

Modelos de regresión simple para variables radiométricas en horas bajo cielo despejado en la ciudad de San Lorenzo - Paraguay

Simple regression models for radiometric variables in hours under clear skies in the city of San Lorenzo – Paraguay

César Daniel Amarilla¹

¹Universidad Nacional de Asunción, Facultad de Ciencias Exactas y Naturales, Departamento de Matemática y Estadística, Campus Universitario, San Lorenzo-Paraguay. Email: ceamarilla@gmail.com.

Resumen: La radiación solar es responsable de la evolución de la vida en el planeta Tierra. Sufre procesos de atenuación en su paso por la atmósfera. Esta atenuación se debe a diversos factores atmosféricos y meteorológicos. Las variables radiométricas, radiación global y ultravioleta tipo B total, cuyas intensidades son generalmente mayores bajo cielo despejado, son variables que evolucionan en el tiempo de manera cíclica afectando de manera directa la salud de los seres humanos. Por esta razón, resulta necesario un estudio estadístico acerca de la evolución de estas variables radiométricas en las horas con cielo despejado. El análisis de regresión proporciona métodos estadísticos válidos que permiten la construcción de modelos, que constituyen formas funcionales, que describen el comportamiento interno de un fenómeno real en ciertas condiciones. Esta investigación plantea y analiza diversos modelos de regresión simple obtenidos a partir del análisis de datos horarios referentes a las radiaciones global y ultravioleta tipo B total, que fueron registrados en el Laboratorio de Investigaciones de la Atmósfera y Problemas Ambientales (LIAPA) de la Facultad de Ciencias Exactas y Naturales (FACEN). Los modelos construidos relacionan, en primer lugar, a la radiación global con la altura solar y, en segundo término, al índice de transmisividad de la UVB total con la altura solar.

Palabras clave: *análisis de regresión simple, variables radiométricas, altura solar, índice de transmisividad de la radiación UV-B total, radiación global.*

Abstract: Solar radiation is responsible for the evolution of life on planet Earth. It undergoes attenuation processes as it passes through the atmosphere. This attenuation is due to various atmospheric and meteorological factors. The radiometric variables, global and ultraviolet radiation type B total, whose intensities are generally greater under clear skies, are variables that evolve in the time of cyclical way affecting directly the health of the human beings. This is why a statistical study is necessary about the evolution of these variables in the open sky hours. Regression analysis provides valid statistical methods that allow the construction of models, which constitute functions forms, which describe the internal behavior of a real phenomenon under certain conditions. This research raises and analyzes several simple regression models obtained from the analysis of hourly data concerning global and ultraviolet radiation type B total, which were registered in the Laboratory of Atmospheric Research and Environmental Problems (LIAPA) of the Faculty of Exact and Natural Sciences (FACEN). The constructed models relate, firstly, the global radiation with the solar height and, secondly, the transmissivity index of the total UVB with the solar height.

Keywords: *simple regression analysis, radiometric variables, solar height, transmissivity index radiation UV-B total, global radiation.*

Introducción

Esta investigación se ha realizado con el objetivo de determinar modelos estadísticos utilizando técnicas del análisis de regresión simple que involucren a la radiación global y la ultravioleta tipo B total, mediante el empleo de series de datos registrados en la estación LIAPA ubicada en la ciudad de San

Lorenzo-Paraguay. Para ello, se han establecido criterios para la caracterización de la evolución temporal de las variables involucradas y para la aplicación y validación de los modelos estadísticos obtenidos. La teoría necesaria para la construcción, aplicación y validación de los modelos estadísticos presentados se basa en el trabajo de investigación

Editor responsable: Nery López Acosta*

*Universidad Nacional de Asunción, Facultad de Ciencias Exactas y Naturales, Dirección de Investigación, San Lorenzo, Paraguay.

Recibido: 19/08/2023

Aceptado: 22/11/2023



2078-399X/2024 Facultad de Ciencias Exactas y Naturales - Universidad Nacional de Asunción, San Lorenzo, Paraguay.

Este es un artículo de acceso abierto bajo la licencia CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/deed.es>).

realizado por Amarilla (2015).

Uno de los temas de interés abordado a nivel mundial es la intensidad con que la radiación solar incide sobre la Tierra. Esta principal fuente de energía, altamente variable, depende de ciertos factores meteorológicos y la posición geográfica del lugar, razón por la cual constituye un factor altamente dependiente del tiempo y del espacio. Además, es directa o indirectamente la responsable de todos los fenómenos que afectan a la meteorología y climatología en la Tierra, y esto a su vez afecta a los seres vivos de dicho planeta (Díaz, 2006). Un tipo de radiación solar que puede ofrecer beneficios como vitamina D, o causar riesgos a la salud bajo alta exposición, a las personas es la radiación ultravioleta (UV).

La radiación UV se divide en las franjas ultravioleta tipo C, abarcando longitudes de onda que van de 100 a los 280 nanómetros, tipo B, que abarca el intervalo espectral 280 a 320 nanómetros, y tipo A, con intervalo espectral 280 a 400 nanómetros (Antón Martínez, 2007). La intensidad de la UV depende de la hora del día y época del año, de la latitud (es más intensa entre el ecuador y los trópicos), de la altura solar (se incrementa con la altura), del espesor de la capa de ozono (a mayor concentración de O_3 menor UV-B), del clima (a mayor cielo despejado mayor radiación incidente), de la contaminación atmosférica (mayor contaminación, menor radiación), del horizonte (a más amplios horizontes corresponde mayor radiación) y del albedo (el pasto refleja menos la UV que la nieve) (Zaratti Sacchetti et al., 2003).

Debido a la influencia de la radiación solar en la evolución de la vida en la Tierra, a nivel mundial se han realizado diversas investigaciones con miras a comprender y predecir su comportamiento. Es casi un 8,3% del espectro solar extraterrestre y un 5% de la radiación que llega a la superficie de la Tierra (Varo Martínez, 2006), es un factor natural de gran interés para la comunidad mundial debido a que es una de las más energéticas de las capas bajas de la atmósfera, razón por la cual influye enormemente sobre ella y los seres vivos (Mateos Villán, 2008). Abarca las longitudes de onda más cortas del espe-

ctro solar, y es responsable de procesos biológicos en los seres vivos, y físico-químicos en la atmósfera (Román Díez, 2010).

Como en Paraguay se registran altas temperaturas durante la mayor parte del año, en el marco de disponer una información continua y actualizada, el Laboratorio de Investigaciones de la Atmósfera y Problemas Ambientales (LIAPA) ejecuta periódicamente mediciones sobre la intensidad de las variables radiométricas en la ciudad de San Lorenzo - Paraguay.

Materiales y métodos

Descripción de la zona de estudio

La estación meteorológica automática, cuya instalación se encuentra ubicada en el Laboratorio de Investigaciones de la Atmósfera y Problemas Ambientales (LIAPA) de la FACEN de la UNA en el Campus Universitario de la ciudad de San Lorenzo - Paraguay, está localizada a 25,34° de latitud sur, 57,52° de longitud oeste y altitud de 123 m.snm. El LIAPA cuenta actualmente con un área aproximada de 80 m². Las mediciones realizadas por los sensores se almacenan en un Microprocesador y Sistema de Almacenamiento de datos.

Recolección de datos

La muestra tomada para llevar adelante esta investigación está compuesta por 1162 datos registrados, en intervalos de tiempo de una hora, de las intensidades de las radiaciones global y ultravioleta tipo B total, en el periodo 1 de enero de 1997 al 31 de diciembre de 2012, en el LIAPA de la FACEN de la UNA.

Para medir la irradiancia solar global en una superficie horizontal, el LIAPA utiliza un piranómetro CMP 11 de marca Kipp & Zonen. Este piranómetro utiliza la tecnología de detección de temperatura y es adecuado para redes meteorológicas y el tiempo de respuesta reducido de 1,66 segundos (63%) cumpliendo con los requisitos para aplicaciones de energía solar. Mientras que para medir la irradiancia solar UV-B total, el LIAPA utiliza un piranómetro UVB-1. Este sensor es utilizado para medir la efectividad biológica de la radiación UV-B solar.

Las técnicas de medición que utiliza el instrumento emplean una combinación de filtros de vidrios de colores para bloquear toda la luz visible del sol y fósforo fluorescente sensible que convierte la luz UV-B a luz visible (Verde). La luz verde resultante es medida con un foto detector de estado sólido. Es un instrumento rústico, diseñando para asegurar la estabilidad de su operación durante largo tiempo, sin necesidad de mantenimiento en las estaciones de campo. Este sensor está estabilizado a una temperatura de 40°C.

Análisis de regresión simple

Todo modelo estadístico de regresión simple es un modelo de carácter empírico y probabilístico que contiene una variable aleatoria dependiente, una variable fija o matemática independiente y parámetros poblacionales, que tratan de relacionar a las variables experimentalmente entre sí, y ajustarlas a funciones que arrojen una alta correlación y bajos errores de estimación. Algunos inconvenientes que presenta este modelo empírico es que no poseen validez universal, solamente es válido localmente, o debe ser reajustado para distintas localidades y en ocasiones carece de sentido físico, aunque los resultados que arroje en su proceso de construcción sean buenos.

Un modelo de regresión simple es una forma funcional dada por $Y = f(X, \varepsilon, \beta_0, \beta_1)$ donde: Y es la variable aleatoria llamada **variable de respuesta**, X es la variable no aleatoria llamada **variable predictor o independiente** considerada como fija dentro del modelo, cuyos valores son fijados por el experimentador, ε es el término de perturbación o **error aleatorio**, que obedece a factores de caracteres aleatorios o variaciones de caracteres sistemáticos a causa de errores en el modelo. Por su parte, $\beta_i, \forall i=0,1$ son conectores o **parámetros poblacionales**, que determinan los aportes de las variables predictoras o transformaciones de ellas en la variable de respuesta. Este modelo puede ser estimado mediante diversos métodos. Los modelos de regresión simple planteados en esta investigación fueron construidos mediante el método de mínimos cuadrados. Según Canavos (1998), este

método realiza las estimaciones de los parámetros intervenientes en el modelo seleccionado mediante la minimización de la suma de los cuadrados de las diferencias entre los valores observados y estimados, mediante el uso del modelo selecto, de la variable de respuesta. La diferencia de cada valor observado y cada valor estimado de la variable respuesta, recibe el nombre de **residual**, es una estimación de ε y se expresa como:

$$e_i = y_i - \hat{y}_i, \text{ para } i = 1, 2, \dots, n \quad (1)$$

De acuerdo al análisis de regresión, cada modelo de regresión simple obtenido debe ser diagnosticado a través del cumplimiento de ciertos supuestos. El cumplimiento de cada uno de los supuestos, en cada modelo planteado, es revisado mediante el **análisis de los residuales**. Los supuestos, según Hair et al. (1999), son:

- Normalidad de los términos de error: Establece que la distribución de probabilidad de los errores aleatorios es normal con media cero y varianza σ^2 .
 - Homogeneidad de varianzas de los términos de error: Establece que los errores aleatorios tienen varianzas constantes. Es decir,
- $$Var(\varepsilon_i) = \sigma^2, \text{ para } i = 1, 2, \dots, n \quad (2)$$
- Independencia de los términos de error: Establece que los errores no están correlacionados. Es decir, cada variable predictor es independiente.
 - Linealidad del fenómeno: Representa el grado de cambio en la variable dependiente asociado con la variable independiente.

Modelos polinómicos, potencial y exponencial de regresión

Los modelos de regresión simple construidos a partir de los datos disponibles fueron de los tipos polinómicos, potencial y exponencial. Según Montgomery et al. (2002), un modelo polinómico

de k-ésimo orden en una variable se define como:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_k X^k + \varepsilon \quad (3)$$

donde ε es el componente aleatorio con media cero y varianza σ^2 . Este modelo es de regresión lineal, ya que es lineal en los parámetros y en la variable de respuesta.

Los modelos potencial y exponencial de regresión simple se definen como:

$$Y = \beta_0 X^{\beta_1} \varepsilon \quad (4)$$

$$Y = \beta_0 \exp(\beta_1 X) \varepsilon \quad (5)$$

Los modelos anteriores constituyen modelos de regresión simple no lineales, ya que no son lineales en los parámetros, pero linealizables bajo transformaciones logarítmicas. La componente aleatoria ε es multiplicativa. Las linealizaciones de los modelos potencial y exponencial están dadas por las expresiones:

$$\ln(Y) = \ln(\beta_0) + \beta_1 \ln(X) + \ln(\mu) \quad (6)$$

$$\ln(Y) = \ln(\beta_0) + \beta_1 X + \ln(\mu) \quad (7)$$

donde $\ln(\varepsilon)$ es una componente aleatoria con media cero y varianza σ^2 . Si se trabaja con estos modelos, se está estimando la media geométrica y no la media aritmética.

Estimaciones de los parámetros de un modelo de regresión lineal simple por Mínimos Cuadrados

En el método de mínimos cuadrados, las estimaciones se realizan mediante la minimización de la suma de los cuadrados de los errores. De esta manera, si un modelo de regresión lineal simple muestral es de la forma:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \forall i = 1, 2, \dots, n \quad (8)$$

este método minimiza la suma de los cuadrados de los errores respecto a los parámetros β_0 y β_1 , y la función de mínimos cuadrados que se expresa como:

$$\sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)]^2,$$

$$\text{para } i = 1, 2, \dots, n \quad (9)$$

Estimación de la varianza del error

En un modelo de regresión simple polinómico lineal, potencial o exponencial se estiman 2 parámetros, y en caso polinómico cuadrático o cúbico 3 o 4, a través de la minimización de la suma de los cuadrados de los errores. De esta forma, al estimar los p parámetros, con $p=2,3,4$, quedan $n-p$ características libres, *grados de libertad*, y se tiene que la estimación insesgada de σ^2 , a través de la suma de los cuadrados de los residuales y los grados de libertad se expresa como:

$$\widehat{\sigma^2} = \text{CME} = \frac{1}{n-p} \sum_{i=1}^n e_i^2 \quad (10)$$

La expresión anterior recibe el nombre de *Cuadrado Medio del Error* (CME) y su raíz cuadrada, *error estándar de estimación* (RCME).

Análisis de varianza y el coeficiente de determinación

El análisis de la varianza consiste en un procedimiento estadístico que divide la variabilidad total de la variable dependiente mediante la utilización de un modelo de regresión estimado, de la siguiente manera:

$$SCT = SCR + SCE \quad (11)$$

El primer miembro de la ecuación anterior se denomina Suma de Cuadrados Total (SCT) y mide la variabilidad total de las observaciones. Mientras que la primera y segunda parte del segundo miembro, son denominadas Suma de Cuadrados Residual (SCE) y Suma de Cuadrados de la Regresión (SCR) respectivamente. La suma de cuadrados de regresión mide la parte de la variabilidad total explicada y la suma cuadrados de error la variabilidad no explicada por la ecuación de regresión. El cociente entre la suma cuadrado de regresión y la suma cuadrado total se la denomina *coeficiente de determinación* (R^2) y es el que mide la proporción de variabilidad de la variable de respuesta explicada por la variable independiente considerada en el modelo de regresión. El coeficiente de determinación se expresa como:

$$R^2 = \frac{SCR}{SCT} \times 100 \% \quad (12)$$

La estadística de Durbin-Watson

La estadística de James Durbin y Geoffrey Watson supone que los errores están autocorrelacionados mediante un modelo autorregresivo de primer orden. Esto es,

$$\varepsilon_t = \rho \varepsilon_{t-1} + \eta_t, \text{ para } t \geq 2 \quad (13)$$

donde η_t , denominada ruido blanco, es libre de autocorrelación y ρ es la constante que mide la correlación entre los errores y se estima mediante:

$$\hat{\rho} = \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2} \quad (14)$$

donde e_{t-1} y e_t son los residuos en los instantes $t-1$ y t . La autocorrelación puede ser negativa, cero o positiva y se mide por medio del estadístico d que está dada por:

$$d \approx 2 - 2\hat{\rho} \quad (15)$$

Al detectarse autocorrelación, Canavos (1998) muestra la utilización de un proceso iterativo que se debe a Cochrane y Orcutt (1949), y que plantea la transformación:

$$y_t - \rho y_{t-1} = \beta_0 (1 - \rho) + \beta_1 (x_t - \rho x_{t-1}) \quad (16)$$

Entonces, el modelo de regresión lineal transformado libre de autocorrelación y los ajustes de los parámetros originales se expresan como:

$$y_t^* = \beta_0^* + \beta_1^* x_t^* + \eta_t, \text{ para } t = 2, 3, \dots, n \quad (17)$$

$$\beta_0^* = \frac{\beta_0^*}{1 - \rho} \text{ y } \text{Var}(\beta_0^*) = \frac{\text{Var}(\beta_0^*)}{(1 - \rho)^2} \quad (18)$$

$$\beta_i^* = \beta_i^* \text{ y } \text{Var}(\beta_i^*) = \text{Var}(\beta_i^*), \\ \text{para } i = 1, 2, 3, \dots, p \quad (19)$$

Criterio de Información de Akaike (AIC)

El criterio de Información de Akaike estudia la calidad relativa de un modelo para un conjunto

dado de datos. Además, constituye un medio para la selección del modelo con mejor ajuste a los datos y se basa en la entropía de la información, que ofrece una estimación relativa de la perdida de información cuando se utiliza un modelo determinado para representar el proceso que realmente genera un conjunto de datos. El mejor modelo es el que tiene el valor mínimo de AIC , que es utilizado para analizar la comparación de las bondades de ajustes de los modelos de regresión e involucra la suma cuadrado del error (SCE) y el número de parámetros estimados (k).

$$AIC_p = n \log \left(\frac{SCE}{n} \right) + 2k \quad (20)$$

Validación cruzada en k iteraciones

La validación cruzada permite estudiar la estabilidad dentro de una determinada muestra considerada para un análisis estadístico. Se utiliza preferentemente en situaciones en donde el objetivo principal es la predicción y se quiere estudiar la precisión de un modelo, construido a partir de una muestra, que se utilizará para la predicción. Consiste en la división de los datos de muestra en dos conjuntos complementarios, una denomina muestra de entrenamiento y la otra de prueba. La muestra de entrenamiento se utiliza para estimar los parámetros correspondientes de un determinado modelo y la de prueba para la validación del mismo.

Medias de las diferencias relativas y absolutas relativas

Las medias de las diferencias relativas (MBE) y absolutas relativas ($MABE$) son indicadores de la bondad de ajuste de los modelos estimados y del error cometido al realizar predicciones de la variable de respuesta a través de un modelo de ajuste respectivamente. Estos indicadores están dados en la siguiente expresión:

$$MBE = \frac{1}{n} \sum_{i=1}^n \frac{(y_i - \hat{y}_i)}{y_i} \text{ y } MABE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (21)$$

donde y_i es un valor de la variable de respuesta obtenida a través de mediciones y \hat{y}_i es el corre-

spondiente valor estimado a través de un modelo de ajuste.

Variables involucradas en los modelos de regresión simple

Los modelos de regresión simple fueron formulados mediante variables que se construyeron a partir de las variables radiométricas, radiación global y UV-B total, y de parámetros astronómicos (Amarilla, 2015).

Índice de claridad o transmisividad solar atmosférica horaria

Esta variable refleja la influencia de los factores atenuantes, en especial el de las nubes y aerosoles, sobre la radiación global antes de incidir sobre la superficie terrestre. Se define la transmisividad solar horaria mediante la siguiente ecuación:

$$k_t^h = \frac{G_{med}^h}{E_{ext}^h} \quad (22)$$

donde G_{med}^h es la irradiancia horaria de la radiación global registrada en el LIAPA y E_{ext}^h es la irradiancia horaria de la radiación solar extraterrestre y viene dada por:

$$E_{ext}^h = I_{sc} E_0 \cos(SZA_M) \quad (23)$$

donde I_{sc} es la constante solar extraterrestre y toma un valor de $1366,1 \text{ Wm}^{-2}$, dando una energía por unidad de área de 4917960 Jm^{-2} por cada hora según el espectro de Gueymard (2003). E_0 constituye el factor de corrección de la excentricidad de la órbita terrestre, cuya expresión en serie de Fourier, según Varo Martínez (2006), es:

$$E_0 = 1,00011 + 0,034221 \cos(\Gamma) + 0,00128 \sin(\Gamma) + 0,000719 \cos(2\Gamma) + 0,000077 \sin(2\Gamma) \quad (24)$$

con Γ medido en radianes, denominado *ángulo diario*, y está relacionado con el día juliano n (sucesivamente se cuentan los días hasta 365 o 366) de la siguiente manera:

$$\Gamma = \frac{2\pi(n-1)}{361} \quad (25)$$

Por último, SZA_M es el ángulo cenital, en el punto medio de un intervalo de tiempo de una hora, medido en radianes. Este parámetro astronómico, siguiendo a Antón Martínez (2007), se puede calcular mediante la expresión:

$$\cos(SZA_M) = \sin(\varphi) \sin(\delta) + \cos(\varphi) \cos(\delta) \cos(\omega_M) \quad (26)$$

En la ecuación anterior, φ , δ y ω_M constituyen la latitud del lugar, declinación de la Tierra y el ángulo horario en el punto medio de un intervalo de una hora. La declinación se puede estimar a través de la serie de Fourier dada por:

$$\begin{aligned} \delta = & 0,00692 - 0,39991 \cos(\Gamma) + 0,07026 \sin(\Gamma) \\ & - 0,00676 \cos(2\Gamma) + 0,00091 \sin(2\Gamma) - 0,0027 \cos(3\Gamma) \\ & + 0,00148 \sin(3\Gamma) \end{aligned} \quad (27)$$

En ángulo horario se estima mediante la expresión

$$\omega_M = \frac{\pi}{12} t - \pi \quad (28)$$

Con t denominado *tiempo solar verdadero* y cuya expresión está dada por:

$$t = t_{civil} + \frac{L_{ref}}{15} \frac{L_{local}}{60} + \frac{E_t}{60} \quad (29)$$

donde t_{civil} , L_{ref} y L_{local} constituyen la hora local, la longitud del mediano de referencia y la longitud del mediano del lugar respectivamente. E_t es denominada la *ecuación del tieepo* y es estimada por serie de Fourier mediante la expresión:

$$\begin{aligned} E_t = & 229,18 [0,000075 + 0,001868 \cos(\Gamma) \\ & + 0,032077 \sin(\Gamma) - 0,014615 \cos(2\Gamma) \\ & - 0,040849 \sin(2\Gamma)] \end{aligned} \quad (30)$$

Índice de transmisividad de la UV-B total

Este índice mide la modulación que sufre la radiación UV-B total en su paso por la atmósfera. Se define la transmisividad ultravioleta tipo B total en el intervalo de tiempo de una hora de la siguiente manera:

$$k_{UVB}^h = \frac{E_{UVB_{med}}^h}{E_{UVB_{ext}}^h} \quad (31)$$

siendo $E_{UVB_{med}}^h$ la irradiancia total en el intervalo de una hora de la radiación UV-B total registrada en la estación LIAPA y $E_{UVB_{ext}}^h$ la irradiancia total en el intervalo de una hora de la UV-B total extra-terrestre, y que puede ser calculada considerando, la parte de la constante solar que corresponda a la banda de la UV-B total y la ecuación:

$$E_{UVB_{ext}}^h = I_{sc_{UVB}} E_0 \cos(SZA_M) \quad (32)$$

donde $I_{sc_{UVB}}$ es la constante solar correspondiente a la irradiancia de la UV-B total y que toma un valor de $21,30 \text{ Wm}^{-2}$ aproximadamente, dando una energía por unidad de área de 76680 Jm^{-2} por cada hora según el espectro de Gueymard (2003).

- Altura solar y masa óptica relativa**

La altura solar es un factor de gran influencia sobre la radiación solar a través del tiempo. Esto se debe a que la cantidad de absorbentes y dispersores de las irradiancia ultravioleta tipo B total presentes en la atmósfera es diferente para cada altura solar (Antón Martínez, 2007). Las revisiones bibliográficas de investigaciones hechas sobre radiación solar muestran que el indicador de la influencia de la altura solar sobre la radiación UV-B total es la **masa óptica relativa** simbolizada por m_r y que se

determina mediante la expresión:

$$m_r = \frac{1}{\cos(SZA)} \quad (33)$$

Software informático

Para la construcción de los modelos estadísticos y el análisis de regresión simple, se ha procedido a la utilización del paquete estadístico SPSS versión 11.0, cuya licencia pertenece a la FACEN de la UNA, y el programa R, función RCommander.

Resultados y discusión

Caracterización de las evoluciones temporales de las variables radiométricas

Para cada variable radiométrica, el análisis exploratorio de los datos da cuenta de los siguientes resultados.

Evolución temporal de la radiación global

De acuerdo a la Fig. 1, la evolución temporal de la radiación global a lo largo del periodo 1997 al 2012 presenta picos máximos en los meses de verano y picos mínimos en los meses de invierno. Su comportamiento evolutivo es sinusoidal, debido a la variación de la altura solar en los diferentes meses del año.

Los meses de diciembre y enero fueron los que arrojaron los valores promedios de radiación global más altos; 23027 kJm^{-2} y 23074 kJm^{-2} re-

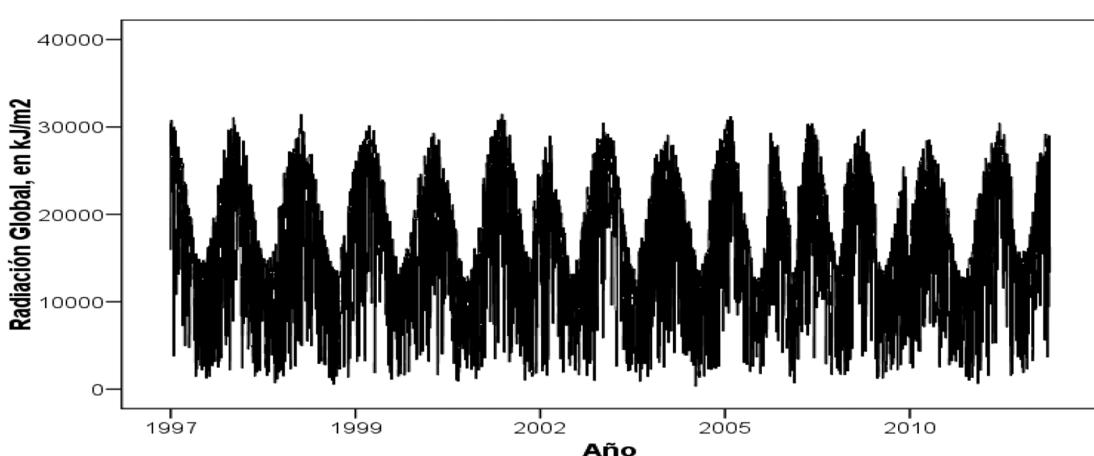


Figura 1. Evolución diaria de la radiación global, en kJm^{-2} , en el periodo 1997 al 2012.

Regresión simple para variables radiométricas bajo cielo despejado en San Lorenzo

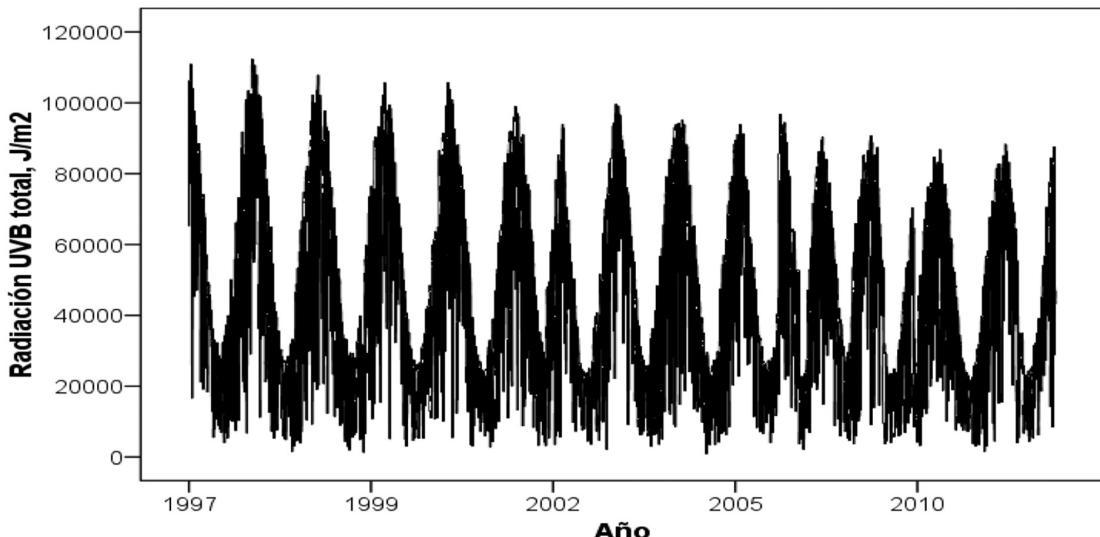


Figura 2. Evolución diaria de la radiación UV-B total, en Jm^{-2} , en el periodo 1997 al 2012.

spectivamente; ya que en estos meses se presentan generalmente mayor intensidad de irradiancia solar y es muy importante adoptar precauciones durante la exposición a los rayos solares. Mientras que los más bajos ocurren en los meses de junio y julio, con 9392 kJm^{-2} y 10532 kJm^{-2} respectivamente. Se observaron valores máximos de 3980 y 3989 kJm^{-2} , que se registraron entre las 11:00 y 13:00 horas el 27 de enero de 1997 y 14 de diciembre de 2012 respectivamente. El menor valor horario fue registrado el 14 de noviembre de 2001 entre las 7:00 y 8:00 horas, y fue de 5 kJm^{-2} .

La variabilidad de la radiación global horaria se puede apreciar a través de la desviación estándar y el rango intercuartil. De acuerdo a los valores un tanto elevados en relación a la media y mediana de estos estadísticos y que el coeficiente de variación en todos los casos es mayor al 40%, se pueden apreciar dispersiones horarias importantes en los distintos intervalos horarios.

Por su parte, se puede constatar que las diferencias entre los valores mínimos y los percentiles 5%; y las diferencias entre los valores máximos y los percentiles 95% son muy elevadas, razón por lo cual, los valores mínimos y máximos no son representativos dentro de las distribuciones horarias de la radiación global, pudiendo representar valores

atípicos en la distribución horaria general de datos.

Evolución temporal de la radiación UV-B total

De acuerdo a la Fig. 2, la radiación UV-B total también presenta un comportamiento del tipo sinusoidal con picos máximos y mínimos en los meses de verano e invierno respectivamente.

El mes de enero fue el que arrojó el valor promedio de UV-B más alto; 77438 Jm^{-2} y se debe a que es el mes veraniego que presenta generalmente mayor intensidad de radiación solar. Mientras que el más bajo tuvo lugar en el mes de Junio, 19319 Jm^{-2} ya que es uno de los meses de invierno en el que la irradiancia solar generalmente no es tan intensa. Los valores diarios que muestran las mayores intensidades se registran en los meses de enero y diciembre, mientras que los más bajos en los meses de mayo y agosto. Por su parte, las intensidades horarias más altas se registraron entre las 9:00 y 15:00 horas. Los valores máximos encontrados fueron de 18193 y 18026 Jm^{-2} ; que se registraron los días 11 y 13 de enero de 1997 en los intervalos horarios de 12:00 a 13:00 horas y de 11:00 a 12:00 horas, respectivamente. El menor valor horario fue de 11 Jm^{-2} , registrado entre las 17:00 y 18:00 horas del día 21 de diciembre de 2010.

Debido a que los valores de las desviaciones

estándares y los rangos intercuartílicos son elevados en relación a las medias y medianas, además de que el coeficiente de variación en todos los casos es mayor al 40% y llegando a casi 70% en algunos casos, se observaron grandes dispersiones en los distintos intervalos horarios.

En el análisis de los valores máximos y mínimos horarios, se constatan que las diferencias, entre los valores mínimos y los percentiles 5% por un lado y las diferencias entre los valores máximos y los percentiles 95% por otro, son muy elevadas. Entonces, los mínimos y máximos no son representativos dentro de las distribuciones horarias de la radiación UV-B total, pudiendo representar valores atípicos en la distribución general de datos de esta variable.

Aplicación del análisis de regresión simple para horas bajo cielo despejado

Modelos de regresión simple propuestos para las variables radiométricas en horas bajo cielo despejado

Por medio de las horas con cielo despejado se puede determinar la influencia de la altura solar sobre las variaciones de las variables radiométricas. Varios investigadores han propuestos modelos polinómicos, potenciales y exponenciales que reflejan la influencia de la altura solar sobre la radiación global. Así, por ejemplo, Román Diez (2010) propone un modelo polinómico de segundo grado para establecer la influencia de la altura solar sobre la radiación global. En esta investigación, se proponen modelos polinómicos de primer, segundo y tercero grado. Además, se presentan modelos de los tipos potencial y exponencial, ambos linealizables mediante la transformación logarítmica:

$$G_h = \beta_0 + \beta_1 \cos(SZA) + \beta_2 \cos^2(SZA) + \beta_3 \cos^3(SZA) + \varepsilon_h \quad (34)$$

$$G_h = \beta_0 \cos^{\beta_1} (SZA) \varepsilon_h \quad (35)$$

$$G_h = \beta_0 \exp[\beta_1 \cos(SZA)] \varepsilon_h \quad (36)$$

En la Tabla 1 se puede observar el valor estimado de los estadísticos R^2 y AIC para los modelos poli-

Tabla 1. Estadísticos R^2 y AIC para los modelos polinómicos que relacionan a la irradiancia global y al ángulo cenital en horas con cielo despejado.

Modelos	n	R^2	AIC
Lineal	1162	0, 989	11313, 214
Cuadrático	1162	0, 990	10119, 130
Cúbico	1162	0, 990	10113, 511

nómicos. En todos los casos, el valor del coeficiente de determinación es superior al 98% y el menor valor calculado del AIC corresponde al modelo polinómico de tercer grado, indicando que este polinomio es el de mayor aporte significativo. Mientras que, en la Tabla 2 se puede observar que

Tabla 2. Estadísticos R^2 y $RCME$ para los modelos cubico, potencial y exponencial que relacionan a la irradiancia global y al ángulo cenital en horas con cielo despejado.

Modelos	n	R^2	$RCME$
Cúbico	1162	0, 990	70784, 105
Potencial	1162	0, 992	0, 025
Exponencial	1162	0, 971	0, 049

los coeficientes de determinación para los modelos cúbico, potencial y exponencial son superiores al 97% y que el menor valor en el $RCME$ lo presenta el modelo potencial. Entonces, el modelo que muestra el mejor ajuste es el potencial. Por su parte, la Fig. 3 muestra la dispersión de los datos en horas con

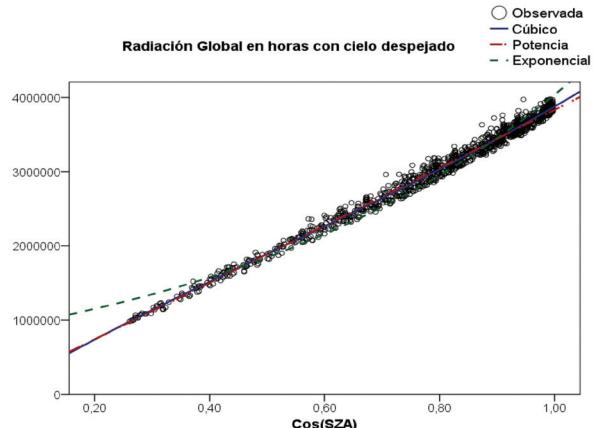


Figura 3. Irradiancia global horaria G_h versus coseno del ángulo cenital, en horas con cielo despejado, en el período 1997 al 2011.

cielo despejado, así como los ajustes de los modelos cubico, potencial y exponencial estimados.

El índice de claridad es una de las variables que puede reflejar la condición del cielo en un cierto intervalo horario. De acuerdo a Román Diez (2010), la condición del cielo según el índice de claridad horario se presenta en la Tabla 3.

Tabla 3. Condición del cielo según el índice de claridad horario k_t^h .

Criterio	Condición de cielo
$0 < k_t^h \leq 0,34$	Hora nubosa
$0,34 < k_t^h \leq 0,75$	Hora parcialmente nubosa
$k_t^h \geq 0,75$	Hora de cielo claro

De acuerdo a la Tabla 3, para proponer y estimar modelos de regresión simple para la UV-B total en horas con cielo despejado se deben utilizar datos para los cuales el índice de claridad horario arroje valores mayores o iguales a 0,75; es decir $k_t^h \geq 0,75$. De acuerdo a Murillo et al. (2003) y Antón Martínez (2007) por un lado y a Dubrovsky et al. (2000) y Antón Martínez (2007) por otro, con el objetivo de analizar el comportamiento de la intensidad de la radiación UV-B total en relación a la altura solar en horas con cielo despejado se proponen los modelos de regresión simple potencial y exponencial que son linealizables bajo transformaciones logarítmicas,

$$k_{UVB}^h = \beta_0 m_r^{\beta_1} \varepsilon_h \quad (37)$$

$$k_{UVB}^h = \beta_0 \exp(\beta_1 m_r) \varepsilon_h \quad (38)$$

La Fig. 4 representa el diagrama de dispersión de datos y en cuyo eje horizontal se ubican los valores de la masa óptica relativa y en cuyo eje vertical se ubican los valores en horas bajo cielo despejado del índice de transmisividad de la UV-B. Las curvas azul y roja representan a los modelos potencial y exponencial estimados por mínimos cuadrados ordinarios.

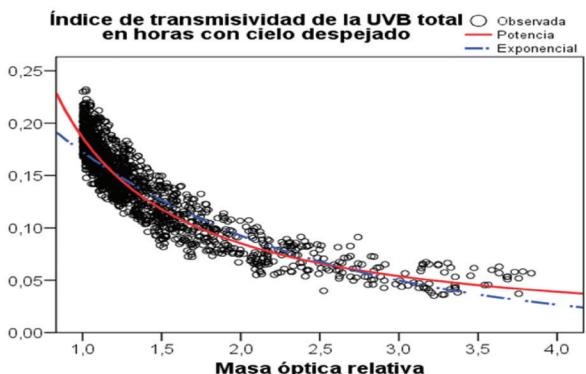


Figura 4. k_{UVB}^h versus m_r , en horas con cielos despejados, en el período 1997 al 2011.

La Tabla 4 muestra los valores obtenidos para los estadísticos R^2 y $RCME$ para los modelos potencial y exponencial. El valor del coeficiente de determinación supera el 80% en ambos modelos y el menor valor calculado del $RCME$ corresponde al modelo potencial, indicando que este modelo es el de mejor ajuste.

Modelos	n	R^2	$RCME$
Potencial	1162	0,880	0,103
Exponencial	1162	0,835	0,120

Tabla 4. Estadísticos R^2 y $RCME$ para los modelos potencial y exponencial que relacionan a las variables k_{UVB}^h y m_r en horas con cielo despejado.

Diagnósticos y validaciones de los modelos de regresión simple lineal aplicados para las estimaciones de las irradiancias de las variables radiométricas en horas bajo cielo despejado

Según el análisis de regresión simple realizado en base a los estadísticos R^2 y $RCME$, los modelos horarios bajo cielo despejado de mejor ajuste fueron potenciales.

Para empezar el proceso de diagnóstico y validación de los modelos potenciales horarios con cielo despejado, resulta conveniente linealizarlos. Por lo que:

$$\ln(G_h) = \ln(\beta_0) + \beta_1 \ln |\cos(SZA)| + \ln(\varepsilon_h) \quad (39)$$

Tabla 5. Contraste de hipótesis realizado sobre el parámetro β_1 para los modelos lineales que relacionan a las variables radiométricas con el ángulo cenital.

Modelo	Coeficientes	Estimación	Error estándar	t_{cal}	$p - valor$
$\ln \cos(\text{SZA}) $	β_1	1, 0215	0,0026	386,9	$2 \cdot 10^{-16}$
$\ln(m_r)$	β_1	-0, 9962	0,0108	-92,3	$2 \cdot 10^{-16}$

$$\ln(k_{UVB}^h) = \ln(\beta_0) + \beta_1 \ln(m_r) + \ln(\varepsilon_h) \quad (40)$$

donde $\ln(\varepsilon_h)$, según uno de los supuestos del análisis de regresión simple, tiene distribución normal con media cero y varianza uno.

El contraste de hipótesis sobre el parámetro β_1 a realizar, en ambos modelos lineales, es:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

El resumen del análisis realizado en base a los datos disponibles para el contraste de hipótesis anterior, en los dos casos, se presenta en la Tabla 5, en el que se observa que en ambos casos el $p - valor = 2 \cdot 10^{-16} << 0,05$, indicando que existen evidencias estadísticamente significativas para rechazar H_0 .

Por otra parte, los respectivos valores del estadístico de Durbin-Watson obtenidos en el análisis son 0,663 y 0,307. Estos valores indican la presencia de autocorrelación positiva entre los errores para los modelos lineales considerados. La solución a este problema, siguiendo las indicaciones de Durbin-Watson, es realizar las estimaciones de los

parámetros $\ln(\beta_0)$ y β_1 mediante de las relaciones:

$$\ln(\beta_0) = \frac{\ln(\beta_0^*)}{1-\rho} \text{ y } \beta_1 = \beta_1^* \quad (41)$$

siendo $\ln(\beta_0^*)$ y β_1^* parámetros de los modelos de regresión lineal transformados:

$$\ln(G_h)^* = \ln(\beta_0^*) + \beta_1^* \ln|\cos(SZA)|^* + \ln(\varepsilon_h)^* \quad (42)$$

$$\ln(k_{UVB}^h)^* = \ln(\beta_0^*) + \beta_1^* \ln(m_r)^* + \ln(\varepsilon_h)^* \quad (43)$$

con $\ln(\varepsilon_h)^*$, denominado ruido blanco, libre de autocorrelación. Además,

$$\ln(G_h)^* = \ln(G_h)_t - \rho \ln(G_h)_{t-1} \quad (44)$$

$$\ln|\cos(SZA)|^* = \ln|\cos(SZA)_t| - \rho \ln|\cos(SZA)_{t-1}| \quad (45)$$

$$\ln(k_{UVB}^h)^* = \ln(k_{UVB}^h)_t - \rho \ln(k_{UVB}^h)_{t-1} \quad (46)$$

Tabla 6. Estimaciones de los parámetros $\ln(\beta_0^*)$ y β_1^* , además de los cálculos de los estadísticos R^2 , $RCME$ y d para los modelos lineales transformados.

Modelo	Coeficientes	Estimación	Error estándar	R^2	$RCME$	d
$\ln \cos(\text{SZA}) ^*$	$\ln(\beta_0^*)$	5, 0312	0, 0006	0,995	0,018	2,18
	β_1^*	1, 0029	0, 0020			
$\ln(m_r)^*$	$\ln(\beta_0^*)$	-0, 2561	0, 0015	0,972	0,0511	2,26
	β_1^*	-1, 0668	0, 0054			

$$\ln(m_r)^* = \ln(m_r)_t - \rho \ln(m_r)_{t-1} \quad (47)$$

$$\ln(\varepsilon_h)^* = \ln(\varepsilon_h)_t - \rho \ln(\varepsilon_h)_{t-1} \quad (48)$$

donde ρ constituye el coeficiente de correlación entre los errores y cuya estimación $\hat{\rho}$ está relacionado con el estadístico d de Durbin-Watson mediante $d \approx 2(1 - \hat{\rho})$.

Las estimaciones por mínimos cuadrados ordinarios de los parámetros y los cálculos de los estadísticos R^2 , $RCME$ y d , de los modelos presentados en las ecuaciones (42) y (43) se observan en la Tabla 6. Los nuevos valores del estadístico d de Durbin-Watson indican la ausencia de autocorrelación de los errores para ambos modelos.

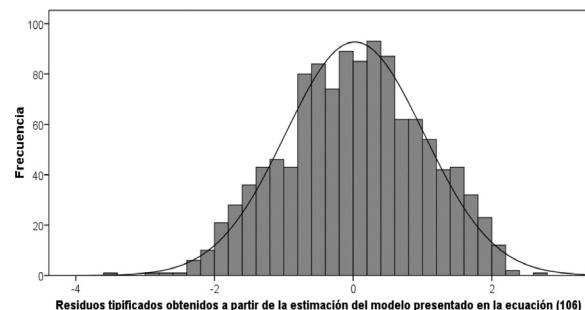
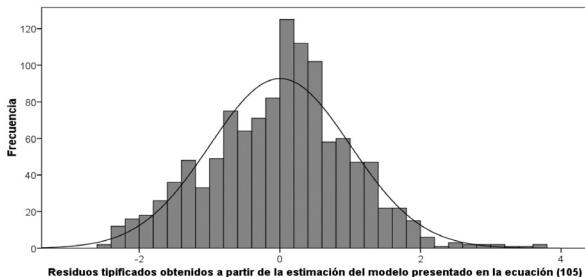


Figura 5. Residuos estandarizados observados en las estimaciones de los modelos lineales que relacionan a las variables radiométricas con el ángulo cenital.

La Fig. 5 muestra los histogramas de los residuos estandarizados observados frente a la curva normal para ambos modelos y la Fig. 6 el comportamiento de dichos residuales frente a los valores de los cuantiles normales esperados. En ambas figuras, se puede observar que los residuales se ubican en su gran mayoría sobre la curva normal y que, además, se presentan leves asimetrías positiva y negativa,

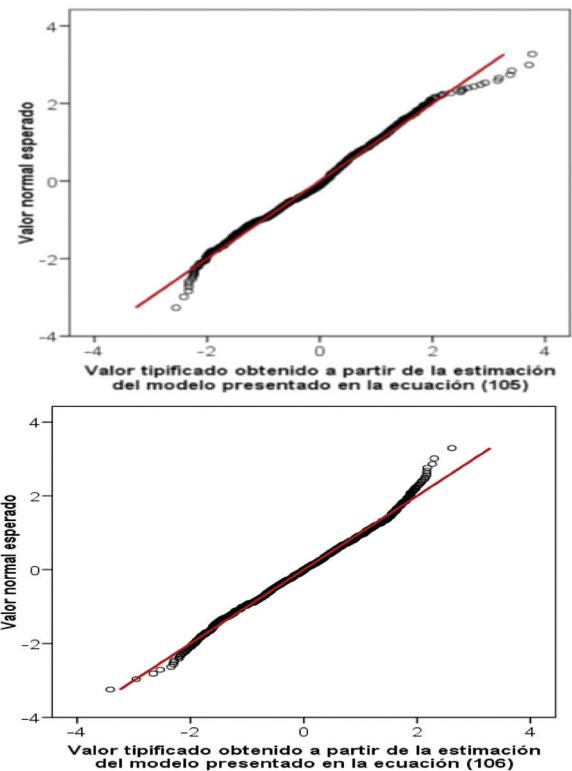


Figura 6. Gráficos Q-Q plot normales de regresión resultantes de las respectivas estimaciones de los modelos lineales que relacionan a las variables radiométricas con el ángulo cenital.

respectivamente. Por lo tanto, según lo observado, podemos asumir que se verifica, de manera aproximada el supuesto de normalidad en ambos modelos.

En base a los resultados anteriores, las estimaciones de los parámetros $\ln(\beta_0)$ y β_1 con sus intervalos de confianza al 95% de fiabilidad se resumen en la Tabla 7.

A continuación, en las regresiones lineales simples, se verifican los supuestos de normalidad, homogeneidad de varianzas y linealidad a través de los gráficos de los residuos estandarizados y estudiantizados.

En cuanto a los supuestos de homogeneidad de varianzas y linealidad de los errores, se puede decir que ambos verifican. En la Fig. 7 se observa que las dispersiones de los residuales frente a los valores predichos no muestran tendencias y casi todas ellas están contenidas entre las franjas -3 y 3.

Al verificar el cumplimiento de los supuestos

Tabla 7. Estimaciones de los parámetros $\ln(\beta_0)$ y β_1 con sus respectivos intervalos de confianza, utilizando las estimaciones de los parámetros $\ln(\beta_0^*)$ y β_1^* .

Modelo	Coeficientes	Estimación	Error estándar	Intervalo de confianza al 95%
$\ln \cos(\text{SZA}) $	$\ln(\beta_0)$	15, 1543	0, 0017	(15, 1511; 15, 1576)
	β_1	1, 0029	0, 0020	(0, 9989; 1, 0068)
$\ln(m_r)$	$\ln(\beta_0)$	-1, 6631	0, 0098	(-1, 6824; -1, 6438)
	β_1	-1, 0668	0, 0054	(-1, 0722; -1, 0563)

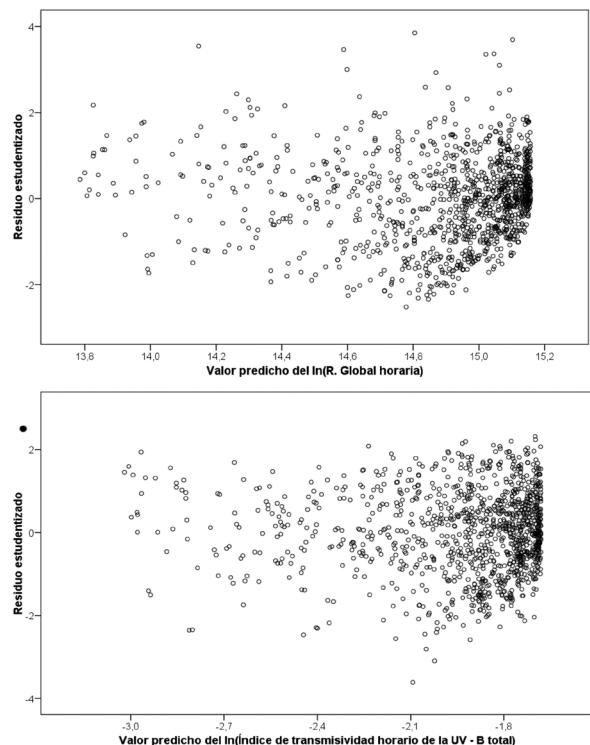


Figura 7: Diagramas de dispersiones de los residuales estandarizados versus los valores predichos obtenidos en las estimaciones por mínimos cuadrados ordinarios de los modelos lineales que relacionan a las variables radiométricas con el ángulo cenital.

básicos, el siguiente paso es el estudio de la estabilidad dentro de la muestra considerada. Para ello, se utiliza el método de la validación cruzada de k iteraciones.

En la Tabla 8 se observa el resumen de las estimaciones de los parámetros involucrados en cada una de las iteraciones. Las estimaciones de los parámetros involucrados no muestran diferencias

significativas en cada iteración, indicando una muy buena estabilidad de la muestra considerada.

Por su parte, la Tabla 9 muestra el resumen de las validaciones de los modelos estimados en cada iteración. Los valores absolutos observados del promedio de las diferencias relativas MBE , inferiores al 0,2% en el primer caso y al 5% en el segundo, muestran ligeros sub o sobre-estimaciones por parte de los modelos estimados en cada iteración. Los valores observados del promedio del valor absoluto de las diferencias relativas $MABE$, indican que las estimaciones en las distintas iteraciones de la radiación global y el índice de transmisividad de la UV-B total en horas con cielo despejado fueron realizadas con errores inferiores al 0,3% en el primer caso y al 6% en el segundo.

Como última etapa del proceso de diagnóstico y validación, se realizan comparaciones a través de diagramas de dispersiones entre datos para horas con cielo despejado, correspondientes al año 2012, que fueron tomados en el LIAPA en el periodo 1 de enero al 31 de diciembre y que no fueron utilizados para las estimaciones de los respectivos modelos, y las correspondientes predicciones para el mismo año obtenidas a partir de dichos modelos. Se cuentan con un total 391 datos.

La Fig. 8 muestra las respectivas asociaciones de los valores medidos y predichos para el año 2012 referentes a las variables $\ln(G_h)$ y $\ln(k_{UVB}^h)$. Las líneas azules corresponden a las respectivas regresiones lineales entre las magnitudes medida y predicha y cuyas expresiones están dadas por las siguientes ecuaciones:

$$\ln(G_h^{\text{predicha}}) = 0,2104 + 0,9855 \ln(G_h^{\text{medida}})$$

Tabla 8: Estimaciones a través de las muestras de entrenamiento para los modelos de las variables radiométricas bajo cielo despejado, en 15 iteraciones.

<i>k</i>	<i>n_e</i>	Modelo $\ln \cos(\text{SZA}) $ Estimaciones				Modelo $\ln(m_r)$ Estimaciones			
		$\ln(\beta_0)$	β_1	<i>RCME</i>	<i>R</i> ²	$\ln(\beta_0)$	β_1	<i>RCME</i>	<i>R</i> ²
1	939	15,1543	1, 0048	0, 018	0, 994	-1, 6875	-1, 0837	0, 051	0, 963
2	1099	15,1546	1, 0019	0, 018	0, 995	-1, 6676	-1, 0638	0,051	0, 971
3	1151	15,1546	1, 0029	0, 018	0, 995	-1, 6626	-1, 0669	0,051	0, 972
4	1148	15,1542	1, 0026	0, 018	0, 995	-1, 6592	-1, 0658	0,051	0, 972
5	1058	15,1539	1, 0022	0, 018	0, 996	-1, 6598	-1, 0635	0,051	0, 974
6	1059	15,1545	1, 0058	0, 018	0, 995	-1, 6607	-1, 0704	0,052	0, 971
7	1042	15,1573	1, 0060	0, 018	0, 995	-1, 6579	-1, 0754	0,051	0, 971
8	1048	15,1532	1, 0019	0, 018	0, 995	-1, 6777	-1, 0651	0,052	0, 972
9	1042	15,1547	1, 0024	0, 018	0, 996	-1, 6532	-1, 0638	0,052	0, 972
10	1117	15,1550	1, 0029	0, 018	0, 995	-1, 6621	-1, 0662	0, 051	0, 973
11	1062	15,1525	1, 0018	0, 018	0, 996	-1, 6563	-1, 0639	0,051	0, 974
12	1048	15,1538	1, 0026	0, 018	0, 995	-1, 6627	-1, 0625	0,051	0, 971
13	1142	15,1546	1, 0024	0, 018	0, 995	-1, 6609	-1, 0669	0,051	0, 971
14	1159	15,1542	1, 0029	0, 018	0, 995	-1, 6630	-1, 0674	0,051	0, 972
15	1154	15,1539	1, 0019	0, 018	0, 995	-1, 6623	-1, 0642	0,050	0, 973

Tabla 9: Validaciones a través de las muestras de prueba para los dos modelos propuestos de las variables radiométricas bajo cielo despejado, en 15 iteraciones.

Año	<i>n_p</i>	Modelo $\ln \cos(\text{SZA}) $		Modelo $\ln(m_r)$	
		<i>MBE</i> (%)	<i>MABE</i> (%)	<i>MBE</i> (%)	<i>MABE</i> (%)
1997	223	-0,020	0,146	-1, 233	3, 926
1998	63	0,073	0,140	3, 804	5, 635
1999	11	0,071	0,139	3, 573	5, 549
2000	14	-0,016	0,099	4, 017	4, 352
2001	104	-0,020	0,105	-2, 567	3, 687
2002	103	-0,045	0,166	-0, 464	4, 722
2003	120	0,141	0,159	-1, 782	3, 824
2004	114	-0,062	0,104	2, 791	3, 239
2005	120	0,045	0,149	-3, 510	4, 016
2006	45	0,079	0,159	-1, 117	2, 663
2007	100	-0,112	0,123	-4, 366	5, 386
2008	114	-0,030	0,117	-1, 078	2, 483
2009	20	0,093	0,174	-2, 876	3, 366
2010	3	0,000	0,203	-3, 625	3, 625
2011	8	-0,183	0,223	-1, 705	3, 862

$$\text{con } R^2 = 0,992 \quad (49)$$

$$\ln(k_{UVB}^{h(\text{predicha})}) = -0,2650 + 0,8151 \ln(k_{UVB}^{h(\text{medida})}) \quad (50)$$

Según el proceso de diagnóstico y validación realizado, las estimaciones de los modelos potenciales para horas con cielo despejado que relacionan a las irradiancias global y UV-B total con la altura solar están dadas por las siguientes expresiones:

$$\widehat{G}_h = 3814607 m_r^{-1,0029} = 3814607 \cos^{1,0029}(SZA) \quad \text{con } R^2 = 0,995 \quad (51)$$

$$\widehat{k}_{UVB}^h = 0,1896 m_r^{-1,0668} = 0,1896 \cos^{1,0668}(SZA) \quad \text{con } R^2 = 0,972 \quad (52)$$

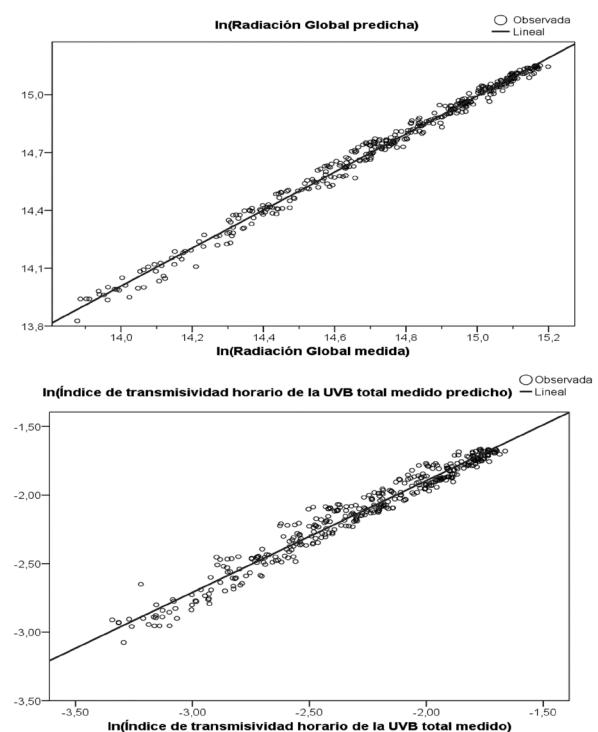
Las variables \widehat{G}_h y \widehat{k}_{UVB}^h son las medias geométricas estimadas de las variables G_h y k_{UVB}^h a partir del ángulo SZA . Estos modelos indican que con cada incremento del 1% de la variable m_r en horas con cielo despejado, se estima que las variables G_h y k_{UVB}^h disminuyen en promedios de 1,003% y 1,067%, respectivamente, y que sus intensidades máximas estimadas son $3814607 Jm^{-2}$ y 0,1896 respectivamente. Estos valores máximos se presentan en las horas con cielo despejado correspondiente a los meses de verano, entre la 11:00 y 12:00 horas locales. Los altos valores de los coeficientes de determinación informan que la variable altura solar, en horas con cielo despejado, explica en más del 97% la variabilidad de las variables radiométricas consideradas.

Debido a la construcción de los modelos potenciales estimados, antes de su utilización, resulta importante tener presente que la interpretación de todo modelo de regresión lineal bajo transformación logarítmica en la variable respuesta debe realizarse en términos de la media geométrica o de la mediana, en lugar de hacerlo en términos de la media aritmética.

Conclusiones

En el proceso de caracterización de las evoluciones temporales de las irradiancias global y UV-B total

Figura 8: Diagramas de dispersiones que reflejan las respectivas asociaciones entre los valores, para horas con cielo despejado, medidas y predichos correspondientes al año 2012 de los logaritmos de las variables radiométricas.



se ha podido constatar que ambas radiaciones tienen un comportamiento sinusoidal con valores máximos muy altos en los meses de verano, alcanzando picos horarios de 3980 y $18,193 kJm^{-2}$ respectivamente, y valores mínimos en los meses invierno, alcanzando valores horarios respectivos de 5 y $0,11 kJm^{-2}$. Estos picos máximos y mínimos son tanto altos debido a la posición geográfica del Paraguay.

En cuanto a la construcción, aplicación y validación de los modelos estadísticos obtenidos mediante el análisis de regresión simple realizado, se propusieron modelos de los tipos polinómicos, potenciales y exponenciales con el objetivo de estimar los comportamientos de las variables radiométricas en horas con cielo despejado, obteniéndose muy buenos resultados. Mediante la utilización de la función RCommander del software estadístico R y el SPSS versión 11.0, los modelos de regresión simple propuestos fueron estimados a través del método

de mínimos cuadrados ordinarios, con altos valores en los coeficientes de determinación y bajos valores en los errores de estimación.

Tomando en cuenta los valores *AIC* en primer término y luego los errores de estimación, se seleccionaron modelos potenciales de regresión, que fueron diagnosticados y validados. En esta etapa, estos modelos potenciales fueron diagnosticados y validados mediante el análisis de los residuales y técnicas como la validación cruzada y la utilización de nuevas muestras. Los coeficientes de determinación obtenidos fueron superiores al 97% y los valores de los errores estándares de estimación calculados de $0,0180 \text{ kJm}^{-2}$ y $0,0511$ respectivamente, con lo cual se pudo constatar que los modelos potenciales estimados ofrecen muy buenos resultados en la estimación de los comportamientos de las variables radiométricas en horas con cielo despejado.

De acuerdo al análisis de regresión simple realizado, resulta factibilidad realizar predicciones acerca de los comportamientos en horas con cielo despejado de las intensidades de las variables radiométricas consideradas en la ciudad de San Lorenzo-Paraguay, mediante modelos potenciales de regresión simple que relacionan a las radiaciones global y UV-B total con la altura solar, construido a partir de mediciones registradas en la estación LIAPA y que muestran los comportamientos de las intensidades en superficie de las variables radiométricas, radiación global y UV-B total, en horas con cielo despejado en la ciudad de San Lorenzo-Paraguay.

Agradecimientos

Al Altísimo, dueño eterno del reino, del poder y la gloria desde que el tiempo no tenía memoria. A este Divino Creador, cuya fuente de luz y de fuerza me impulsa a seguir adelante y me levanta cuando tropiezo por los caminos de la vida.

A mi amada familia. En especial a mi abuela Avelina, a Francisca Ofelia, mis hijas Jessica Larissa y Lisset Fiorella, y mis hermanas Máxima y Nancy por el apoyo de siempre.

Al Laboratorio de Investigaciones de la Atmós-

fera y Problemas Ambientales por cederme tan amablemente los datos que fueron utilizados para la construcción de la base de datos y posterior análisis descriptivo e inferencial.

Financiamiento

El trabajo se realizó con financiamiento propio.

Literatura citada

- Amarilla, C.D. (2015). Modelización de la radiación ultravioleta tipo B total en la ciudad de San Lorenzo – Paraguay. *Reportes Científicos de la FACEN*, 6(2): 5–21.
- Martínez, M.A. (2007). *Modelos empíricos para la estimación de la Irradiancia Solar Ultravioleta*. (Tesis Doctoral). Badajoz: Universidad de Extremadura. 212 pp.
- Canavos, G.C. (1988). *Probabilidad y Estadística. Aplicaciones y Métodos*. (1^a Ed.). México DF: McGraw-Hill/Interamericana. 651 pp.
- Cochrane, D. & Orcutt, G.H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *Journal of the American Statistical Association*, 44(245): 32-61.
- Díaz Rodríguez, A.M. (2006). *Caracterización de los aerosoles atmosféricos y su influencia en los niveles de radiación UV en la región de Canarias*. (Tesis Doctoral). San Cristóbal de la Laguna: Universidad de la Laguna. iv + 279.
- Dubrovsky, M. (2000). Analysis of UV-B irradiances measured simultaneously at two stations in the Czech Republic. *Journal Geophysical Research*, 105: 4907–4913.
- Hair, J.F., Anderson, R.E., Tatham, R.L. & Black, W.C. (1999). *Análisis Multivariante*. (5^a Ed.). Madrid: Pearson Prentice Hall. 799 pp.
- Mateos Villán, D.D. (2008). *Medida, Análisis Climatológico y Modelización de la Radiación Solar Ultravioleta Total en Valladolid*. (Tesis de Maestría). Valladolid: Universidad de Valladolid. 111 pp.
- Montgomery, D.C., Peck, E.A. & Vining, G.G. (2002). *Introducción al análisis de regresión*

- lineal.* (1^a Ed.). México: Compañía Editorial Continental. 588 pp.
- Murillo, W., Cañada, J. & Pedrós, G. (2003). Correlation between global ultraviolet (290-385 nm) and global irradiation in Valencia and Córdoba (Spain). *Renewable Energy*, 28(3): 409–418.
- R Development Core Team. (2012). *R: Un lenguaje y un entorno para computación estadística*. Viena: R Foundation for Statistical Computing. [Consulted: 17.xii.2024]. <<https://www.r-project.org/>>.
- Román Diez, R. (2010). *Efectos de los Componentes Atmosféricos sobre la Radiación Eritemática y Modelo propuesto para la reconstrucción de series*. (Tesis de Maestría). Valladolid: Universidad de Valladolid. 111 pp.
- SPSS Inc. (2002). *IBM SPSS Statistics* (Vers. 11.0). Chicago: SPSS Inc. [Consulted: 17.xii.2024]. <<https://www.ibm.com/es-es/products/spss-statistics>>.
- Varo Martínez, M.M. (2006). *Modelización de la Radiación Ultravioleta Solar*. (Tesis Doctoral). Córdoba: Universidad de Córdoba. 329 pp.
- Zaratti Sachetti, F. & Forno Gisbert, R. (Eds.). (2003). *La Radiación Ultravioleta en Bolivia*. La Paz: Organización Panamericana de la Salud / Organización Mundial de la Salud. 468 pp.