

MODELACIÓN ESTOCÁSTICA DE LLUVIAS HORARIAS

Ana Martínez¹ y José D. Salas²

Resumen:

En este artículo se presenta un modelo estocástico para analizar y simular lluvias horarias. El modelo que genera la ocurrencia de lluvias es un modelo periódico discreto autoregresivo de orden 1, denominado PDAR y aquél que genera las intensidades es uno estacionario autoregresivo (AR) de orden 1. Éste último requiere que las lluvias se transformen a una variable normal. La estimación de parámetros del modelo PDAR se hizo por el método de momentos mientras que para la estimación del modelo AR se utilizaron dos métodos: en el Método 1, las medias y desviaciones típicas horarias de la variable transformada (parámetros del modelo) se estimaron a fin de preservar las medias y las desviaciones típicas horarias obtenidas de la muestra histórica; en el Método 2, las medias y desviaciones típicas horarias de la variable transformada se estimaron directamente a partir de la muestra transformada. El modelo y métodos descritos anteriormente se probaron utilizando los datos horarios de lluvias del mes de Julio en una estación de Denver, Colorado del período 1949-1990. Además se analizaron los datos de 4, 6, 12, y 24 horas. La capacidad del modelo y métodos de estimación se probaron sobre la base de comparar estadísticas obtenidas del modelo ya sea directamente o por simulación con aquéllas obtenidas de la muestra histórica. Dichas estadísticas incluyeron medias y desviaciones típicas horarias y las probabilidades horarias de periodos secos (o aquéllas para la escala de tiempo considerada) así como la media y desviación típicas de las lluvias diarias. Los resultados obtenidos demuestran que el modelo PDAR-AR con el Método 1 de estimación es capaz de reproducir razonablemente bien las estadísticas históricas de las lluvias para varias escalas de tiempo.

Palabras clave: Lluvias horarias, modelos matemáticos, lluvias agregadas.

INTRODUCCIÓN

Los modelos matemáticos para representar la precipitación dependen de la escala de tiempo considerada. Para escalas de tiempo grandes, como anuales o mensuales, se puede asumir que la autocorrelación de la precipitación no existe o es despreciable (Roesner y Yevjevich, 1966). En este caso, basta con encontrar una distribución de probabilidad que sea capaz de reproducir la variabilidad de la precipitación en la escala de tiempo considerada.

Cuando la escala de tiempo es menor, o sea semanas, días u horas, la autocorrelación de la precipitación ya no es despreciable. En estos casos, se requiere de modelos más complejos que tengan en cuenta no sólo la periodicidad anual sino también la dependencia en el tiempo. Modelos estocásticos, como las cadenas de Markov, se han utilizado con

éxito para modelar la precipitación diaria (Katz, 1977a, b). Dichos modelos reproducen los estados de ocurrencia y no ocurrencia de la precipitación mediante una matriz de transición de probabilidad y la cantidad de lluvia, o intensidad, con una cierta función de probabilidad sesgada (generalmente lognormal o gamma).

Para escalas de tiempo menores que 24 horas, por ejemplo para lluvias horarias, también se han usado procesos Markovianos estacionarios, sin embargo los modelos más comunes han sido aquéllos basados en procesos puntuales (LeCam, 1961; Kavvas y Delleur, 1975) que modelan la precipitación en tiempo continuo, o los modelos que parten de una tormenta y desagregan la lluvia a una escala de tiempo menor mediante un hietograma (Huff, 1967).

¹ Centro de Estudios de Experimentación de Obras Públicas (CEDEX). Centro de Estudios Hidrográficos, Laboratorio de Hidráulica. Paseo Bajo Virgen del Puerto 3, 28005, Madrid, Spain

² Departamento de Ingeniería Civil. Colorado State University. Fort Collins, Colorado 80523, USA

Artículo recibido el 17 de junio de 2002, recibido en forma revisada el 20 de marzo de 2003 y aceptado para su publicación el 21 de agosto de 2003. Pueden ser remitidas discusiones sobre el artículo hasta seis meses después de la publicación del mismo siguiendo lo indicado en las "Instrucciones para autores". En el caso de ser aceptadas, éstas serán publicadas conjuntamente con la respuesta de los autores.

Sin embargo, Obeysekera y col. (1987) encontraron que las lluvias horarias convectivas tienen características periódicas no sólo relacionadas al ciclo anual causadas por la rotación de la tierra alrededor del sol, sino también a un ciclo diario causado por la rotación de la tierra alrededor de su eje. Por lo tanto, los modelos Markovianos y aquéllos basados en procesos puntuales adolecen de tales características, sobre todo aquéllas relacionadas al ciclo diario.

Katz y Parlange (1995) demostraron cómo los modelos Markovianos simples de dos estados pero con matrices de transición que varíen cíclicamente en el día son capaces de reproducir la precipitación horaria de una manera más realista que los modelos estacionarios. En este artículo se modela la precipitación horaria siguiendo un procedimiento similar al de Katz y Parlange (1995) pero usando un modelo matemático parametrizado en forma distinta y modificando la estimación de parámetros a fin de que el modelo sea capaz de reproducir no sólo estadísticas a escala horaria sino también a escala diaria.

BREVE REVISIÓN BIBLIOGRÁFICA

A lo largo de los años se han aplicado satisfactoriamente modelos estocásticos para representar series de lluvia diarias y horarias. Los modelos más utilizados se han basado en los procesos puntuales y las cadenas de Markov. Los primeros modelan la precipitación en tiempo continuo, asumiendo por ejemplo que la llegada de tormentas sigue un proceso de Poisson y que la cantidad de agua asociada a cada tormenta sigue una cierta distribución de probabilidad, como por ejemplo la gamma. En cambio, en los segundos, se asume que la ocurrencia de lluvias en el tiempo es de tipo Markoviano y luego se modela la intensidad. Estos modelos son más sencillos que los primeros. Gabriel y Neumann (1962) fueron los primeros en utilizar una cadena de Markov estacionaria de primer orden para modelar los estados de ocurrencia (lluvia o no lluvia) a escala diaria. Años más tarde Todorovic y Woolhiser (1975) utilizaron las cadenas de Markov como parte del modelo de lluvia diaria para representar los estados de ocurrencia, siendo la otra componente la cantidad de lluvia o intensidad, normalmente representada por la distribución gamma.

Desde entonces diversos autores (ver por ejemplo Katz, 1977; Roldán y Woolhiser, 1982;

Mimikou, 1983; Wilks, 1998; Sanso y Guenni, 1999) han seguido utilizando este tipo de modelos, permitiendo la periodicidad anual al dividir el año en varios periodos de tiempo (meses, estaciones), y estimando diferentes parámetros para cada periodo.

Sin embargo, la aplicación directa de las cadenas de Markov a series de lluvia horaria no fue tan satisfactoria, bien porque para simplificar el modelo las suposiciones que se hacían no eran realistas (como considerar las intensidades horarias independientes, Nguyen y Rousselle, 1981), o porque la estructura impuesta al modelo para permitir dicha dependencia le hacía demasiado complicado (Nguyen, 1984). En cambio, Katz y Parlange (1995) demostraron cómo la cadena de Markov periódica podía reproducir satisfactoriamente las estadísticas de la lluvia horaria, siendo a la vez consistentes con los modelos convencionales utilizados para la lluvia diaria. En su artículo mostraron la habilidad del modelo para reproducir los marcados ciclos diarios de la lluvia horaria, permitiendo a los parámetros variar en el día mediante una cadena de Markov periódica de primer orden, y autocorrelacionando las intensidades mediante un modelo autoregresivo de primer orden, o AR(1). También examinaron el comportamiento de este tipo de modelos cuando la lluvia horaria se agrega a lluvia diaria.

Islam y col. (1990) y Rodríguez-Iturbe y col. (1987, 1988) mostraron cómo los modelos basados en procesos puntuales tienden a subestimar la varianza de la lluvia diaria, así como que no pueden reproducir la frecuencia de días de lluvia. Además estos modelos puntuales han ignorado los ciclos diarios (incluso cuando se sabe que existen, como por ejemplo al final de la primavera y en verano en Denver), en parte porque su incorporación complicaría mucho, e incluso en algunos casos imposibilitaría, la estimación de los parámetros. Katz y Parlange (1995) encontraron fallos similares con la cadena de Markov periódica, sin embargo, demostraron su superioridad en cuanto a su capacidad para reproducir los marcados ciclos diarios de la precipitación en determinadas zonas y estaciones. Como mostraron en su artículo, para reproducir adecuadamente estadísticas como la duración de periodos secos o de lluvia, es necesario poder tratar explícitamente estos ciclos diarios. Con relación a la subestimación de la varianza cuando se agrega la lluvia horaria a lluvia diaria, Katz y Parlange (1995) comprobaron que autocorrelacionar o no las intensidades,

no contribuye a una mejora significativa en la estimación de la varianza de la precipitación diaria. Alternativamente, proponen probar modelos capaces de aumentar el grado de persistencia en los estados de ocurrencia de la lluvia horaria, como medida para mejorar las estadísticas del proceso agregado.

En este artículo se sigue un procedimiento similar al de Katz y Parlange (1995) pero con una modelación del proceso de ocurrencia de lluvias parametrizado de forma distinta. Además se encuentra un método de estimación de parámetros más eficiente en términos de reproducir estadísticas a escala diaria.

MODELO MATEMÁTICO DE LA LLUVIA HORARIA

Como se dijo anteriormente, la precipitación horaria se puede asumir como la combinación de dos procesos: (1) aquél que determina la ocurrencia de la precipitación (o sea la secuencia de horas de lluvia y horas secas) y (2) el de la intensidad (cantidad de agua en las horas en que llueve). Además se asume que ambos procesos son mutuamente independientes.

En este artículo, por razones de simplicidad y compacidad, el modelo de la precipitación horaria se expresa igual que el modelo producto introducido por Chebaane y col. (1995). La ecuación básica de dicho modelo es

$$Y_{v,\tau} = X_{v,\tau} Z_{v,\tau}; \quad v=1,2,\dots; \quad \tau=1,2,\dots,\omega \quad (1)$$

donde $\{Y_{v,\tau}\}$ es una variable positiva intermitente que representa la precipitación horaria, v representa los días, τ la escala de tiempo dentro del día (por ejemplo horas), ω es el número de periodos en el día, es decir, $\omega=24$ para lluvias horarias, $\{X_{v,\tau}\}$ es una variable periódica binaria (0,1) que representa el proceso de ocurrencia, y $\{Z_{v,\tau}\}$ es una variable periódica continua que representa la intensidad. Los procesos $\{X_{v,\tau}\}$ y $\{Z_{v,\tau}\}$ se suponen independientes. Esta suposición puede no ser válida pero el no suponerla conduciría a una modelación mucho más compleja.

Alternativamente, la ecuación (1) se puede reescribir como

$$Y_{v,\tau} = \begin{cases} Z_{v,\tau} & \text{si } X_{v,\tau} = 1 \\ 0 & \text{si } X_{v,\tau} = 0 \end{cases} \quad (2)$$

Además se asume que este modelo es válido durante una estación del año (periodo de tiempo dentro del año que puede ser de uno o más meses). Entonces para la descripción completa del modelo será necesario modelar cada componente, es decir, $X_{v,\tau}$ y $Z_{v,\tau}$. Éstas se describen en las siguientes secciones.

MODELACIÓN DE LA OCURRENCIA DE LLUVIAS HORARIAS, $\{X_{v,\tau}\}$

El proceso que define la ocurrencia de las lluvias horarias se representa por

$$\begin{cases} X_{v,\tau} = 1 & \text{si en la hora } \tau \text{ del día } v \text{ ha llovido.} \\ X_{v,\tau} = 0 & \text{de lo contrario} \end{cases} \quad (3)$$

Un modelo periódico discreto autoregresivo de primer orden, o PDAR(1), se define como (Chebaane y col. 1995)

$$X_{v,\tau} = V_{v,\tau} X_{v,\tau-1} + (1 - V_{v,\tau}) W_{v,\tau}; \quad (4)$$

$$v=1,2,\dots; \quad \tau=1,2,\dots,\omega$$

donde $\{X_{v,\tau}\}$ es un proceso binario periódico dependiente que toma valores $\{0,1\}$, y con la convención $X_{v,0} = X_{v-1,\omega}$. Además $\{V_{v,\tau}\}$ y $\{W_{v,\tau}\}$ son procesos Bernoulli independientes con probabilidades

$$P\{W_{v,\tau} = 1\} = \pi_{1,\tau}; \quad 0 \leq \pi_{1,\tau} \leq 1 \quad (5)$$

$$P\{W_{v,\tau} = 0\} = \pi_{0,\tau} = 1 - \pi_{1,\tau} \quad (6)$$

y

$$P\{V_{v,\tau} = 1\} = \lambda_{\tau}; \quad 0 \leq \lambda_{\tau} \leq 1 \quad (7)$$

Por lo tanto, el proceso $\{X_{v,\tau}\}$ también se puede representar de la siguiente manera

$$X_{v,\tau} = \begin{cases} X_{v,\tau-1} & \text{con probabilidad } \lambda_{\tau} \\ W_{v,\tau} & \text{con probabilidad } 1 - \lambda_{\tau} \end{cases} \quad (8)$$

Al mismo tiempo, se puede demostrar que el proceso PDAR(1) es equivalente a una cadena de Markov periódica con matriz de transición P_{τ} cuyos elementos están definidos por

$$p_{\tau}(i,j) = P\{X_{v,\tau} = j | X_{v,\tau-1} = i\}; \quad (9)$$

$$i, j = 0, 1$$

la cual depende de la hora τ , permitiendo por lo tanto tener en cuenta el ciclo diario.

Combinando las ecuaciones (5) - (9) la matriz de transición se puede expresar en función de los parámetros del modelo (Chebaane y col. 1995)

$$P_{\tau} = \begin{bmatrix} 0 & 1 \\ \lambda_{\tau} + (1 - \lambda_{\tau})\pi_{0,\tau} & (1 - \lambda_{\tau})\pi_{1,\tau} \\ (1 - \lambda_{\tau})\pi_{0,\tau} & \lambda_{\tau} + (1 - \lambda_{\tau})\pi_{1,\tau} \end{bmatrix} \quad (10)$$

La probabilidad de lluvia en la hora τ , o sea $P[X_{v,\tau} = 1] = E[X_{v,\tau}] = \mu_{\tau}(X)$ se puede determinar teniendo en cuenta la ecuación (8) como

$$P[X_{v,\tau} = 1] = \mu_{\tau}(X) = \lambda_{\tau}\mu_{\tau-1}(X) + (1 - \lambda_{\tau})\pi_{1,\tau}; \quad \tau = 1, \dots, \omega \quad (11)$$

El sistema de ω ecuaciones se puede resolver para $\mu_{\tau}(X)$ en función de los parámetros del modelo λ_{τ} y $\pi_{1,\tau}$ (Chung, 1999)

$$\mu_{\tau}(X) = \frac{\sum_{i=1}^{\tau} [\pi_{1,i}(1 - \lambda_i) \prod_{t=i+1}^{\tau} \lambda_t] + \sum_{i=\tau+1}^{\omega} [\pi_{1,i}(1 - \lambda_i) \frac{\prod_{t=1}^{\omega} \lambda_t}{\prod_{t=\tau+1}^{\omega} \lambda_t}]}{1 - \prod_{t=1}^{\omega} \lambda_t} \quad (12)$$

Así mismo, la duración de periodos secos o lluviosos (secuencia de horas secas o de lluvia, respectivamente) refleja la presencia del ciclo diario en las probabilidades de transición. Definiendo $T_{i,\tau}$ como la duración del periodo seco ($i=0$) o húmedo ($i=1$) originado en la hora τ , la probabilidad de que uno de los periodos, por ejemplo el seco, originado en la hora τ dure por lo menos n horas, se puede obtener fácilmente mediante

$$P[T_{0,\tau} \geq n] = P[X_{v,\tau+n-1} = 0, \dots, X_{v,\tau} = 0, \dots, X_{v,\tau-1} = 1 | X_{v,\tau} = 0, X_{v,\tau-1} = 1] = \prod_{j=0}^{n-1} p_{\tau+j}(0,0) \quad n \geq 2$$

$$= 1 \quad n = 1 \quad (13)$$

Para estimar los parámetros del modelo, es decir λ_{τ} y $\pi_{0,\tau}$, primeramente se calculan las probabilidades de transición y luego se utiliza la ecuación (10).

Las probabilidades de transición se estiman de la muestra mediante

$$\hat{p}_{\tau}(i, j) = n_{\tau}(i, j) / n_{\tau}(i); \quad i, j = 0, 1; \quad \tau = 1, \dots, \omega \quad (14)$$

donde $n_{\tau}(i, j)$ representa el número de veces que la lluvia pasa del 'estado i ' en la hora $\tau - 1$ al 'estado j ' en la hora τ , y $n_{\tau}(i) = n_{\tau}(i, 0) + n_{\tau}(i, 1)$ representa el número de veces que en la hora τ la lluvia está en el 'estado i '. Dado que generalmente los estimados $\hat{p}_{\tau}(i, j)$ muestran cierta variabilidad o fluctuación además de una variación periódica, uno podría utilizar las series de Fourier para eliminar dichas fluctuaciones y al mismo tiempo mantener la periodicidad. Por ejemplo, Katz y Parlange (1995) siguieron este procedimiento. En este caso la representación paramétrica de $\hat{p}_{\tau}(i, j)$ usando series de Fourier es (Salas y col. 1980)

$$p_{\tau}(i, 1) = \bar{p}(i, 1) + \sum_{m=1}^h [A_m^i \cos(2\pi m \tau / \omega) + B_m^i \sin(2\pi m \tau / \omega)] \quad i = 0, 1; \quad \tau = 1, \dots, \omega \quad (15)$$

donde $\bar{p}(i, 1)$ es la media de $\hat{p}_{\tau}(i, 1)$, A_m^i y B_m^i son los coeficientes de la serie de Fourier, m es el armónico y h es el número total de armónicos, que es igual a $\omega/2$ cuando ω es par y $(\omega-1)/2$ cuando es impar. Notar que en la práctica generalmente se utiliza un número menor de armónicos $h^* < h$.

La media, $\bar{p}(i, 1)$ y los coeficientes de Fourier, A_m^i , B_m^i , se calculan mediante

$$\bar{p}(i, 1) = \sum_{\tau=1}^{\omega} \hat{p}_{\tau}(i, 1); \quad i = 0, 1 \quad (16)$$

$$\hat{A}_m^i = (2/\omega) \sum_{\tau=1}^{\omega} \hat{p}_{\tau}(i, 1) \cos(2\pi m \tau / \omega); \quad m = 1, \dots, h; \quad i = 0, 1 \quad (17)$$

$$\hat{B}_m^i = (2/\omega) \sum_{\tau=1}^{\omega} \hat{p}_{\tau}(i, 1) \sin(2\pi m \tau / \omega); \quad m = 1, \dots, h; \quad i = 0, 1 \quad (18)$$

Cuando ω es par, los últimos coeficientes vienen dados por

$$\hat{A}_h^i = (1/\omega) \sum_{\tau=1}^{\omega} \hat{p}_{\tau}(i, 1) \cos(2\pi h \tau / \omega); \quad i = 0, 1 \quad (19)$$

$$\hat{B}_h^i = 0; \quad i = 0, 1 \quad (20)$$

Una vez conocidos $\bar{p}(i, l)$, y , los estimados de las probabilidades de transición se obtienen mediante la ecuación (15).

Finalmente, teniendo en cuenta la ecuación (10) y usando las probabilidades de transición obtenidas de la ecuación (15) los parámetros λ_τ y $\pi_{1,\tau}$ del modelo PDAR(1) vienen dados por

$$\lambda_\tau = p_\tau(0,0) - p_\tau(1,0) = p_\tau(1,1) - p_\tau(0,1) \quad (21)$$

$$\pi_{1,\tau} = \frac{p_\tau(0,1)}{1 - \lambda_\tau} \quad (22)$$

y el parámetro $\pi_{0,\tau}$ se determina por la ecuación (6).

MODELACIÓN DE LA INTENSIDAD DE LLUVIAS, $\{Z_{v,\tau}\}$

El proceso $Z_{v,\tau}$ representa la cantidad de lluvia en las horas que llueve, es decir, si $X_{v,\tau}=1$ entonces $Z_{v,\tau} > 0$. La variable $Z_{v,\tau}$ es generalmente sesgada y autocorrelacionada. Primero examinaremos algunas de sus propiedades y luego una forma de transformarla a una variable normal a fin de que sea más fácil el modelar su dependencia en el tiempo.

Las medias y varianzas de las intensidades horarias se definen como

$$\mu_\tau(Z) = E[Z_{v,\tau} | X_{v,\tau} = 1] \quad (23)$$

$$\sigma_\tau^2(Z) = Var[Z_{v,\tau} | X_{v,\tau} = 1] \quad (24)$$

y al igual que en el proceso de ocurrencia, las medias y varianzas no dependen del día. Debido a la distribución sesgada de las intensidades conviene normalizarlas, elevándolas a una determinada potencia (Katz y Garrido, 1994). Por ejemplo, dado que $X_{v,\tau} = 1$,

$$N_{v,\tau} = Z_{v,\tau}^{1/c} \quad \text{para algún } c, c > 1 \quad (25)$$

donde la variable aleatoria $N_{v,\tau}$ sigue una distribución normal con media $\mu_\tau(N)$ y varianza $\sigma_\tau^2(N)$. Por lo tanto, la media y la varianza de $Z_{v,\tau}$, $\mu_\tau(Z)$ y $\sigma_\tau^2(Z)$, se pueden expresar en función de $\mu_\tau(N)$, $\sigma_\tau^2(N)$, y c . El parámetro c se supone independiente de las horas y toma el valor que minimice el índice de simetría de Hinkley (Katz y Parlange, 1993), o dicho de otra manera, aquél que minimice el sesgo. Para series de precipitación los valores de c que normalmente se tantean son 2, 3, 4, 5, 8 y 10.

Para nuestra serie de precipitación horaria el menor índice se obtuvo para $c = 8$.

Para relacionar la distribución de la precipitación horaria $Z_{v,\tau}$ con la de la variable transformada $N_{v,\tau}$, se lleva a cabo un cambio de variable (Lindgren, 1968, p. 70). La variable estandarizada $Q_{v,\tau} = [N_{v,\tau} - \mu_\tau(N)] / \sigma_\tau(N)$ sigue una distribución normal con media cero y varianza uno y sus momentos se pueden expresar como

$$\mu_{r,\tau} = E[Q_{v,\tau}^r] = \begin{cases} 0; & \text{si } r \text{ es impar} \\ (r-1) \times (r-3) \times \dots \times 3 \times 1; & \text{si } r \text{ es par} \end{cases} \quad (26)$$

Entonces las relaciones entre los momentos de las variables $Z_{v,\tau}$ y $N_{v,\tau}$ se pueden obtener de

$$E[Z_{v,\tau}^r] = E[N_{v,\tau}^{rc}] = E\{[\mu_\tau(N) + \sigma_\tau(N)Q_{v,\tau}]^{rc}\} \quad (27)$$

Por ejemplo, para el caso $c = 8$, las medias y varianzas de la precipitación horaria en función de aquellas de la variable transformada son:

$$\begin{aligned} \mu_\tau(Z) &= \\ &= \mu_\tau^8(N) + 105\sigma_\tau^8(N) + 28\mu_\tau^6(N)\sigma_\tau^2(N) + \\ &+ 420\mu_\tau^2(N)\sigma_\tau^6(N) + 210\mu_\tau^4(N)\sigma_\tau^4(N) \end{aligned} \quad (28a)$$

$$\begin{aligned} \sigma_\tau^2(Z) &= \\ &= 2016000\sigma_\tau^{16}(N) + 64\mu_\tau^{14}(N)\sigma_\tau^2(N) + \\ &+ 4256\mu_\tau^{12}(N)\sigma_\tau^4(N) + 107520\mu_\tau^{10}(N)\sigma_\tau^6(N) + \\ &+ 1283520\mu_\tau^8(N)\sigma_\tau^8(N) + 7385280\mu_\tau^6(N)\sigma_\tau^{10}(N) + \\ &+ 18698400\mu_\tau^4(N)\sigma_\tau^{12}(N) + 16128000\mu_\tau^2(N)\sigma_\tau^{14}(N) \end{aligned} \quad (28b)$$

A escala horaria es importante tener en cuenta la dependencia en el tiempo entre las intensidades. Para ello la variable estandarizada $Q_{v,\tau}$ se modela utilizando el modelo autoregresivo de primer orden, AR(1). En este artículo se presentan directamente las ecuaciones, sin entrar en cómo se obtienen. Para una descripción más detallada el lector puede referirse por ejemplo a Salas (1993). Entonces el modelo AR(1) escrito para las variables $N_{v,\tau}$ y $Q_{v,\tau}$ están dadas por

$$\begin{cases} N_{v,\tau} = \mu_\tau(N) + \sigma_\tau(N)Q_{v,\tau} \\ Q_{v,\tau} = \phi_l Q_{v,\tau-1} + \varepsilon_{v,\tau} \end{cases} \quad (29)$$

donde $\varepsilon_{v,\tau}$ es una serie independiente que sigue una distribución normal $N[0, 1 - \phi_1^2]$ y ϕ_1 es el coeficiente de autoregresión.

Los parámetros del modelo son $\mu_\tau(N)$, $\sigma_\tau^2(N)$ y ϕ_1 . La media y la varianza de las intensidades transformadas, $\mu_\tau(N)$, $\sigma_\tau^2(N)$ se estiman resolviendo el sistema de ecuaciones no lineales, ecuaciones (28a) y (28b), reemplazando $\mu_\tau(Z)$ y $\sigma_\tau^2(Z)$, la media y la varianza de las intensidades de lluvia horaria por sus estimados a partir de la muestra histórica. Estas estadísticas varían con la hora y generalmente fluctúan alrededor de una variación periódica. Entonces como se hizo con las probabilidades de transición, se pueden utilizar las series de Fourier para representar dichas variaciones periódicas y eliminar el ruido. Así las ecuaciones toman la misma forma que aquellas numeradas de la (15) a la (20), con los cambios correspondientes. Este método de estimación de $\mu_\tau(N)$ y $\sigma_\tau^2(N)$ a partir de las estadísticas de la muestra $\mu_\tau(Z)$ y $\sigma_\tau^2(Z)$ y ecuaciones (28) se refiere como Método 1 de estimación.

En cambio, Katz y Parlange (1995) siguieron el proceso inverso, es decir, obtuvieron la media $\hat{\mu}_\tau(N)$ y la varianza $\hat{\sigma}_\tau^2(N)$ a partir de las intensidades transformadas (ecuación 25), luego ajustaron las series de Fourier a dichas estadísticas, para obtener $\mu_\tau(N)$ y $\sigma_\tau^2(N)$ y a manera de comprobación utilizaron las ecuaciones (28) para determinar las estadísticas $\mu_\tau(Z)$ y $\sigma_\tau^2(Z)$. Este método alternativo de estimación de $\mu_\tau(N)$ y $\sigma_\tau^2(N)$ a partir de la muestra transformada se refiere como Método 2 de estimación. Más adelante, cuando se calculan las estadísticas del proceso agregado, se muestra la ventaja de seguir los pasos descritos en este artículo (Método 1) al comparar los resultados con los obtenidos por los citados autores (Método 2).

Por último, el coeficiente de autoregresión, ϕ_1 se estima mediante la función de autocorrelación del modelo (conocida como la ecuación de Yule-Walker), $\rho_k = \phi_1 \rho_{k-1} = \phi_1^k$; $k \geq 0$. Aplicando el Método de los Momentos y particularizando para $k = 1$ tenemos

$$\hat{\phi}_1 = r_1 \tag{30}$$

donde r_1 es el coeficiente de autocorrelación de retardo 1 que se estima de la muestra (Salas, 1993). Luego, la varianza de la serie independiente $\{\varepsilon_{v,\tau}\}$ se estima por

$$\hat{\sigma}_\varepsilon^2 = (1 - \hat{\phi}_1^2) \tag{31}$$

RESULTADOS

Los modelos se aplicaron a una serie de precipitación horaria de 42 años, correspondiente al periodo 1949-1990, del mes de Julio en la estación del Aeropuerto de Denver, Colorado. En esta zona, la precipitación durante el verano es normalmente de origen convectivo, por lo que se puede anticipar ciclos diarios en las estadísticas horarias.

Lluvias horarias

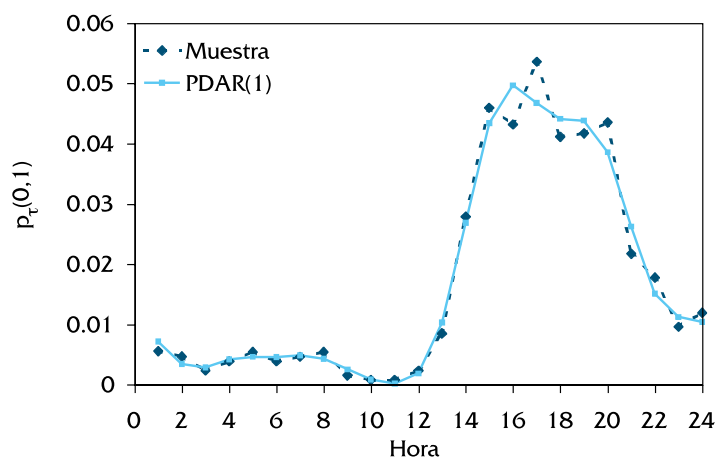
Como se muestra a continuación el modelo PDAR(1) reproduce satisfactoriamente los estados de ocurrencia de la lluvia horaria. Una vez estimadas de la muestra las probabilidades de transición $p_\tau(i,j)$, éstas se ajustaron mediante las series de Fourier. La tabla 1 da los parámetros (de las series de Fourier) con 6 armónicos para $p_\tau(0,1)$ y 2 armónicos para $p_\tau(1,1)$. Luego los parámetros del modelo PDAR(1) estimados por las ecuaciones (21) y (22) se resumen en la tabla 2.

Tabla 1. Parámetros de las series de Fourier para las probabilidades de transición del proceso de los estados de ocurrencia de lluvias horarias.

Probabilidad de transición $p_\tau(0,1)$: media $\bar{p}(0,1) = 0.00170$						
Armónicos: m	1	2	3	4	5	6
\hat{A}_m^0	-0.0063	-0.0071	0.0025	-0.0004	0.0022	-0.0007
\hat{B}_m^0	-0.0203	0.0083	-0.0021	0.0018	-0.0013	-0.0012
Probabilidad de transición $p_\tau(1,1)$: media $\bar{p}(1,1) = 0.4931$						
Armónicos: m	1	2				
\hat{A}_m^0	0.0718	-0.0471				
\hat{B}_m^0	0.0326	-0.0122				

Tabla 2. Parámetros del modelo PDAR(1)-AR(1) para la serie de lluvias horarias.

Tiempo τ (horas)	Parámetros				Tiempo τ (horas)	Parámetros			
	λ_τ	$\pi_{1,\tau}$	$\mu_\tau(N)$	$\sigma_\tau(N)$		λ_τ	$\pi_{1,\tau}$	$\mu_\tau(N)$	$\sigma_\tau(N)$
1	0.5106	0.0147	0.6338	0.0907	13	0.3640	0.0162	0.6674	0.1000
2	0.5205	0.0072	0.6248	0.0832	14	0.3416	0.0408	0.6684	0.1106
3	0.5346	0.0061	0.6158	0.0800	15	0.3372	0.0654	0.6685	0.1201
4	0.5505	0.0093	0.6089	0.0807	16	0.3575	0.0773	0.6681	0.1276
5	0.5656	0.0106	0.6065	0.0827	17	0.3952	0.0773	0.6673	0.1327
6	0.5733	0.0107	0.6098	0.0834	18	0.4337	0.0779	0.6662	0.1352
7	0.5679	0.0112	0.6182	0.0817	19	0.4638	0.0817	0.6644	0.1351
8	0.5486	0.0096	0.6294	0.0785	20	0.4885	0.0754	0.6619	0.1324
9	0.5170	0.0052	0.6412	0.0761	21	0.5087	0.0534	0.6585	0.1274
10	0.4768	0.0016	0.6516	0.0767	22	0.5180	0.0313	0.6541	0.1200
11	0.4341	0.0003	0.6595	0.0813	23	0.5148	0.0231	0.6485	0.1108
12	0.3957	0.0031	0.6647	0.0896	24	0.5089	0.0212	0.6418	0.1006
$\phi_1 = 0.327$									



En la Figura 1(a) se muestra la probabilidad de transición $p_\tau(0,1)$ obtenida de la muestra (ecuación 14), así como la del modelo PDAR(1) (ecuación 10). Es evidente que existe un ciclo diario en las probabilidades, con una amplitud de dos órdenes de magnitud aproximadamente, 0.0002 a las 10 hasta 0.049 a las 15 horas. La amplitud de la probabilidad de transición, $p_\tau(1,1)$, es menor (no se muestra), variando entre 0.368 a la 13 y 0.579 a las 5 horas. Además la Figura 1(b) da la comparación de la probabilidad de lluvia en cada hora, $\mu_\tau(X)$, estimada de la muestra histórica y aquella del modelo PDAR(1) obtenida de la ecuación (12). Aquí nuevamente se comprueba la existencia del ciclo diario con un mínimo de 0.002 a las 10 y un máximo de 0.080 a las 19 horas.

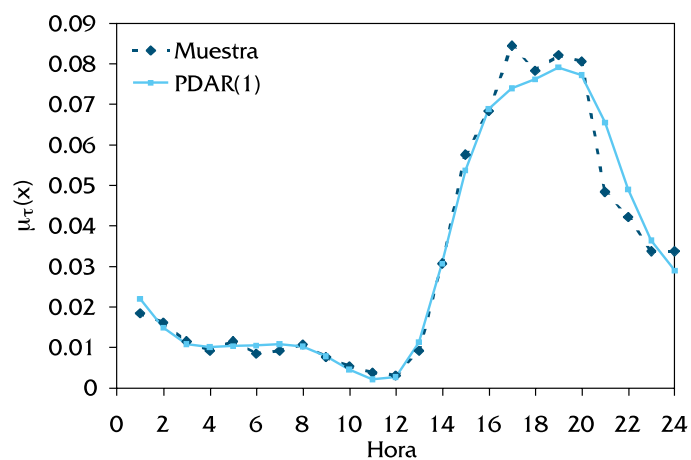


Figura 1. Probabilidades observadas y modeladas del proceso de ocurrencia de las lluvias horarias: (a) probabilidad de transición, $p_\tau(0,1)$ y (b) probabilidad de lluvia cada hora, $\mu_\tau(X)$.

La Figura 2 da la probabilidad $P(T_{0,\tau} \geq n)$ de que un periodo seco originado en la hora τ dure por lo menos n horas, tanto para el modelo PDAR(1) (ecuación 13) así como para la muestra. Se puede ver como el ciclo diario en las probabilidades (por ejemplo ver Figura 1(a) para $p_\tau(0,1)$) produce variaciones en $P(T_{0,\tau} \geq n)$ que dependen de la hora τ en la que el periodo seco comienza. Para periodos secos cortos, por ejemplo de 3 y 6 horas, el modelo sobreestima aquellas derivadas de la muestra especialmente en las horas del día donde la intensidad es mayor (después del medio día). A medida que la duración del periodo seco aumenta, por ejemplo a 9 y 15 horas, vemos que el modelo reproduce mejor los periodos observados.

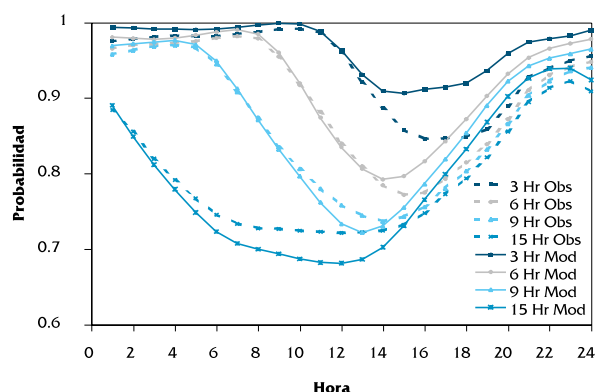


Figura 2. Variación de la probabilidad de periodos secos en función de la hora en la que comienzan para el modelo PDAR(1) y para la muestra observada.

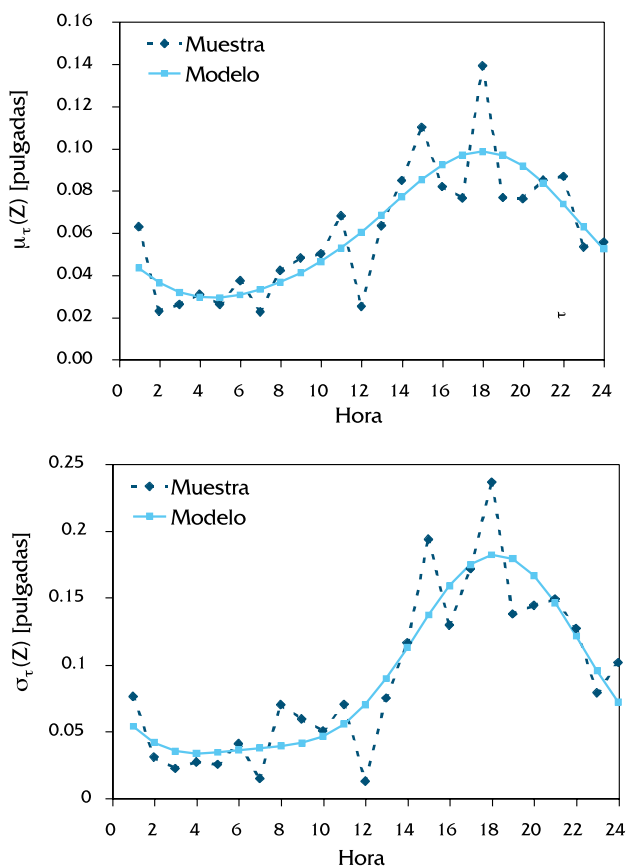


Figura 3. Estadísticas de lluvias horarias $Z_{v,\tau}$: (a) media $\mu_\tau(Z)$ y (b) desviación típica $\sigma_\tau(Z)$. La línea quebrada corresponde a las estimadas de la muestra y la continua a aquellas ajustadas con las series de Fourier.

Con respecto a la modelación de la intensidad, como ya se adelantó en la introducción, sólo se muestran los resultados de aplicar el modelo AR(1) ya que los otros dos modelos que se probaron, AR(2) y ARMA(1,1), no mejoran los resultados del primero.

Las curvas de las Figuras 3(a) y 3(b) representan la media, $\mu_\tau(Z)$, y la desviación típica horaria, $\sigma_\tau(Z)$, de la muestra junto con aquellas ajustadas mediante las series de Fourier con dos armónicos, respectivamente. Una vez conocidas $\mu_\tau(Z)$ y $\sigma_\tau^2(Z)$, las estadísticas de las intensidades transformadas, $\mu_\tau(N)$ y $\sigma_\tau(N)$ se obtienen al resolver las ecuaciones (28) y se muestran en las Figuras 4(a) y 4(b), respectivamente.

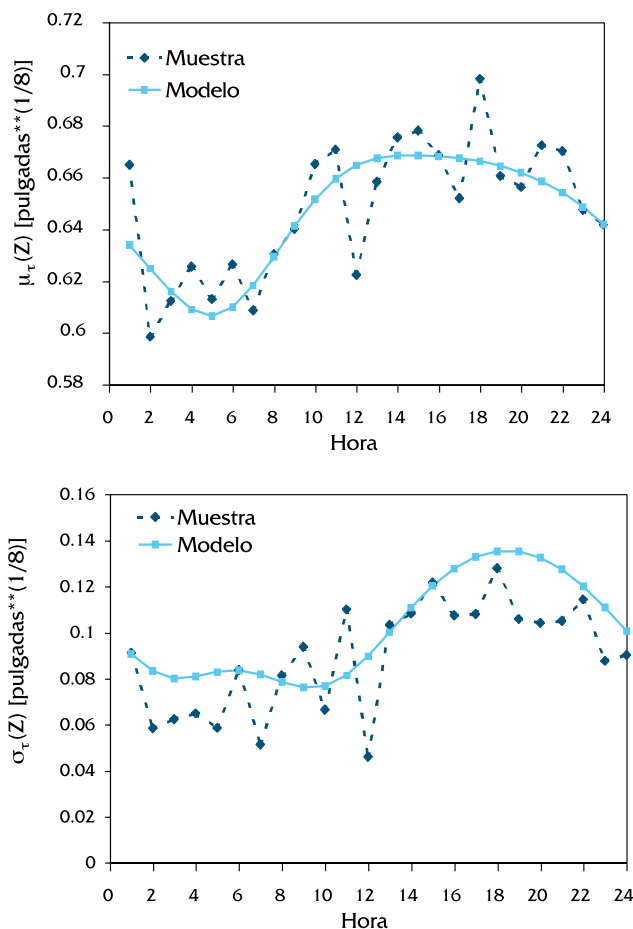


Figura 4. Estadísticas horarias de las intensidades de lluvias transformadas: (a) media $\mu_\tau(N)$ y (b) desviación típica $\sigma_\tau(N)$.

Como se puede observar en la Figura 4(a) la media de la muestra $\mu_\tau(N)$ se reproduce adecuadamente. Para la desviación típica de las intensidades transformadas, si bien es cierto que la tendencia cíclica es reproducida, hay una ligera sobreestimación. Esto ocurre por el efecto de la transformación de las intensidades de lluvias. Sin embargo, como veremos más adelante tal sobreestimación más bien tiene un efecto positivo en la reproducción de estadísticas a escalas de tiempo agregadas.

Lluvias Agregadas

Otra manera de sancionar cualquier modelo estocástico para la precipitación horaria es comprobar su capacidad de reproducir las estadísticas para escalas de tiempo más grandes, como por ejemplo días. Aún más, para comprobar la aplicabilidad del modelo PDAR(1)-AR(1) a otras escalas de tiempo, a partir de los datos de lluvias horarias ya descritos anteriormente, se obtuvieron datos de lluvias a escalas de 4 hr., 6 hr., 12hr., y 24 hr., y luego se modelaron siguiendo el mismo procedimiento descrito antes en detalle para las lluvias horarias. Los parámetros del modelo PDAR(1)-AR(1) obtenidos en cada caso se dan en la tabla 3. Nótese que para el caso de datos de escala de 24 horas (datos diarios) el modelo es más bien estacionario pero con la misma estructura, es decir, DAR(1)-AR(1).

Para verificar el comportamiento del modelo y métodos de estimación referidos para lluvias horarias y lluvias de escalas de tiempo de 4, 6, 12 y hasta 24 horas (diarias), se simularon en cada caso 100 series de 4000 días y se estimaron a partir de ellas la media y la desviación típica de las lluvias diarias. Éstas se compararon respectivamente con aquéllas estimadas de la muestra histórica. Los resultados se muestran en la tabla 4. Se puede notar que en general la media diaria es reproducida satisfactoriamente con ambos métodos de estimación. Sin embargo éste no es el caso para la desviación típica. Por ejemplo, para la escala de 1 hora (lluvias horarias), el Método 2 subestima significativamente la desviación típica de las lluvias diarias. En cambio, con el Método 1 la subestimación es menor o despreciable. Además como era de esperar, se puede notar que la referida subestimación disminuye cuando la escala de tiempo se incrementa o se aproxima a la escala diaria.

COMENTARIOS FINALES

En este artículo se ha examinado la capacidad de un modelo estocástico para representar series de precipitación horaria. El modelo tiene en cuenta el hecho de que en algunas épocas del año las lluvias varían en forma cíclica a lo largo del día, es decir que la ocurrencia de la lluvia así como su intensidad dependen de la hora. Por ejemplo, las lluvias convectivas generalmente ocurren después de medio día. Así, el modelo de lluvias se ha conceptualizado como el producto de dos procesos componentes, el primero que genera la ocurrencia

de lluvias dentro del día (o sea si hay o no hay lluvia) y el segundo que genera la intensidad cuando llueve.

El modelo que genera la ocurrencia de lluvias es un modelo periódico discreto autoregresivo de orden 1, denominado PDAR y aquel que genera las intensidades es un modelo estacionario autoregresivo (AR) de orden 1. Este último requiere que las intensidades de lluvia se transformen previamente a una variable normal. La estimación de parámetros del modelo PDAR se hizo por el método de momentos haciendo previamente un ajuste de las matrices de transición de probabilidades horarias mediante las series de Fourier. En cuanto a la estimación de parámetros del modelo AR se utilizaron dos métodos. En el primero (Método 1), las medias y desviaciones típicas horarias de la variable transformada (parámetros del modelo) se estimaron a fin de preservar directamente las medias y las desviaciones típicas horarias obtenidas de la muestra histórica. En el segundo (Método 2), las medias y desviaciones típicas horarias de la variable transformada se estimaron directamente a partir de la muestra transformada.

El modelo y métodos descritos anteriormente se probaron utilizando los datos horarios de lluvias del mes de Julio de la estación del aeropuerto de Denver, Colorado, en el periodo 1949-1990. Además de analizar los datos horarios, el modelo se aplicó para datos de 4, 6, 12, y 24 horas. La capacidad del modelo y métodos de estimación se probaron sobre la base de comparar estadísticas obtenidas del modelo ya sea directamente o por simulación con aquellas obtenidas de la muestra histórica. Dichas estadísticas incluyeron medias y desviaciones típicas horarias y las probabilidades horarias de periodos secos (o aquellas para la escala de tiempo considerada) así como la media y desviación típica de las lluvias diarias. Los resultados obtenidos demuestran que el modelo PDAR-AR con el Método 1 de estimación es capaz de reproducir razonablemente bien las estadísticas históricas de las lluvias para varias escalas de tiempo.

Por último, cabe hacer notar que el modelo utilizado en este artículo podría modificarse utilizando modelos que sean capaces de reproducir una mayor persistencia tanto en la ocurrencia de las lluvias así como en las intensidades. Por ejemplo, utilizando el modelo PDARMA para la ocurrencia y modelos AR(p) para las intensidades. Sin embargo, en el caso de las lluvias de Julio en Denver, el modelo utilizado dio un buen resultado.

Tabla 3. Parámetros del modelo PDAR(1)-AR(1) para lluvias agregadas a escalas de tiempo de 4, 6, 12 y 24 horas.

AR(1)	ESCALAS DE TIEMPO DE LLUVIAS AGREGADAS															
	4 horas				6 horas				12 horas				24 horas			
	$\phi_1 = 0.300$				$\phi_1 = 0.269$				$\phi_1 = 0.116$				$\phi_1 = 0.079$			
$\tau \backslash$ PDAR(1)	λ_τ	$\pi_{1,\tau}$	$\mu_\tau(N)$	$\sigma_\tau(N)$	λ_τ	$\pi_{1,\tau}$	$\mu_\tau(N)$	$\sigma_\tau(N)$	λ_τ	$\pi_{1,\tau}$	$\mu_\tau(N)$	$\sigma_\tau(N)$	λ	π	$\mu(N)$	$\sigma(N)$
1	0.213	0.015	0.672	0.101	0.135	0.015	0.687	0.097	0.096	0.021	0.688	0.113	0.213	0.298	0.751	0.128
2	0.382	0.021	0.633	0.111	0.211	0.014	0.674	0.116	0.175	0.322	0.751	0.130				
3	0.319	0.006	0.698	0.088	0.188	0.230	0.740	0.125								
4	0.335	0.158	0.715	0.117	0.309	0.148	0.708	0.134								
5	0.335	0.213	0.720	0.136												
6	0.206	0.057	0.708	0.131												

Tabla 4. Estadísticas de lluvias diarias estimadas de la muestra histórica y de aquéllas que resultan de modelar y simular lluvias a escalas de tiempo de 1, 4, 6, 12 y 24 horas.

Método 1 de estimación	Muestra histórica	1 hora	4 horas	6 horas	12 horas	1 día
Media diaria (pulgadas)	0.0607	0.0619	0.0648	0.0615	0.0588	0.0598
Desviación típica diaria (pulgadas)	0.1924	0.1785	0.1923	0.1886	0.1906	0.1980
Método 2 de estimación	Muestra histórica	1 hora	4 horas	6 horas	12 horas	1 día
Media diaria (pulgadas)	0.0607	0.0527	0.0602	0.0638	0.0555	0.0570
Desviación típica diaria (pulgadas)	0.1924	0.1452	0.1775	0.1865	0.1906	0.1890

AGRADECIMIENTOS.

Se agradece la financiación de los proyectos “Uncertainty and Risk Analysis Under Extreme Hydrologic Events” de la Fundación Nacional de Ciencias (NSF) de los Estados Unidos de América y “Predictability of Extreme Hydrologic Events Related to Colorado’s Agriculture” de la Estación Experimental de Agricultura para Colorado, que hicieron posible la investigación materia de este artículo. También se agradece a R. Kaz y M. Parlange por habernos proporcionado los datos de lluvias horarias.

LISTA DE SÍMBOLOS

- $\{Y_{v,\tau}\}$: serie de la precipitación horaria.
- $\{X_{v,\tau}\}$: serie del proceso de estados de ocurrencia.
- $\{Z_{v,\tau}\}$: serie de las intensidades.
- v : días.
- τ : horas.
- $p_\tau(i,j)$: matriz de transición de probabilidad para el proceso de los estados de ocurrencia.
- $\mu_\tau(X)$: probabilidad de lluvia en la hora τ .
- $T_{i,\tau}$: duración en horas de los periodos secos ($i=0$) o de lluvia ($i = 1$) en función de la hora τ en la que empieza.

- A_m y B_m : coeficientes de la serie de Fourier utilizados para ajustar estadísticas ya sea de $X_{v,\tau}$ o de $Z_{v,\tau}$.
- $\{V_{v,\tau}\}, \{W_{v,\tau}\}, \{U_{v,\tau}\}$: procesos Bernoulli independientes.
- $\pi_{1,\tau}, \lambda_\tau$: parámetros del modelo PDAR(1).
- $\mu_\tau(Z)$: media de las intensidades horarias.
- $\sigma_\tau^2(Z)$: varianza de las intensidades horarias.
- $N_{v,\tau}$: intensidades horarias transformadas.
- $\mu_\tau(N)$: media de las intensidades horarias transformadas.
- $\sigma_\tau^2(N)$: varianza de las intensidades horarias transformadas.
- $Q_{v,\tau}$: variable normal estandarizada, $N[0,1]$.
- ϕ_1 : coeficiente de autoregresión del modelo AR(1).
- $\epsilon_{v,\tau}$: variable independiente con distribución normal $N[0, 1 - \phi_1^2]$.
- r_1 : coeficiente de autocorrelación de retardo 1.

REFERENCIAS

- Chebaane, M., J. D. Salas, y D. C. Boes (1995), *Product Periodic Autoregressive Processes for Modeling Intermittent Monthly Streamflows*, Water Resour. Res., Vol. 31(6), 1513-1518.
- Chung, Chen-hua (1999), *Probability Distribution, Risk, and Return Period of Dependent Hydrologic events*, Dissertation of Civil Engineering, Colorado State University.
- Gabriel, K. R y Neumann, J. (1962), *A Markov Chain Model for Daily Rainfall Occurrence at Tel Aviv*, Quart. J. Royal Met. Soc., Vol. 88, 90-95.
- Huff, F. A. (1967), *Time Distribution of Rainfall in Heavy Storms*, Water Resour. Res., Vol. 3, 1007-1019.
- Islam, S., Entekhabi, D., Bras, R. L., y Rodríguez-Iturbe, I. (1990), *Parameter Estimation and Sensitivity Analysis for the Modified Bartlett-Lewis Rectangular Pulses Model of Rainfall*, J. Geophys. Res., Vol. 95, 2093-2100.
- Johnson, N.L., y Kotz, S. (1970), *Continuous Univariate Distributions*, Vol.1, 300 pp., Houghton Mifflin, Boston, Mass.
- Katz, R. W., y Parlange M. B. (1995), *Generalizations of Chain-Dependent Processes: Application to Hourly Precipitation*, Water Resources Research, Vol. 31(5), 1331-1341.
- Katz, R. W. (1977a) *Precipitation as a Chain-Dependent Process*, J. Appl. Meteorol., Vol. 16, 671-676.
- Katz, R. W. (1977b) *An Application of Chain-Dependent Processes to Meteorology*, J. Appl. Meteorol., Vol. 14, 598-603.
- Katz, R. W., y Garrido, J. (1994), *Sensitivity Analysis of Extreme Precipitation Events*, Int. J. Climatol., Vol. 14, 985-999.
- Katz, R. W., y Parlange, M. B. (1993), *Effects on an Index of Atmospheric Circulation on Stochastic Properties of Precipitation*, Water Resour. Res., Vol. 29, 2335-2344.
- Kavvas, M. L. y Delleur, J. W. (1975), *The Stochastic and Chronological Structure of Rainfall Sequences-Application to Indiana*, technical report 57, Purdue University, Water Resources Center.
- LeCam, L. (1961), *A Stochastic Description of Precipitation*, in Fourth Berkeley Symposium on Mathematics, Statistics, and Probability Proceedings, edited by J. Neyman, 165-186, University of California Press, Berkeley.
- Lindgren, B. W. (1968), *Statistical Theory*, 2nd edn, Macmillan, London, pp. 521.
- Nguyen, V.-T.-V. (1984), *On Stochastic Characterization of Temporal Storm Patterns*, Water Sci. Technol., Vol. 16, 147-153.
- Nguyen, V.-T.-V. y Rousselle, J. (1981), *A Stochastic Model for the Time Distribution of Hourly Rainfall Depth*, Water Resour. Res., Vol. 17, 399-409.
- Obeyssekera, J., Tabios, G. y Salas, J.D. (1987), *On Parameter Estimation of Temporal Rainfall Models*, Water Resour. Res., Vol. 23(10), 1837-1850.
- Rodríguez-Iturbe, I., Febres de Power, B., y Valdés. J. B. (1987), *Rectangular Pulses Point Process Models for Rainfall: Analysis of Empirical Data*, J. Geophys. Res., Vol. 92, 9645-9656.
- Rodríguez-Iturbe, I., Cox, D. R., y Isham, V. (1988), *A Point Process Model for Rainfall: Further Developments*, Proc. R. Soc. London A, Vol. 417, 283-298.
- Roesner, L. A. y Yevjevich, V. (1966), *Mathematical Models for Time Series of Monthly Precipitation and Monthly Runoff*, Hydrology Paper 15, Colorado State University, Fort Collins, Colorado.
- Roldán, J. y Woolhiser, D. A. (1982), *Stochastic Daily Precipitation Models*, Water Resour. Res., Vol. 18(5), 1451-1459.
- Salas, J. D., Delleur, J. W., Yevjevich, V., y Lane, W. L. (1980), *Applied Modeling of Hydrologic Time Series*, Water Resources Publications.
- Salas, J. D. (1993), *Chapter 19: Analysis and Modeling of Hydrologic Time Series*, Handbook of Hydrology, ed. David R. Maidment, McGraw-Hill.
- Sanso, B., y Guenni, L. (1999), *A Stochastic Model for Tropical Rainfall at a Single Location*, J. of Hydrol., Vol. 214(1-4), 64-73.
- Stuart, A., y Ord, J. K. (1987), *Kendall's Advanced Theory of Statistics*, Vol.1, 5th ed., 604 pp., Oxford University Press, New York.
- Todorovic, P., y Woolhiser, D. A. (1975), *A Stochastic Model of n-day Precipitation*, J. Appl. Meteorol., Vol. 14, 17-24.
- Wilks, D. S. (1998), *Multisite Generalization of a Daily Stochastic Precipitation Generation Model*, J. of Hydrol., Vol. 210(1-4), 178-191.