.; GARCÍA O., H. Validación en el campo de un método de muestreo para evaluar la calidad en la recolección de café. Revista Cenicafé 66 (2): 110-120. 2015

Dado que el caficultor no cuenta con un método de muestreo ágil y confiable con respaldo estadístico, que le permita evaluar labores de recolección para establecer políticas de incentivo y pagos o ejercer acciones correctivas, se validó un método de muestreo denominado “plan surco” en la Estación Central Naranjal de Cenicafé, donde se registraron el número de frutos maduros dejados en el árbol y el número de frutos dejados en el suelo después de realizada la labor de recolección, en 17 jornadas de cosechas principales. Dicho método se validó a partir de la estimación de los parámetros promedio de frutos maduros dejados en el árbol y promedio de frutos dejados en el suelo, frente a la estimación por intervalos obtenida para estos mismos parámetros con el muestreo aleatorio simple. Con los resultados se estableció que de 17 jornadas evaluadas, en 13 de ellas, la diferencia absoluta entre las estimaciones con los dos métodos fue menor de un fruto, para el promedio de frutos maduros dejados en el árbol, mientras para el promedio de frutos dejados en el suelo, en 14 de las 17 jornadas, la diferencia absoluta fue menor o igual a 1, lo que indica que el método plan surco, es una herramienta administrativa con la cual puede evaluarse la calidad de recolección, tomando una muestra de 60 árboles en un lote.

Palabras clave: Plan surco, muestreo no probabilístico, estimación.

FIELD VALIDATION OF A SAMPLING METHOD TO EVALUATE HARVESTING QUALITY

Since coffee growers do not have a fast and reliable method of sampling with statistical support that enable them to assess harvest work to establish incentive and payments policies or to exercise corrective actions, a sampling method called "Plan surco" was validated at the Naranjal Central Station of Cenicafé. The number of ripe fruits left on the tree and on the ground after harvesting during 17 working days was recorded. This method was validated by estimating the average parameters of ripe fruits left on the tree and the average of ripe fruits left on the ground compared with the interval estimation obtained for these parameters with simple random sampling. The results established that in 13 out of the 17 evaluated workdays the absolute difference between the estimates with both methods was less than one fruit for the average ripe fruits left on the tree, while in 14 out of the 17 evaluated workdays the absolute difference for the average fruit left on the ground was less or equal to 1. This indicates that the method “Plan surco” is a management tool that can be used to assess the harvest quality taking a sample of 60 trees in a plot.

Keywords: Plan surco, non-probability sampling, estimation.

*Investigador Científico I, Investigador Científico III y Auxiliar III de Investigación, respectivamente. Disciplina de Biometría. Centro Nacional de Investigaciones de Café, Cenicafé. Manizales, Caldas, Colombia.

El cultivo del café es un proceso donde intervienen factores climáticos, biológicos y sociales entre otros, que hacen de ésta una actividad compleja y de gran dedicación. Para lograr que su negocio sea rentable, el productor debe conocer las condiciones y estado actual de las actividades que se realizan, como lo es el proceso de recolección, el cual representa entre 38,8% y 52,3% de los costos totales de producción (4).

Conocer qué tan bien se realiza la labor de recolección, en cuanto al número de frutos maduros que se dejan en el árbol (eficacia), la cantidad de frutos dejados en el suelo (pérdidas) y el número de frutos verdes en la masa cosechada (calidad), permite saber la cantidad de ingresos que el caficultor deja de percibir y advertir las posibles disminuciones futuras en calidad y producción como consecuencia de una mala recolección. Así mismo, a partir de esta información es posible obtener indicadores, los cuales se constituyen en una fuente de información que, integrada a los procesos de toma de decisiones, pueden favorecer el desempeño de las empresas cafeteras, al contribuir a la eficacia de los actores del mismo y a la mejor asignación y distribución de los recursos humanos, materiales y financieros, fijando además, políticas de pago a partir de estándares de calidad.

Teniendo en cuenta que los métodos utilizados por algunos caficultores para obtener esta información es subjetivo y no se conoce la confiabilidad de los mismos, se validó el método de muestreo por surcos con respaldo estadístico y con una probabilidad conocida de aciertos con respecto a lo obtenido con el muestreo aleatorio, para entregar una herramienta al caficultor que le permita evaluar la recolección, de manera rápida, confiable y menos costosa.

Partiendo de la situación anterior y dada la importancia y necesidad de contar con un método de muestreo ágil y confiable para evaluar los indicadores de eficacia y pérdidas en el proceso de la recolección del café, Gutiérrez (3) determinó estadísticamente el número de árboles necesarios a seleccionar para obtener las estimaciones del promedio de las variables frutos maduros dejados en el árbol y el suelo, utilizando dos trayectorias no probabilísticas, con respecto a las estimaciones obtenidas con un muestreo probabilístico. La primera trayectoria (plan surco), consistió en recorrer en línea recta y a través de los surcos el lote recolectado (Figura 1a), una vez ubicado en uno de los árboles centrales del primer surco, se seleccionó el número de árboles a evaluar, según el número de surcos en el lote. La segunda trayectoria (Figura 1b), consistió en recorrer en diagonal los surcos del lote recolectado, seleccionando un número establecido de árboles. Con los resultados obtenidos, Gutiérrez (3), concluyó que el número de árboles para estimar los promedios de las variables asociadas a la eficacia (número de frutos maduros dejados en el árbol después de la recolección) y pérdidas (número de frutos dejados en el suelo), en una jornada de recolección de café es de 90 árboles. En cuanto a la evaluación de los dos métodos de muestreo propuestos, en las Tablas 1 y 2, se presenta el promedio por jornada para ambas trayectorias, para las variables número de frutos dejados en el árbol y en el suelo.

Para la variable, número de frutos dejados en el árbol, con la trayectoria en línea recta se presentó un mayor número de casos (sitio y jornada) donde el promedio cayó en los intervalos de estimación para el promedio obtenido, con el diseño de muestreo aleatorio (Tabla 1).

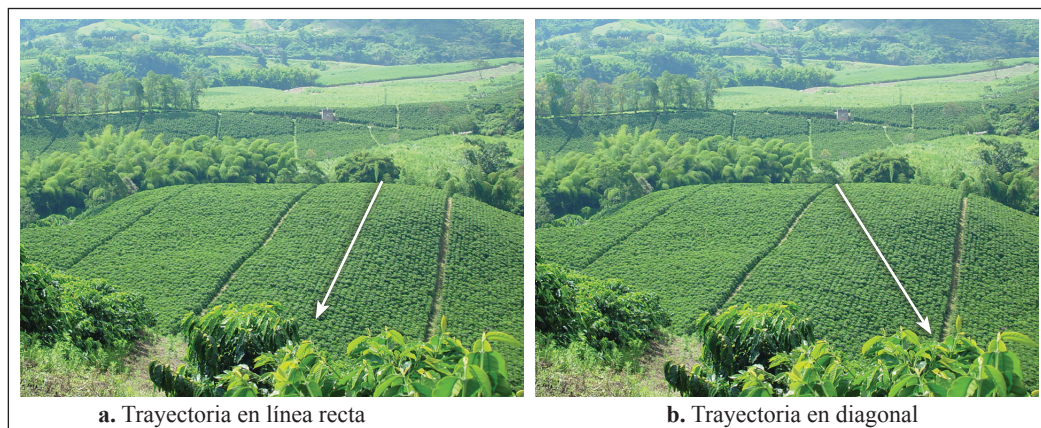


Figura 1. Trayectorias para evaluar indicadores de eficacia y pérdidas en el proceso de la recolección del café.

Para la variable número de frutos dejados en el suelo, las dos trayectorias tienen el mismo comportamiento con respecto a la estimación obtenida con el diseño de muestreo aleatorio (Tabla 2).

Por otra parte, se determinó que el tiempo medio invertido con las dos trayectorias para contar el número de frutos maduros dejados en el árbol y en el suelo en 81 árboles, estuvo entre 68 y 95 min, con una confiabilidad del 95%.

Se recomendó finalmente, la trayectoria en línea recta, porque además de tener un soporte estadístico con respecto al muestreo aleatorio, fue fácil de aplicar en el campo.

A partir de dicha recomendación, se hizo la validación de la trayectoria en línea recta, la cual será referida en adelante como muestreo en surcos o plan surco.

MÉTODO

En la Estación Experimental Naranjal de Cenicafé, en 17 jornadas de la cosecha principal, se evaluó la labor de recolección en lotes de café, con diferentes condiciones

de variedad, densidad y edad utilizando dos métodos de muestreo: el muestreo aleatorio simple (método 1) y el método del plan surco (método 2) siendo este último de tipo no probabilístico, de la siguiente manera:

Método 1. En el lote, del total de árboles recolectados en la jornada se seleccionaron aleatoriamente 60 árboles, y en ellos se contó el número de frutos maduros sin recolectar en el árbol (incluyendo sobremaduros y secos) y el número de frutos dejados en el suelo.

Método 2. Identificada el área del lote que fue recolectado durante la jornada, se seleccionó al azar el árbol central del primer surco, para iniciar la trayectoria en línea recta a través de los surcos, en los cuales se seleccionaron 90 árboles de acuerdo con los siguientes criterios:

- Si el número de surcos es menor que 90 (tamaño de muestra), entonces se toma en cada surco m árboles, de tal manera que al final del recorrido se tengan los 90 árboles, donde m es el número entero más cercano al cociente entre 90 y el número de surcos recolectados ($90/\text{número de surcos recolectados}$).

Tabla 1. Intervalo para la estimación del promedio del número de frutos dejados en el árbol, según el muestreo aleatorio simple y el promedio obtenido con las trayectorias, para cada sitio y cada jornada.

| Sitio | Jornada | Número promedio de frutos dejados en el árbol | | | |
|------------|---------|---|-----|--------------|----------|
| | | Aleatorio | | Trayectorias | |
| | | LI | LS | Recta | Diagonal |
| La Soledad | 1 | 0,6 | 1,8 | 0,8 | 0,6 |
| | 2 | 0,5 | 1,1 | 1,8 | 1,6 |
| | 3 | 0,8 | 1,5 | 1,3 | 2,4 |
| | 4 | 0,4 | 1,2 | 1,5 | 1,6 |
| Naranjal | 1 | 2,1 | 4,4 | 3,2 | 2,9 |
| | 2 | 1,4 | 2,8 | 3,4 | 3,5 |
| | 3 | 2,9 | 5,3 | 3,9 | 5,1 |
| | 4 | 2,8 | 4,8 | 5,0 | 6,0 |

LI: Límite inferior; LS: Límite superior.

Tabla 2. Intervalo para la estimación del promedio del número de frutos dejados en el suelo, según el muestreo aleatorio simple y el promedio obtenido con las trayectorias, para cada sitio y cada jornada.

| Sitio | Jornada | Número promedio de frutos dejados en el suelo | | | |
|------------|---------|---|------|--------------|----------|
| | | Aleatorio | | Trayectorias | |
| | | LI | LS | Recta | Diagonal |
| La Soledad | 1 | 1,1 | 2,6 | 1,0 | 1,9 |
| | 2 | 2,3 | 4,1 | 6,2 | 5,2 |
| | 3 | 2,6 | 7,8 | 3,4 | 3,7 |
| | 4 | 0,7 | 1,4 | 1,1 | 0,7 |
| Naranjal | 1 | 7,6 | 10,3 | 10,8 | 11,9 |
| | 2 | 6,6 | 9,6 | 10,8 | 7,1 |
| | 3 | 5,2 | 7,3 | 9,5 | 7,5 |
| | 4 | 1,9 | 3,4 | 4,1 | 4,6 |

LI: Límite inferior; LS: Límite superior.

- Si el número de surcos es mayor que 90, entonces se toma un árbol cada k surcos, donde k es el entero más cercano al cociente entre el número de surcos recolectados y 90 (número de surcos recolectados/90).
- Si el número de surcos es igual al tamaño de muestra (90), entonces se toma un árbol por surco.

Con los registros de cada método de muestreo y jornada, se realizó el siguiente análisis estadístico utilizando el SAS (5) versión 9.4.

Con la información del método 1, se obtuvieron los promedios y errores de estimación asociados al promedio de frutos maduros dejados en el árbol y el promedio de frutos dejados en el suelo, y se construyó la estimación por intervalos de los parámetros correspondientes, con un coeficiente de confianza del 95%.

Con el promedio de los errores de estimación asociados al promedio del número de frutos en el árbol y frutos en el suelo, se procedió a verificar si éste fue mayor a un fruto, según

la prueba de t al 5%, en caso contrario se validó el error de estimación para el tamaño de muestra fijado en 90 árboles.

Con los datos obtenidos en el método 2, se obtuvo la estimación del promedio de frutos maduros dejados en el árbol y el promedio de frutos dejados en el suelo.

Para cada parámetro se identificaron cuántas de las estimaciones obtenidas con el método 2 cayeron en los respectivos intervalos de confianza construidos con el método 1, para obtener así, la proporción de estimaciones que caen dentro del intervalo y se verificó si esta proporción es mayor al 90%, según la prueba t al 5%, con lo cual se habrá cumplido con la validación del método en surcos.

RESULTADOS Y DISCUSIÓN

En las Tablas 3 a la 6, se presenta para cada jornada, independientemente de las características de los lotes, los percentiles 25, 50, 75, mínimo y máximo observados, para las variables número de frutos maduros dejados en el árbol y número de frutos dejados en el suelo por árbol, después de la recolección, obtenidos con el muestreo aleatorio simple y el muestreo en surcos.

Para la variable número de frutos maduros dejados en el árbol, se presentó que en la jornada 2, el 50% de los árboles evaluados tenían aproximadamente hasta 8 frutos maduros dejados en el árbol, utilizando los dos métodos de muestreo (Tabla 3 y 4). El número máximo de frutos maduros dejados en el árbol por jornada estuvo entre 8 y 58 frutos para el muestreo aleatorio simple (Tabla 3), mientras que con el muestreo en surcos, el número máximo de frutos maduros dejados en el árbol osciló entre 10 y 55 frutos (Tabla 4).

Con el muestreo aleatorio simple, se observó para la jornada 8, que el 75% de los árboles registraban 18 ó menos frutos dejados en el suelo (Percentil 75) y para al menos uno de los árboles se registró un máximo de 55 frutos dejados en el suelo (Tabla 5).

Para el muestreo en surcos, se observó en la jornada 8, hasta 11 frutos dejados en el suelo para el 75% de los árboles y un máximo de 35 frutos dejados en el suelo (Tabla 6), lo que indica que el método de muestreo en surcos coincide con las descripciones obtenidas al utilizar el muestreo aleatorio.

Con los registros obtenidos utilizando el método de muestreo aleatorio simple, para cada una de las jornadas se construyó la estimación del promedio de número de frutos maduros dejados en el árbol y número de frutos dejados en el suelo por árbol, con un coeficiente de confianza del 95% (Tablas 7 y 8).

El promedio de frutos maduros dejados en el árbol obtenido con el método de muestreo en surco cayó en 15 de los 17 intervalos de estimación, para el promedio de frutos maduros dejados en el árbol obtenido con el diseño de muestreo aleatorio (Tabla 7), mientras que para la variable número de frutos dejados en el suelo, 12 de los 17 promedios obtenidos con el método de muestreo en surco cayeron en los intervalos de estimación del muestreo aleatorio simple (Tabla 8).

De las 17 jornadas evaluadas, en 13 de ellas, la diferencia absoluta entre las estimaciones con los dos métodos, es menor de un fruto, para los frutos maduros dejados en el árbol, a excepción de una de las jornadas, cuya diferencia fue de 4,1 frutos. Para los frutos dejados en el suelo, en 14

Tabla 3. Mínimo, percentiles y máximo para el número de frutos maduros dejados en el árbol por jornada, obtenidos con el muestreo aleatorio simple.

| Jornada | Mínimo | Percentil 25 | Percentil 50 | Percentil 75 | Máximo |
|----------------|---------------|---------------------|---------------------|---------------------|---------------|
| 1 | 0 | 0 | 2 | 4,5 | 14 |
| 2 | 0 | 3 | 7,5 | 13 | 54 |
| 3 | 0 | 0 | 2 | 5,5 | 38 |
| 4 | 0 | 2 | 4 | 10 | 40 |
| 5 | 0 | 0 | 1 | 3 | 17 |
| 6 | 0 | 0 | 1 | 2 | 57 |
| 7 | 0 | 1 | 3 | 10 | 30 |
| 8 | 0 | 1 | 3 | 6 | 58 |
| 9 | 0 | 0 | 1 | 3 | 13 |
| 10 | 0 | 0 | 1 | 3 | 8 |
| 11 | 0 | 0,5 | 2 | 4,5 | 18 |
| 12 | 0 | 0 | 1 | 2,5 | 9 |
| 13 | 0 | 0 | 1 | 3 | 19 |
| 14 | 0 | 0 | 1 | 4 | 24 |
| 15 | 0 | 0 | 0 | 1,5 | 45 |
| 16 | 0 | 1 | 2 | 4 | 9 |
| 17 | 0 | 0 | 1 | 2 | 22 |

Tabla 4. Mínimo, percentiles y máximo para el número de frutos maduros dejados en el árbol por jornada, obtenidos con el muestreo en surco.

| Jornada | Mínimo | Percentil 25 | Percentil 50 | Percentil 75 | Máximo |
|----------------|---------------|---------------------|---------------------|---------------------|---------------|
| 1 | 0 | 1 | 3 | 5 | 35 |
| 2 | 0 | 4 | 7 | 11 | 22 |
| 3 | 0 | 0 | 1 | 3 | 47 |
| 4 | 0 | 1 | 3 | 8 | 55 |
| 5 | 0 | 0 | 2 | 4 | 25 |
| 6 | 0 | 0 | 1 | 3 | 23 |
| 7 | 0 | 0 | 2 | 7 | 34 |
| 8 | 0 | 1 | 2,5 | 5 | 45 |
| 9 | 0 | 0 | 1 | 2 | 22 |
| 10 | 0 | 0 | 1 | 4 | 12 |
| 11 | 0 | 1 | 2 | 8 | 34 |
| 12 | 0 | 0 | 2 | 3 | 10 |
| 13 | 0 | 0 | 1 | 3 | 34 |
| 14 | 0 | 0 | 1 | 3 | 17 |
| 15 | 0 | 0 | 0 | 3 | 21 |
| 16 | 0 | 1 | 2 | 4 | 13 |
| 17 | 0 | 0 | 1 | 2 | 13 |

Tabla 5. Mínimo, percentiles y máximo para el número de frutos dejados en el suelo por jornada, obtenidos con el muestreo aleatorio simple.

| Jornada | Mínimo | Percentil 25 | Percentil 50 | Percentil 75 | Máximo |
|----------------|---------------|---------------------|---------------------|---------------------|---------------|
| 1 | 0 | 0 | 0 | 1 | 7 |
| 2 | 0 | 1 | 5,5 | 12 | 29 |
| 3 | 0 | 0 | 0 | 3 | 11 |
| 4 | 0 | 2 | 3 | 6,5 | 15 |
| 5 | 0 | 0 | 2 | 4 | 11 |
| 6 | 0 | 0 | 1 | 4 | 9 |
| 7 | 0 | 2 | 3 | 7 | 15 |
| 8 | 0 | 4 | 10 | 18 | 55 |
| 9 | 0 | 0 | 0 | 1,5 | 11 |
| 10 | 0 | 0 | 2 | 3 | 11 |
| 11 | 0 | 0 | 2 | 3 | 8 |
| 12 | 0 | 0,5 | 2,5 | 4 | 15 |
| 13 | 0 | 0 | 1 | 4,5 | 16 |
| 14 | 0 | 2 | 3 | 5 | 14 |
| 15 | 0 | 0 | 1 | 5,5 | 15 |
| 16 | 0 | 1 | 2 | 4 | 12 |
| 17 | 0 | 1 | 2 | 4,5 | 31 |

Tabla 6. Mínimo, percentiles y máximo para el número de frutos dejados en el suelo por jornada, obtenidos con el muestreo en surco.

| Jornada | Mínimo | Percentil 25 | Percentil 50 | Percentil 75 | Máximo |
|----------------|---------------|---------------------|---------------------|---------------------|---------------|
| 1 | 0 | 0 | 0 | 1 | 4 |
| 2 | 0 | 0 | 1 | 2 | 13 |
| 3 | 0 | 0 | 1 | 3 | 17 |
| 4 | 0 | 0 | 1 | 2 | 9 |
| 5 | 0 | 1 | 3 | 5 | 26 |
| 6 | 0 | 1 | 2 | 3 | 12 |
| 7 | 0 | 2 | 5 | 8 | 24 |
| 8 | 0 | 3 | 7 | 11 | 35 |
| 9 | 0 | 0 | 0 | 1 | 17 |
| 10 | 0 | 0 | 2 | 3 | 10 |
| 11 | 0 | 1 | 2 | 4 | 11 |
| 12 | 0 | 0 | 1 | 3 | 14 |
| 13 | 0 | 0 | 0 | 3 | 10 |
| 14 | 0 | 0 | 2 | 5 | 19 |
| 15 | 0 | 0 | 2 | 5 | 16 |
| 16 | 0 | 0 | 2 | 4 | 16 |
| 17 | 0 | 1 | 2 | 5 | 20 |

Tabla 7. Intervalo para la estimación del promedio del número de frutos maduros dejados en el árbol, según el muestreo aleatorio simple y el promedio obtenido con el plan surco para cada jornada.

| Jornada | Muestreo aleatorio | | Plan surco |
|---------|--------------------|-----------------|------------|
| | Límite inferior | Límite superior | Promedio |
| 1 | 2,2 | 3,9 | 3,5 |
| 2 | 6,6 | 16,6 | 7,5 |
| 3 | 2,4 | 6,3 | 3,5 |
| 4 | 5,1 | 9,3 | 6,1 |
| 5 | 1,5 | 3,3 | 2,9 |
| 6 | 0,4 | 4,2 | 2,1 |
| 7 | 4,1 | 7,7 | 5,0 |
| 8 | 4,1 | 8,1 | 4,1 |
| 9 | 1,5 | 3,0 | 2,1 |
| 10 | 1,4 | 2,6 | 2,7 |
| 11 | 2,3 | 4,3 | 5,2 |
| 12 | 1,2 | 2,3 | 2,2 |
| 13 | 1,2 | 3,1 | 3,0 |
| 14 | 1,8 | 3,8 | 2,3 |
| 15 | 0,6 | 4,0 | 2,2 |
| 16 | 2,1 | 3,3 | 2,7 |
| 17 | 0,9 | 2,6 | 1,5 |

Tabla 8. Intervalo para la estimación del promedio del número de frutos dejados en el suelo, según el muestreo aleatorio simple y el promedio obtenido con el plan surco para cada jornada.

| Jornada | Muestreo aleatorio | | Plan surco |
|---------|--------------------|-----------------|------------|
| | Límite inferior | Límite superior | Promedio |
| 1 | 0,3 | 1,2 | 0,6 |
| 2 | 4,6 | 10,7 | 1,4 |
| 3 | 1,2 | 2,6 | 2,1 |
| 4 | 3,5 | 5,4 | 1,3 |
| 5 | 1,9 | 3,1 | 3,5 |
| 6 | 1,4 | 2,7 | 2,4 |
| 7 | 3,4 | 5,5 | 5,3 |
| 8 | 10,5 | 15,5 | 7,6 |
| 9 | 0,7 | 2,0 | 1,5 |
| 10 | 1,5 | 2,8 | 2,3 |
| 11 | 1,5 | 2,6 | 2,5 |
| 12 | 2,2 | 3,7 | 2,0 |
| 13 | 1,6 | 3,3 | 1,8 |
| 14 | 2,8 | 4,4 | 3,3 |
| 15 | 2,1 | 3,9 | 3,4 |
| 16 | 1,9 | 3,2 | 2,9 |
| 17 | 2,0 | 4,3 | 3,9 |

de las 17 jornadas, la diferencia absoluta fue menor o igual a 1, y para las demás jornadas osciló entre 3 y 6 frutos.

En cuanto al promedio de los errores de estimación asociados al promedio del número de frutos maduros dejados en el árbol y frutos dejados en el suelo obtenidos con el muestreo aleatorio simple, fue estadísticamente igual a un fruto, según prueba de t al 5% (Tabla 9), con lo cual se estaría verificando lo encontrado por Gutiérrez (3) en cuanto al error de estimación para el tamaño de muestra fijado en 90 árboles.

Por otra parte, el tiempo medio invertido con el método en surcos para contar en 90 árboles el número de frutos maduros dejados en el árbol y en el suelo, fluctuó entre 39 y

91 min./jornada, con dos operarios, teniendo en cuenta que fueron los mismos operarios en igualdad de condiciones, quienes realizaron todas las evaluaciones.

Pero, hasta dónde es posible disminuir el número de árboles a seleccionar para realizar el plan de muestreo en surco, de tal manera que el error de estimación no supere al obtenido con el muestreo aleatorio?

Para resolver este interrogante, se aplicó un método de simulación basado en remuestreo. Dicho método denominado *Bootstrap*, propuesto por Efron (1) y Efron y Tibshirani (2), tiene como idea central aproximar la distribución de muestreo de un estadístico y de sus propiedades, basado en una muestra empírica a partir de simulación, esto es, a partir de

Tabla 9. Promedio y error estándar de los errores de estimación obtenidos con el muestreo aleatorio simple, para cada variable asociada a los indicadores de calidad de recolección.

| Variable | Promedio | Error estándar |
|------------------------------------|----------|----------------|
| Frutos maduros dejados en el árbol | 1,44* | 0,259 |
| Frutos dejados en el suelo | 0,99* | 0,171 |

* Promedio errores de estimación igual a 1, según prueba t al 5%.

una muestra aleatoria con n observaciones, se extraen B muestras independientes con reposición. Para cada una de las B muestras se obtiene el valor del estadístico (estimador), obteniendo una aproximación a su distribución muestral, a partir de la cual es posible valorar el error estadístico, bien sea en cuanto a sesgo, el error estándar o tasa de error en la estimación con el estadístico definido (2).

La estimación *Bootstrap* del error estándar del estadístico, se obtiene mediante la Ecuación <1>.

$$\hat{\sigma}_{Boot} = \left(\sum_{i=1}^B (\hat{\theta}^{*i} - \hat{\theta}^{**})^2 / (B - 1) \right)^{1/2} <1>$$

$$\hat{\theta}^{**} = \frac{\sum_{i=1}^B \hat{\theta}^{*i}}{B}$$

Donde $\hat{\theta}^{**}$ corresponde a la estimación *Bootstrap* de la media de la distribución muestral *Bootstrap*.

Para obtener estimaciones del sesgo del estadístico de interés, se toma la diferencia entre la estimación del estadístico objetivo, obtenido con los datos de la muestra original, y la obtenida con el *Bootstrap* (Ecuación <2>).

$$Sesgo_{Boot}(\hat{\theta}) = \hat{\theta}^{**} - \hat{\theta} <2>$$

Para cada jornada y con la muestra de árboles seleccionados con el muestreo aleatorio simple, en la cual se hizo la evaluación de las variables número de frutos maduros dejados en el árbol y número de frutos dejados en el suelo por árbol (variables de interés) después de la recolección, se analizó la distribución de los errores de estimación con diferentes tamaños de muestra de 40, 45, 50, 55, 60, 70, 80 y 90. Al no asumir una determinada distribución teórica *a priori*, se utilizó la muestra original y se generaron 2.000 submuestras para cada una de las 17 jornadas, que sirvieron de base para estimar de forma inductiva la distribución muestral de las variables de interés, con sus respectivos intervalos de confianza, utilizando el método de los percentiles.

En la Figura 2, se ilustra la distribución muestral del estadístico número promedio de frutos dejados en el árbol, con muestras de 40, 50, 55 y 60 árboles, a partir de los datos observados en la jornada 1, con el muestreo aleatorio simple. Es posible verificar, que la distribución del estadístico es normal, pero que al aumentar el tamaño de la muestra, disminuye la desviación de los estadísticos (error estándar).

En la Figura 3, se presenta para cada una de las variables de respuesta número de frutos en el árbol (NFA) y número de frutos en el suelo (NFS) el error promedio de estimación, de acuerdo al tamaño de las submuestras extraídas de las muestras aleatorias originales de cada jornada. Pasar de una selección de 90 árboles en el lote a una de 60 árboles, implica aumentar el error de estimación en 0,22 frutos en la variable número de frutos en el árbol, mientras en la variable frutos dejados en el suelo dicho error de estimación se incrementa en 0,18 frutos. Dado que la magnitud de los aumentos en el error de estimación no supera los

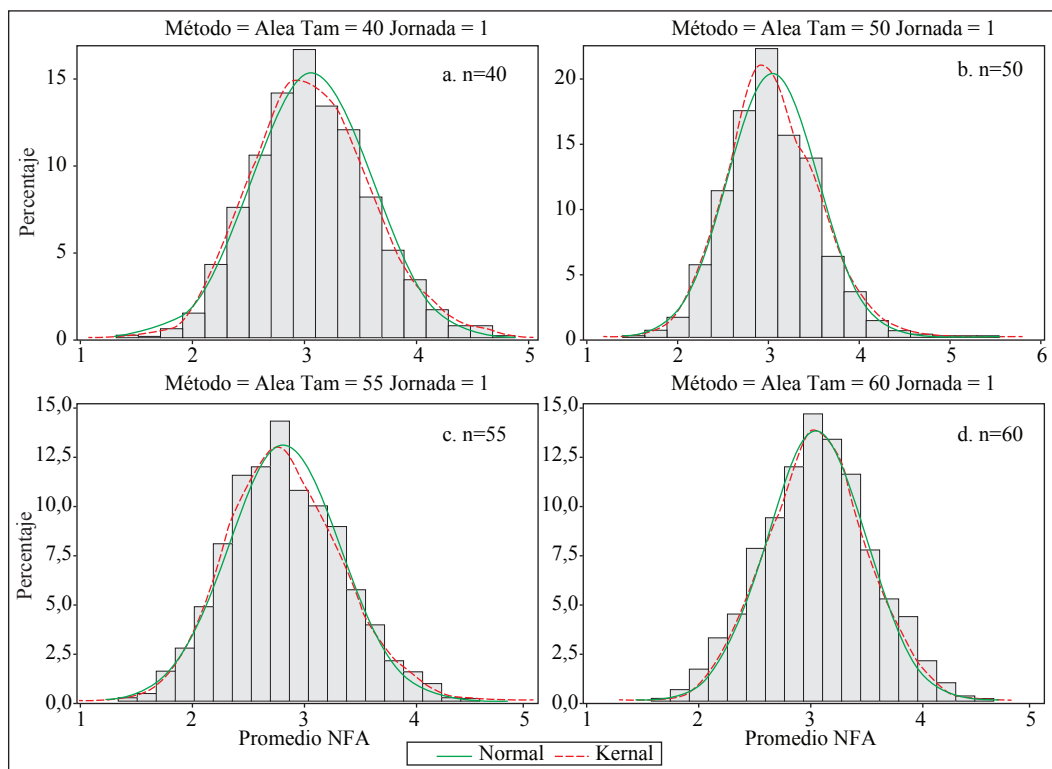


Figura 2. Distribución muestral del número promedio de frutos dejados en el árbol, para la jornada 1, con muestras de 40, 50, 55 y 60 árboles seleccionados.

0,3 frutos y que disminuir la muestra en 30 árboles se traduce en menor tiempo y facilidad de evaluación, se establece que 60 árboles por lote es un tamaño de muestra adecuado para la aplicación del muestreo en surcos. Además, con este número de plantas se mantiene la probabilidad entre el 70% y 83% de que las estimaciones con el muestreo en surco queden incluidas en el intervalo de las estimaciones obtenidas con el muestreo aleatorio. En cuanto al sesgo del promedio obtenido con el muestreo en surcos y el muestreo aleatorio, éste no supera los 0,22 frutos.

En general, con este trabajo se validó un método de muestreo denominado plan surco,

propuesto y evaluado para la estimación de indicadores de recolección de café. Esta validación permite sugerir al caficultor una herramienta administrativa para evaluar las labores de recolección, con el fin de establecer políticas de incentivo y pagos e incluso tomar correctivos.

Puede concluirse que los errores de estimación asociados al promedio del número de frutos maduros dejados en el árbol y frutos dejados en el suelo, obtenidos con el muestreo aleatorio simple, no superan un fruto, con un tamaño de muestra de 60 árboles. Teniendo en cuenta estos resultados, el tamaño de muestra para utilizar el muestreo en surcos puede ser de 60 árboles, dado que el sesgo

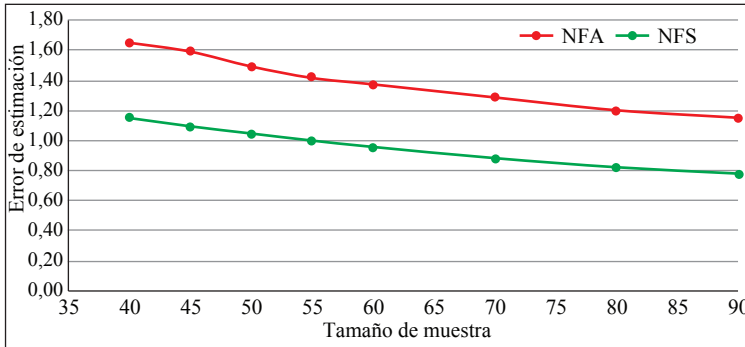


Figura 3. Error de estimación de acuerdo al número de árboles seleccionados en las submuestras a partir de la muestra aleatoria original.

no supera los 0,22 frutos y dicho método puede ser utilizado independientemente de las características del lote.

Para la variable número de frutos maduros dejados en el árbol, el método de muestreo en surcos tiene el mismo comportamiento con respecto a la estimación obtenida con el diseño de muestreo aleatorio, con una probabilidad del 82% de que las estimaciones queden incluidas en los intervalos de confianza construidos con el muestreo aleatorio.

LITERATURA CITADA

1. EFRON, B. Bootstrap methods: Another look at the Jackknife. *The annals of statistics* 7:1-26. 1979.
2. EFRON, B.; TIBSHIRANI, R. [Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy]: *Rejoinder. Statistical science* 1(1):77. 1986.
3. GUTIÉRREZ, M., A. Propuesta de un método de muestreo para la evaluación de indicadores asociados a la recolección de café. Cali : Universidad del Valle. Facultad ingeniería. Escuela de ingeniería industrial y estadística, 2006, 81 p. Tesis: Estadístico.
4. MADR. CCI. Servicio de información agropecuaria. Estructura de costos II semestre 2010. [En línea]. Bogotá : Agronet, (s.f.). Disponible en internet: [Http://www.agronet.gov.co/www/htm3b/public/boletines/Costos2010trim2/costos2010T2_archivos/frame.htm](http://www.agronet.gov.co/www/htm3b/public/boletines/Costos2010trim2/costos2010T2_archivos/frame.htm). Consultado el 25 de Octubre de 2014.
5. SAS 9.4 TS Level 1M1 X64_7PRO platform. *North Carolina* : SAS Institute : Cary, 2012.