

ESTIMACIÓN DE VALORES FALTANTES EN SERIES HISTÓRICAS DE BRILLO SOLAR

Rubén Darío Medina-Rivera*; Esther Cecilia Montoya-Restrepo*; Álvaro Jaramillo-Robledo**

RESUMEN

MEDINA R., R.D.; MONTOYA R., E.C.; JARAMILLO R., A. Estimación de valores faltantes en series históricas de brillo solar. Revista Cenicafé 63(1): 59-69 2012

Los registros de brillo solar en la red meteorológica de la Federación Nacional de Cafeteros de Colombia (FNC), son una de las fuentes para la realización de estudios climáticos, planificación en sistemas agrícolas y actividades de seguimiento para fundamentar la toma de decisiones en materia de recursos, costos de producción y establecimiento de política ambiental. En ocasiones no se tienen registros completos de brillo solar debido a mediciones inoportunas o errores en el cambio de la gráfica del heliógrafo. Con el fin de tener series históricas completas de brillo solar se evaluó el "método de la regresión acumulada" para estimar valores faltantes de estas series, en las estaciones de la red meteorológica de la FNC. Este método se basa en series de datos acumulados por ciclos, en períodos mensuales, teniendo en cuenta la agrupación por años dominados por los eventos climáticos de El Niño, La Niña y Neutro. Se obtuvo un error relativo promedio, estadísticamente inferior al 10%, cuando se estiman hasta diez datos faltantes para un mismo mes, independientemente del grupo, y un error relativo promedio, estadísticamente inferior al 6%, cuando se estiman hasta seis datos faltantes para un mismo mes de los grupos El Niño y Neutro. Para lograr estos errores de estimación, se requiere de series de brillo solar diario superiores a 25 años. El método de la regresión acumulada es flexible para estimar datos faltantes de horas de brillo solar, dado que se ajusta a las características propias de cada estación y su estimación depende única y exclusivamente de su historia.

Palabras clave: Red meteorológica, regresión acumulada.

ABSTRACT

Sunshine records obtained from the meteorological network of the National Federation of Coffee Growers of Colombia (FNC) are one of the sources used to conduct climate studies, agricultural planning, and follow-up actions to support decision-making regarding resources, production costs, and the establishment of an environmental policy. Sometimes you do not have complete sunshine records due to untimely measurement or to errors in the change of the heliograph graph. In order to have complete time series of sunshine, the "method of cumulative regression" was evaluated to estimate the missing values of these series in the meteorological network stations of the FNC. This method is based on a series of data gathered in monthly periods, taking into account the grouping into years dominated by the climatic events of El Niño, La Niña and Neutro. An average relative error, statistically less than 10% is obtained when up to ten missing data are estimated for a given month regardless of the group, and an average relative error, statistically less than 6% is obtained when up to six missing data are estimated for a given month of the El Niño and Neutro groups. To achieve these estimation errors, series of daily sunshine over 25 years are necessary. The cumulative regression method is versatile to estimate missing data in hours of sunlight since it fits the characteristics of each station and its estimation depends exclusively on its history.

Keywords: Meteorological network, cumulative regression.

* Investigador Científico I, Investigador Científico III, respectivamente, Disciplina de Biometría Cenicafé.

** Investigador Científico III, Disciplina de Agroclimatología, Centro Nacional de Investigaciones de Café, Cenicafé. Manizales, Caldas, Colombia.

El servicio meteorológico de la Federación Nacional de Cafeteros de Colombia (FNC), operado por Cenicafé, consciente de la necesidad de contar con información de calidad como factor clave para los estudios climáticos en las áreas donde se cultiva el café, y ante la dificultad de no tener siempre información completa de datos de precipitación, horas de brillo solar, temperatura, humedad del aire y evaporación, entre otros, establece la prioridad de disponer de un método con fundamento estadístico, para completar algunas de estas series históricas de clima. De las condiciones atmosféricas, interesa en este trabajo, especialmente la insolación, conocida como brillo solar, que se expresa como el número de horas en las cuales el sol brilla en un período de tiempo dado. Desde el punto de vista meteorológico el brillo solar es toda aquella radiación solar superior a 120 w.m^{-2} que tiene la capacidad de quemar el papel del heliógrafo (9).

La radiación solar es uno de los principales elementos que determina el microclima del cultivo, la productividad potencial del sistema, su energía condiciona la temperatura del aire y del suelo, el movimiento del viento, la evapotranspiración y la fotosíntesis, de tal manera que su cantidad, el grado de interceptación y la eficiencia en su uso, son factores fundamentales en el crecimiento de las plantas (6).

Por lo anterior, tener series históricas completas de brillo solar, se convierte en una tarea imprescindible y más aún, cuando a lo largo del tiempo la pérdida de datos es un aspecto de difícil control. Esto hace que sea necesario tener procedimientos confiables, sin embargo, la elección del procedimiento para el manejo de datos incompletos resulta una tarea compleja, pues un mismo método, en determinadas situaciones, produce estimaciones precisas y en otras, no, esto sugiere a los

investigadores que, cuando manejen datos incompletos, valoren previamente el uso de más de una alternativa para tratarlos y realicen un análisis de sensibilidad que les permita una mejor selección del procedimiento a implementar (1).

Flores *et al.* (4), estudiaron en Honduras el comportamiento de brillo solar (heliofanía) y aunque el objetivo del estudio no fue el de estimar datos faltantes de brillo solar, los autores completaron para algunas estaciones, las series históricas de brillo solar, correlacionando las series de las estaciones con datos faltantes, con aquellas estaciones cercanas que tuvieran series extensas de datos completos y condiciones similares en cuanto a vegetación, posición geográfica y altura. Además, incluyeron registros de brillo solar de estaciones de países vecinos como El Salvador y Guatemala. Como consecuencia del control de calidad series no se tuvieron en cuenta las estaciones con un coeficiente de correlación menor a 0,81. En el estudio no se presentan los errores asociados a la estimación de datos faltantes con esta metodología.

Guzmán y Gómez (5), estudiaron las características del brillo solar a partir de los datos registrados durante 36 años (1956–1991), en dos estaciones climatológicas situadas en el municipio de Chinchiná, Caldas. A estas series anuales se les aplicó una prueba de rachas para verificar la homogeneidad de las series, encontrando una discontinuidad en éstas, lo cual se explica por el cambio en la metodología de evaluación de las bandas de brillo solar en 1972 por la OMM, de allí que se hiciera una nueva evaluación diaria de todos los heliogramas originales, de 1972 hacia atrás, con el fin de contar con series correctas y homogéneas. Para cada estación, se hizo el análisis descriptivo y de normalidad de las series tanto anual como mensual.

Para Colombia se encontraron 29 regiones de acuerdo con la distribución del brillo solar (3), las cuales se obtuvieron mediante un análisis de agrupamiento (*Cluster analysis*) según la distribución promedio mensual de brillo solar de series de estaciones con al menos 10 diez años de datos diarios, la ubicación geográfica y el promedio total de brillo anual en cada estación. Para cada uno de los grupos determinados se obtuvo un patrón mensual con su correspondiente intervalo, con un coeficiente de confianza del 95%. Los resultados de este trabajo pueden ser tenidos en cuenta para estudiar la bondad de algunos métodos de interpolación para estimar datos faltantes de brillo solar.

Chaves (2), a partir de la información de brillo solar obtenida en 32 estaciones meteorológicas de la Federación Nacional de Cafeteros de Colombia, aplicó la metodología Box y Jenkins y ajustó modelos ARIMA a series mensuales con la estructura $ARIMA(0,1,1)_s(0,1,1)_{12}$ de promedios móviles con componente estacional de 12 meses. Los pronósticos obtenidos fueron cercanos de los valores observados, actualizados mensualmente, con errores de pronóstico menores del 23%, el cual se considera aceptable. En este trabajo, se mencionan las ventajas de la aplicación de la metodología Box y Jenkins para describir y modelar series de datos de variables climáticas, con el fin de hacer pronósticos a corto plazo.

Se hace evidente en la literatura, los esfuerzos de diferentes autores por resolver el problema de los datos faltantes, pero más enfocados hacia la modelación del fenómeno, partiendo de series históricas completas o a partir de la simulación y generación de series utilizando algoritmos iterativos o técnicas Montecarlo, incluso existe un generador estocástico de datos meteorológicos (10), el cual es un modelo

numérico que genera series diarias sintéticas de variables de clima como precipitación, temperatura y radiación solar. Es éste, los estados para la precipitación (Seco/húmedo) se genera con Cadenas de Markov y la serie de precipitación con función Gamma; para la radiación solar se asume una distribución normal. Para estos trabajos, no se determinan errores de estimación, dado que se generan escenarios de datos diarios.

A partir de lo anterior, y dada la necesidad de tener series históricas completas de brillo solar, se quiso evaluar el “Método de la regresión acumulada” para estimar valores faltantes en las series históricas de brillo solar de estaciones pertenecientes a la red climatológica de la Federación Nacional de Cafeteros-FNC, sin pretender comparar los resultados obtenidos con la bondad de otros métodos que puedan existir para este propósito.

MATERIALES Y MÉTODOS

Se utilizaron registros de horas de brillo solar diario, de diez estaciones climáticas de la red meteorológica de la FNC, con más de 25 años de información, operada por la Disciplina de Agroclimatología de Cenicafé (Tabla 1).

Para el análisis se trabajó con las series a partir del año 1974 (34 años), debido al cambio en la metodología de medición para esta variable. Se cuantificó el número de registros faltantes por estación y mes, con el fin de obtener series completas de brillo solar. En cada estación y para cada mes, se establecieron los episodios de El Niño, de La Niña y condición Neutra (grupos), de acuerdo con los Índices Oceánicos de El Niño-ONI, establecidos internacionalmente y publicados en la página del National Weather Service, Climate Prediction Center-NOAA (8). Estos

valores ONI, indican las desviaciones de la temperatura del agua superficial del océano Pacífico ecuatorial, donde temperaturas iguales o mayores a los $0,5^{\circ}\text{C}$ por encima de un valor medio normal (media del año 1971 al 2000) durante 3 meses consecutivos, indica la ocurrencia de El Niño; por el contrario, si la temperatura es de $-0,5^{\circ}\text{C}$ o menos, es un período de La Niña, si no se reúnen estas condiciones se considera un período Neutro (normal).

Para cada estación se seleccionaron de manera aleatoria, 10 meses de cada condición (El Niño, La Niña y Neutro) para validar el método de regresión acumulada, en la estimación de datos faltantes de horas de brillo solar diario.

Con la información restante se hizo lo siguiente:

1. Se tomó como variable de interés o endógena el número de horas de brillo solar acumulado transcurridas desde el inicio hasta el final de cada mes.

2. Para cada estación y grupo se hizo un análisis descriptivo con la variable de interés, y se estimaron los intervalos para el promedio del brillo solar acumulado mensual (histórico), por estación y grupo, con un coeficiente de confianza del 95%.

3. Para cada estación, grupo y mes, a través de una regresión lineal simple, entre el brillo solar acumulado y el tiempo transcurrido en días, se obtuvo la tasa diaria de brillo solar acumulado (Z).

4. En cada estación y en cada grupo, con la información para la validación, se tomó aleatoriamente 1 día o tiempo t (caso 1), 2 días (caso 2), 3 días (caso 3), 4 días (caso 4) hasta completar 10 días (caso 10), en los cuales se asume (en cada caso) como dato faltante. Una vez identificados los días con datos faltantes, se aplicó el método de

la regresión acumulada para su estimación (7), el cual consiste en:

- Identificar el grupo al cual pertenece el mes con el día del dato faltante: El Niño, La Niña, o Neutro.
- Obtener el acumulado de horas de brillo solar (X)
- Si X no es mayor descriptivamente, que el límite superior para el intervalo del acumulado histórico, del mes y grupo, respectivamente, la estimación del dato faltante es Z , de lo contrario será cero.

La validación del método consistió en:

- Obtener el acumulado para el mes, incluyendo la estimación del dato faltante (valor estimado).
- Con el valor real (acumulado incluyendo el dato que fue retirado para asumir el dato faltante), se establece la diferencia absoluta y relativa, del valor estimado con respecto al valor real, los cuales son a su vez, los errores de estimación absoluto y relativo, respectivamente.
- Con el error de estimación relativo y el número de datos faltantes (casos), se evaluó la hipótesis de investigación: “El error relativo promedio de estimación es menor del 13%”, con el estadístico de prueba de t , al 5%.

RESULTADOS Y DISCUSIÓN

En la Tabla 2 se presenta el número de meses y total de días con faltantes de datos de brillo solar, para las estaciones de estudio. Al revisar las series de datos, se encontró que las estaciones Chapetón, Pueblo Bello y Naranjal, en 34 años, presentaron de 546 a 859 días (18–29 meses) sin registro de

Tabla 1. Localización de las estaciones de la red climática de la FNC y número de años con datos registrados de brillo solar.

| Estación | Municipio (Departamento) | Latitud Norte | | Longitud Oeste | | Altitud (m) | Número de años observados | Período de la serie | |
|-----------------|-----------------------------|------------------|----|-------------------|----|----------------|---------------------------------|------------------------|-------|
| | | ° | ' | ° | ' | | | Inicio | Final |
| Bertha | Moniquirá (Boyacá) | 05 | 53 | 73 | 34 | 1.677 | 34 | 1953 | 2007 |
| Chapetón | Ibagué (Tolima) | 04 | 28 | 75 | 16 | 1.353 | 34 | 1955 | 2007 |
| Julio Fernández | Restrepo (Valle) | 03 | 49 | 76 | 32 | 1.381 | 34 | 1954 | 2007 |
| Jorge Villamil | Gigante (Huila) | 02 | 20 | 75 | 31 | 1.420 | 34 | 1955 | 2007 |
| Manuel Mejía | El Tambo (Cauca) | 02 | 24 | 76 | 44 | 1.735 | 34 | 1953 | 2007 |
| Naranjal | Chinchiná (Caldas) | 04 | 58 | 75 | 39 | 1.381 | 34 | 1956 | 2007 |
| Ospina Pérez | Consacá (Nariño) | 01 | 15 | 77 | 29 | 1.603 | 34 | 1953 | 2007 |
| Paraguaicito | Buenavista (Quindío) | 04 | 24 | 75 | 44 | 1.203 | 34 | 1963 | 2007 |
| Pueblo Bello | Valledupar (Cesar) | 10 | 25 | 73 | 34 | 1.134 | 34 | 1958 | 2007 |
| Santa Helena | Marquetalia (Caldas) | 05 | 19 | 75 | 00 | 1.395 | 27 | 1981 | 2007 |

brillo solar, a diferencia de Paraguaicito y Santa Helena, con 1 y 2 meses sin registro, respectivamente.

Con base en los datos de la Tabla 1, se eliminaron de las series de datos, aquellos meses con al menos un dato faltante de brillo solar, obteniendo series completas. Así, teniendo en cuenta el período de estudio de las series de datos para cada estación (Tabla 1), se ilustra el número de meses o ciclos estudiados en cada grupo, sin contar con los 10 meses seleccionados aleatoriamente de cada condición de El Niño, La Niña y Neutro (Tabla 3), para la validación.

En la Tabla 4, se presentan los valores mínimos, máximos y los percentiles 25, 50 y 75, para la estación Pueblo Bello, en la cual se han registrado como máximo 11,3 horas de brillo solar en al menos un día del mes de febrero del grupo La Niña; del

Tabla 2. Número de meses y total de días con faltantes de datos de brillo solar, por estación.

| Estación | Meses | Días |
|-----------------|-------|------|
| Bertha | 7 | 215 |
| Chapetón | 29 | 859 |
| Julio Fernández | 3 | 91 |
| Jorge Villamil | 4 | 123 |
| Manuel Mejía | 7 | 216 |
| Naranjal | 18 | 546 |
| Ospina Pérez | 7 | 199 |
| Paraguaicito | 1 | 28 |
| Pueblo Bello | 27 | 820 |
| Santa Helena | 2 | 61 |

total de días analizados, el 75% de éstos, no superan las 9,6 horas de sol al día; además, independientemente del grupo, en el 25% de los días del mes de enero, se registran hasta 7,0 horas de brillo solar diarias, es decir, que en el 75% de los días de este

Tabla 3. Número de meses por estación y grupo, utilizados para la estimación de datos faltantes de brillo solar.

| Estación | Grupo | | |
|-----------------|---------|---------|--------|
| | El Niño | La Niña | Neutro |
| Bertha | 67 | 49 | 111 |
| Chapetón | 62 | 55 | 131 |
| Julio Fernández | 101 | 74 | 187 |
| Jorge Villamil | 101 | 75 | 187 |
| Manuel Mejía | 96 | 75 | 188 |
| Naranjal | 99 | 57 | 180 |
| Ospina Pérez | 101 | 75 | 183 |
| Paraguaicito | 101 | 75 | 189 |
| Pueblo Bello | 84 | 72 | 171 |
| Santa Helena | 87 | 57 | 146 |

mes, como mínimo se presentan 7,0 horas de brillo solar diario, siendo éste, el único mes, que en todos los grupos, registra en al menos uno de los días como mínimo 0,1 horas de brillo solar.

En general, independiente del grupo analizado, en el 50% de los días de cada mes, para la estación Pueblo Bello, el número de horas de brillo solar que se ha registrado

en un día, alcanza las 8,7 horas, mientras para las demás estaciones analizadas, no se superan las 7,0 horas de brillo solar en un día.

Al analizar el brillo solar acumulado mensual por grupo, en siete de las diez estaciones, el evento La Niña se diferencia por ser el de menor número de horas de brillo solar (Figura 1), de tal manera que las estaciones Chapetón, Julio Fernández, Manuel Mejía, Naranjal, Ospina Pérez y Paraguaicito, presentaron en promedio un brillo solar mensual entre 90 y 129 horas, con errores de estimación entre 4,0 y 6,6 horas para este mismo grupo. La estación Jorge Villamil, recibe en promedio 100 horas de brillo solar, independientemente del grupo (Tabla 5).

Las estaciones Bertha, Pueblo Bello y Santa Helena, presentan en promedio la misma cantidad de horas de brillo solar acumuladas en el mes para los tres grupos (Figura 1). De las estaciones analizadas, Pueblo Bello, es la de mayor número de horas de brillo solar (más de 191 horas), independientemente del grupo.

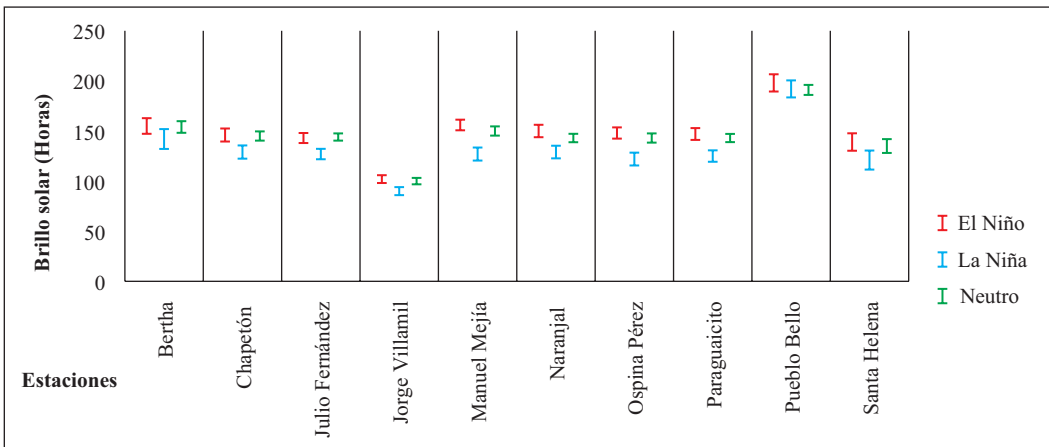


Figura 1. Intervalo, con un coeficiente de confianza del 95%, para el promedio de brillo solar acumulado mensual por evento y por estación.

Tabla 4. Percentiles de horas de brillo solar diario por mes y grupo. Estación Pueblo Bello. 1974- 2007.

| Mes | Grupo | Mínimo observado | Percentil | | | Máximo observado |
|------------|---------|------------------|-----------|-----|-----|------------------|
| | | | 25 | 50 | 75 | |
| Enero | El Niño | 0,5 | 7,5 | 8,6 | 9,5 | 10,6 |
| Febrero | | 0,0 | 7,1 | 8,7 | 9,6 | 10,8 |
| Marzo | | 0,2 | 5,5 | 7,2 | 8,5 | 10,1 |
| Abril | | 0,0 | 3,9 | 5,6 | 6,9 | 9,3 |
| Mayo | | 0,0 | 3,2 | 5,1 | 6,6 | 10,3 |
| Junio | | 0,0 | 4,3 | 6,2 | 7,7 | 10,6 |
| Julio | | 0,0 | 5,3 | 7,1 | 8,3 | 10,8 |
| Agosto | | 0,0 | 5,3 | 6,6 | 7,8 | 10,4 |
| Septiembre | | 0,0 | 3,9 | 6,0 | 7,0 | 9,9 |
| Octubre | | 0,0 | 4,1 | 5,6 | 7,2 | 10,7 |
| Noviembre | | 0,0 | 5,6 | 7,2 | 8,7 | 10,5 |
| Diciembre | | 0,5 | 7,4 | 8,6 | 9,4 | 10,5 |
| Enero | La Niña | 1,0 | 7,5 | 8,6 | 9,4 | 10,1 |
| Febrero | | 0,0 | 6,2 | 8,3 | 9,4 | 11,3 |
| Marzo | | 0,0 | 5,2 | 6,6 | 8,4 | 10,2 |
| Abril | | 0,1 | 4,8 | 6,5 | 8,1 | 10,7 |
| Mayo | | 0,0 | 4,0 | 6,0 | 7,4 | 10,4 |
| Junio | | 0,0 | 3,7 | 5,9 | 7,5 | 10,5 |
| Julio | | 0,7 | 4,5 | 6,5 | 8,3 | 10,9 |
| Agosto | | 0,0 | 3,7 | 5,8 | 7,0 | 10,5 |
| Septiembre | | 0,0 | 4,2 | 5,8 | 7,2 | 10,3 |
| Octubre | | 0,0 | 3,3 | 5,6 | 6,8 | 9,2 |
| Noviembre | | 0,0 | 4,1 | 5,7 | 7,4 | 10,4 |
| Diciembre | | 0,2 | 5,5 | 7,6 | 8,6 | 10,1 |
| Enero | Neutro | 0,1 | 7,2 | 8,5 | 9,2 | 10,5 |
| Febrero | | 1,6 | 6,6 | 8,2 | 9,4 | 10,7 |
| Marzo | | 0,2 | 5,6 | 7,3 | 8,9 | 10,7 |
| Abril | | 0,0 | 3,7 | 5,7 | 7,2 | 10,3 |
| Mayo | | 0,0 | 3,3 | 5,4 | 7,1 | 10,6 |
| Junio | | 0,0 | 4,0 | 6,0 | 7,6 | 10,7 |
| Julio | | 0,0 | 4,8 | 6,9 | 8,2 | 10,9 |
| Agosto | | 0,0 | 4,8 | 6,3 | 7,7 | 10,6 |
| Septiembre | | 0,0 | 4,1 | 6,0 | 7,3 | 10,5 |
| Octubre | | 0,0 | 4,5 | 5,8 | 7,2 | 10,0 |
| Noviembre | | 0,0 | 4,7 | 6,2 | 7,7 | 10,1 |
| Diciembre | | 0,6 | 6,5 | 8,1 | 9,0 | 10,3 |

Tabla 5. Promedio y error de estimación para el brillo solar acumulado mensual (horas), en cada estación y grupo.

| Estación | El Niño | | La Niña | | Neutro | |
|-----------------|----------|---------------------|----------|---------------------|----------|---------------------|
| | Promedio | Error de estimación | Promedio | Error de estimación | Promedio | Error de estimación |
| Bertha | 155 | 7,8 | 142 | 9,9 | 154 | 5,8 |
| Chapetón | 146 | 6,5 | 129 | 6,6 | 145 | 4,6 |
| Julio Fernández | 143 | 5,0 | 127 | 5,2 | 144 | 3,6 |
| Jorge Villamil | 102 | 3,9 | 90 | 4,0 | 100 | 3,2 |
| Manuel Mejía | 156 | 5,3 | 127 | 6,5 | 150 | 4,6 |
| Naranjal | 150 | 6,2 | 129 | 6,3 | 143 | 4,2 |
| Ospina Pérez | 148 | 5,6 | 122 | 6,4 | 143 | 4,6 |
| Paraguaicito | 147 | 5,9 | 125 | 5,7 | 143 | 4,1 |
| Pueblo Bello | 198 | 8,6 | 192 | 8,5 | 191 | 5,0 |
| Santa Helena | 139 | 8,7 | 121 | 9,5 | 135 | 6,9 |

Al analizar por grupo y mes, el brillo solar acumulado mensual de cada una de las estaciones, para la estación Chapetón, en el evento El Niño, el número promedio de horas de brillo solar acumulado, fluctuó entre 121 y 170 horas (Figura 2). Para el evento La Niña, descriptivamente, los promedios de brillo solar acumulado están entre 101 y 163 horas (Figura 2b), mientras en el evento Neutro, los promedios de brillo solar acumulado fluctúan entre 117 y 174 horas, presentando los menores errores de estimación con respecto al grupo El Niño y La Niña que van de 3,5 a 12,1 horas (Figura 2c).

En la Tabla 6, se ilustra la estimación de la tasa diaria de brillo solar (Z), para la estación Bertha, para cada mes, en cada evento, cuya interpretación es la siguiente:

Por ejemplo, para la estación Bertha, si se llegara a tener como dato faltante el día 8 del mes de enero clasificado en el evento El Niño, la estimación del número de horas de brillo solar para este día sería de 7,0 mm (Tabla 6), con un error de estimación de 0,10 horas, asumiendo que el acumulado de este mes no supera el límite

superior del promedio histórico. Así, para el mismo grupo en el mes de mayo, siempre y cuando se cumplan las condiciones, la estimación sería 4,0 horas con un error de estimación de 0,09.

Al analizar la tasa diaria para el acumulado de brillo solar mensual (Z) de otras estaciones, tales como Jorge Villamil, se observó que oscila entre 2,4 y 4,4 horas, siendo ésta la estación de menor número de horas de brillo solar; y como caso opuesto, se presenta la estación Pueblo Bello, donde la tasa diaria está entre 5,0 y 8,3 horas de brillo solar. Para las demás estaciones, la tasa diaria es como mínimo 3,4 sin que se superen las 7,0 horas.

Con respecto a la validación del método de la regresión acumulada, en la Tabla 7, se ilustran para cada estación y evento, los errores de estimación absoluto y relativo, independientemente del número de datos faltantes en la serie.

En general, para los grupos de El Niño y Neutro, los errores relativos no superan estadísticamente el 8%, mientras que en el evento La Niña, específicamente en la

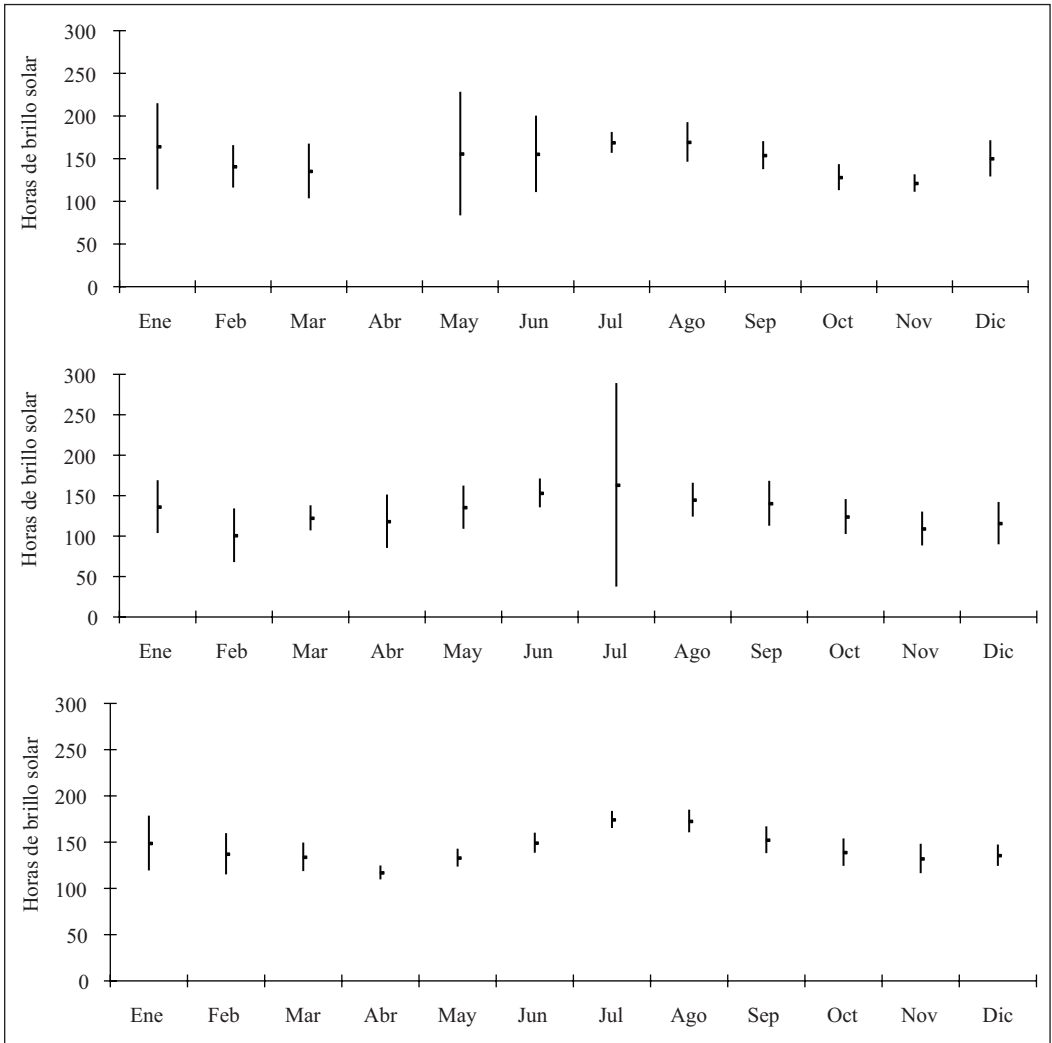


Figura 2. Brillo solar acumulado promedio por mes y su intervalo, con un coeficiente de confianza del 95%. Estación Chapetón. **a.** Evento El Niño; **b.** Evento La Niña; **c.** Evento Neutro.

estación Santa Helena, el error relativo puede ser hasta del 12%.

Al realizar la distribución de los errores relativos para cada caso (número de datos faltantes), por evento, se observó que para el evento El Niño y Neutro, estos errores son inferiores al 8%, según la prueba Z, al 5%, cuando se estiman hasta diez datos faltantes;

mientras que en el evento la Niña, los errores relativos no superan el 10% (Tabla 8).

En la Figura 3, se ilustra el comportamiento de los errores relativos promedio para cada grupo y en general, en función del número de datos faltantes, donde se observa, descriptivamente que a mayor número de datos faltantes mayor error de estimación.

Tabla 6. Tasa diaria para el acumulado de horas de brillo solar mensual, por grupo y mes. Estación Bertha.

| Grupo | Mes | Tasa diaria | |
|---------|------------|----------------|---------------------|
| | | Estimación (Z) | Error de estimación |
| El Niño | Enero | 7,0 | 0,10 |
| | Febrero | 6,0 | 0,12 |
| | Marzo | 5,5 | 0,14 |
| | Abril | 3,9 | 0,05 |
| | Mayo | 4,0 | 0,09 |
| | Junio | 4,5 | 0,09 |
| | Julio | 4,9 | 0,13 |
| | Agosto | 5,2 | 0,19 |
| | Septiembre | 4,5 | 0,11 |
| | Octubre | 4,5 | 0,13 |
| | Noviembre | 5,0 | 0,11 |
| | Diciembre | 5,7 | 0,11 |
| La Niña | Enero | 6,2 | 0,13 |
| | Febrero | 5,0 | 0,20 |
| | Marzo | 4,3 | 0,25 |
| | Abril | 4,2 | 0,28 |
| | Mayo | 4,2 | 0,11 |
| | Junio | 3,6 | 0,11 |
| | Julio | 4,3 | 0,22 |
| | Agosto | 4,3 | 0,16 |
| | Septiembre | 4,0 | 0,12 |
| | Octubre | 4,2 | 0,12 |
| | Noviembre | 4,9 | 0,19 |
| | Diciembre | 4,9 | 0,20 |
| Neutro | Enero | 6,4 | 0,13 |
| | Febrero | 6,9 | 0,18 |
| | Marzo | 4,9 | 0,12 |
| | Abril | 4,2 | 0,07 |
| | Mayo | 4,0 | 0,08 |
| | Junio | 4,5 | 0,09 |
| | Julio | 5,1 | 0,11 |
| | Agosto | 5,0 | 0,09 |
| | Septiembre | 4,7 | 0,09 |
| | Octubre | 4,9 | 0,09 |
| | Noviembre | 5,2 | 0,07 |
| | Diciembre | 5,7 | 0,11 |

Tabla 7. Promedio y error estándar de los errores de estimación (EE) absoluto y relativo, para cada estación y evento.

| Estación | Grupo | Error absoluto | | Error relativo | |
|-----------------|---------|----------------|-------|----------------|-------|
| | | Promedio | EE | Promedio | EE |
| Bertha | El Niño | 6,34 | 0,591 | 3,59* | 0,302 |
| | La Niña | 5,02 | 0,429 | 3,71* | 0,343 |
| | Neutro | 5,61 | 0,616 | 4,15* | 0,498 |
| Chapetón | El Niño | 8,32 | 0,908 | 6,46* | 0,806 |
| | La Niña | 7,08 | 0,778 | 6,33* | 0,640 |
| | Neutro | 7,87 | 0,656 | 6,54* | 0,590 |
| Julio Fernández | El Niño | 8,35 | 0,618 | 4,98* | 0,349 |
| | La Niña | 8,16 | 0,588 | 5,84* | 0,405 |
| | Neutro | 6,22 | 0,623 | 4,41* | 0,391 |
| Jorge Villamil | El Niño | 4,44 | 0,332 | 5,02* | 0,375 |
| | La Niña | 6,61 | 0,610 | 6,65* | 0,566 |
| | Neutro | 6,24 | 0,518 | 5,97* | 0,444 |
| Manuel Mejía | El Niño | 9,47 | 0,859 | 4,83* | 0,386 |
| | La Niña | 7,01 | 0,850 | 5,30* | 0,555 |
| | Neutro | 4,78 | 0,317 | 3,63* | 0,257 |
| Naranjal | El Niño | 4,93 | 0,379 | 3,10* | 0,241 |
| | La Niña | 4,36 | 0,389 | 3,33* | 0,290 |
| | Neutro | 6,01 | 0,525 | 4,12* | 0,363 |
| Ospina Pérez | El Niño | 7,88 | 0,838 | 5,19* | 0,480 |
| | La Niña | 9,48 | 1,044 | 6,55* | 0,593 |
| | Neutro | 5,50 | 0,442 | 3,69* | 0,310 |
| Paraguaicito | El Niño | 5,52 | 0,430 | 3,81* | 0,350 |
| | La Niña | 5,57 | 0,483 | 4,07* | 0,338 |
| | Neutro | 8,57 | 0,917 | 5,56* | 0,499 |
| Pueblo Bello | El Niño | 5,46 | 0,481 | 2,75* | 0,244 |
| | La Niña | 7,31 | 0,713 | 3,70* | 0,342 |
| | Neutro | 6,28 | 0,752 | 3,14* | 0,360 |
| Santa Helena | El Niño | 9,28 | 0,679 | 6,86* | 0,451 |
| | La Niña | 9,49 | 0,793 | 9,91* | 1,103 |
| | Neutro | 6,99 | 0,795 | 4,79* | 0,515 |

* Promedios menores estadísticamente al 13%, según prueba t al 5%.

Cuando las estimaciones se hacen, hasta seis datos faltantes de un mes del grupo El Niño o Neutro, los errores relativos promedio no superan el 6%; mientras que este mismo error se logra cuando se estiman hasta cuatro datos faltantes cuando el mes pertenece al grupo La Niña; sin embargo, es de anotar que estadísticamente, los errores relativos promedio

no superan el 10%, independientemente del grupo y el número de datos faltantes (Figura 2), con lo cual fue corroborada la hipótesis de investigación: “El error relativo promedio de estimación es menor del 13%”.

Puede concluirse que el método de la regresión acumulada, para estimar datos

Tabla 8. Promedio y error estándar para errores de estimación (EE) absolutos y relativos según el número de datos faltantes por evento.

| Grupo | Número de datos faltantes | Error absoluto | | Error relativo | |
|---------|---------------------------|----------------|--------|----------------|-------|
| | | Promedio | EE | Promedio | EE |
| El Niño | 1 | 2,86* | 0,229 | 1,89** | 0,139 |
| | 2 | 4,41* | 0,367 | 2,94** | 0,235 |
| | 3 | 5,67* | 0,512 | 3,66** | 0,293 |
| | 4 | 5,90* | 0,616 | 3,72** | 0,337 |
| | 5 | 7,35* | 0,674 | 4,93** | 0,432 |
| | 6 | 7,67* | 0,675 | 5,00** | 0,431 |
| | 7 | 8,26* | 0,747 | 5,68** | 0,523 |
| | 8 | 8,49* | 0,683 | 5,80** | 0,473 |
| | 9 | 9,82* | 0,805 | 6,63** | 0,601 |
| La Niña | 10 | 9,57* | 0,755 | 6,37** | 0,518 |
| | 1 | 2,48* | 0,192 | 1,96** | 0,142 |
| | 2 | 3,73* | 0,360 | 2,90** | 0,264 |
| | 3 | 5,10* | 0,482 | 3,88** | 0,357 |
| | 4 | 6,14* | 0,556 | 4,87** | 0,462 |
| | 5 | 6,81* | 0,651 | 5,24** | 0,458 |
| | 6 | 7,82* | 0,702 | 6,20** | 0,585 |
| | 7 | 8,97* | 0,784 | 7,04** | 0,589 |
| | 8 | 9,41* | 0,829 | 7,45** | 0,668 |
| | 9 | 9,32* | 0,890 | 7,53** | 0,819 |
| Neutro | 10 | 10,28* | 0,876 | 8,31** | 0,838 |
| | 1 | 2,57* | 0,177 | 1,86** | 0,128 |
| | 2 | 3,93* | 0,340 | 2,73** | 0,212 |
| | 3 | 5,10* | 0,440 | 3,63** | 0,276 |
| | 4 | 5,49* | 0,552 | 3,87** | 0,340 |
| | 5 | 6,51* | 0,600 | 4,69** | 0,392 |
| | 6 | 7,22* | 0,627 | 5,13** | 0,434 |
| | 7 | 7,63* | 0,726 | 5,55** | 0,510 |
| | 8 | 8,15* | 0,787 | 5,92** | 0,554 |
| | 9 | 8,58* | 0,807 | 6,17** | 0,575 |
| 10 | 8,91* | 0,772 | 6,45** | 0,536 | |

* Promedios menores estadísticamente a 12 horas, según prueba t al 5%.

** Promedios menores estadísticamente al 10%, según prueba t al 5%.

faltantes de horas de brillo solar, es flexible dado que se ajusta a las características propias de cada estación y grupo y su estimación depende única y exclusivamente de su historia, por lo tanto puede ser difundido como una técnica viable para la estimación

de ellos, con un error relativo promedio, estadísticamente inferior al 10%, cuando se estiman hasta diez datos faltantes para un mismo mes, independientemente del grupo y un error relativo promedio, estadísticamente inferior al 6%, cuando se estiman hasta

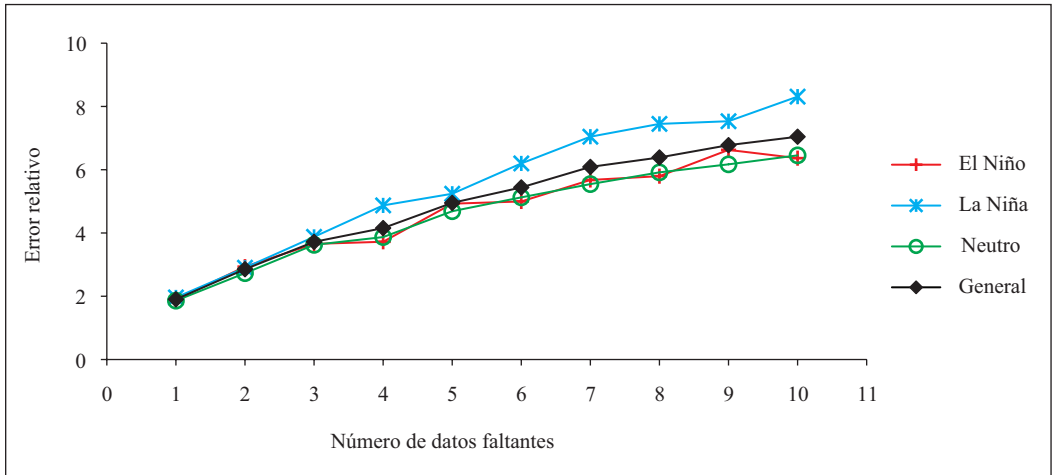


Figura 3. Error relativo promedio de acuerdo con el número de datos faltantes, por evento y en general.

seis datos faltantes para un mismo mes del grupo El Niño o Neutro. Para lograr estos errores de estimación, se requiere de series de brillo solar diaria superiores a 25 años.

LITERATURA CITADA

- CAÑIZARES, M.; BARROSO, I.; ALFONSO, K. Datos incompletos: Una mirada crítica para su manejo en estudios sanitarios. [En línea]. *Gaceta sanitaria* 18(1):58-63. 2004. Disponible en internet: <http://www.scielo.isciii.es/pdf/gsv/v18n1/metodologica.pdf>. Consultado el 4 de Abril de 2009.
- CHAVES C., B. Stochastic modeling of monthly sun bright in coffee growing areas. *Revista colombiana de estadística* 25(1):59-71. 2002.
- CHAVES C., B.; JARAMILLO R., A. Regionalización de la distribución de brillo solar en Colombia por métodos de conglomeración estadística. *Cenicafé* 48(2):120-132. 1997.
- FLORES, B.M.; LAGOS, I.J.; FLORES, J.M. Estudio preliminar de la Heliofanía en Honduras. [en línea]. Disponible en internet: <http://www.bvsde.paho.org/bvsaidis/centroa22/Ponencia25.pdf>. Consultado el 3 de Septiembre de 2008.
- GUZMÁN M., O.; GÓMEZ G., L. Caracterización del brillo solar en Chinchiná, Caldas. *Cenicafé* 48(1):26-39. 1997.
- JARAMILLO R., A. *Clima andino y café en Colombia*. Chinchiná : CENICAFÉ, 2005. 192 p.
- MEDINA R., R.; MONTOYA R., E.C.; JARAMILLO R., A. Estimación estadística de valores faltantes en series históricas de lluvia datos faltantes. *Cenicafé* 59(3):260-272. 2008.
- NATIONAL WEATHER SERVICE. Climate prediction center. [En línea]. Disponible en Internet: http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml. Consultado en Marzo de 2010.
- RAMÍREZ B., V.H.; JARAMILLO R., A.; PEÑA Q., A.; VALENCIA, A.J. El brillo solar en la zona cafetera colombiana, durante los eventos El Niño y La Niña. Chinchiná : CENICAFÉ, 2012. 11 p. (Avances Técnicos No. 421).
- SEMENOV, M.; BARROW, E.M.; RICHARDSON, C.W. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. [En línea]. Disponible en internet: <http://www.rothamsted.ac.uk/mas-models/SupportFiles/Papers/ComparisonWGENandLARSWG.pdf>. Consultado el 5 de Noviembre de 2008.