

## SIMULACIÓN MONTE CARLO PARA REGISTROS DE PRECIPITACIÓN PLUVIAL

### MONTE CARLO SIMULATION FOR RAINFALL RECORDS

V. M. PEÑAFIEL<sup>†</sup> & M. ANDRADE<sup>‡</sup>

Laboratorio de Física de la Atmósfera  
Instituto de Investigaciones Físicas, Carrera de Física  
Universidad Mayor de San Andrés  
c. 27 Cota-Cota, Campus Universitario, Casilla de Correos 8639  
La Paz – Bolivia

Se propone un modelo Monte Carlo simple orientado a simular los patrones anuales de precipitación pluvial diaria. El modelo reproduce el comportamiento cualitativo de los veinte registros (anuales) empíricos disponibles para tres estaciones colocadas a diferentes alturas sobre el nivel del mar y, aproximadamente, a lo largo del eje transversal del territorio boliviano. Se hizo énfasis, sobre todo, en las variaciones de la precipitación acumulada anual, la frecuencia anual de días secos, la correlación lineal entre registros anuales y las propiedades de la transformada de Fourier tomada sobre el conjunto total de datos.

*Descriptor:* Aplicaciones del método Monte Carlo— precipitación pluvial

Código(s) PACS: 87.55.kh, 92.40.eg

#### ABSTRACT

We propose a simple Monte Carlo model for the purpose of simulating annual profiles of daily rainfall. This model reproduces the qualitative behaviour of 20 annual records for 3 stations located at different heights above the sea level and along the transversal axis of the Bolivian territory. We have concentrated mainly on the accumulated annual rainfall, the annual frequency of dry days, the linear correlation of annual records and the properties of the Fourier transform taken over the totality of the data.

*Subject headings:* Monte Carlo method and applications — rainfall

#### 1. INTRODUCCIÓN

La principal característica de los registros anuales de precipitación diaria es la estocasticidad. No es posible, sin embargo, atribuirles alguna distribución probabilística, conocida o no (al menos en la región aquí estudiada). La razón de ello es la ocurrencia, también marcadamente aleatoria, de días “secos” (precipitación igual a cero), aislados o en serie, cuyo principal efecto es de anular cualquier tendencia consecutiva en los valores para los días “húmedos”. El recurrir, más bien, a un *algoritmo* de simulación parece inevitable y, efectivamente, es la opción adoptada en una ya muy abundante literatura al respecto.

Un modelo, ampliamente aceptado, recurre a procesos de Markov de primer o segundo orden para simular la transición de días secos a húmedos y viceversa (por ejemplo, Katz (1974), Katz (1977a), Katz (1977b), Hayhoe (2000) y Deni & Jemain (2009)); los valores de los días húmedos se extraen de alguna distribución probabilística ajustada a la

precipitación acumulada desde un registro empírico dado.

Para los registros sobre territorio boliviano, disponibles para el presente trabajo SENAMHI (2010), ese modelo tiende a una innecesaria complicación, al parecer, debido a ciertas estructuras con series persistentes de días secos aún en las regiones más húmedas de las series temporales.

En las siguientes secciones se describe en detalle tanto las características de tales registros empíricos cuanto el método alternativo -más simple- empleado para la simulación. Y, claro, se comparan resultados mostrando algunas de las implicaciones.

#### 2. REGISTROS EMPÍRICOS

Se han elegido datos de precipitación adquiridos en tres estaciones colocadas, aproximadamente, sobre el eje transversal del territorio boliviano, como se puede apreciar de los valores de latitud, longitud y altura sobre el nivel del mar que se muestran en la tabla 1. Se trata de series temporales con 7304 valores cada una correspondientes a 20 años, desde el 1 de Enero de 1989 hasta el 31 de Diciembre de 2008.

La inspección de los datos para E1, mostrados en las figuras 1, 2(a), y 2(b), revela: (i) que la única pe-

<sup>†</sup>vmiguel@fumsa.edu.bo

<sup>‡</sup>mandrade@fumsa.edu.bo

TABLA 1  
POSICIÓN DE LAS ESTACIONES.

Estación	Lat.	Long.	a.s.n.m. [m]
E1 (Ayo Ayo)	-17.083	-68.000	3880
E2 (Cochabamba)	-17.417	-66.183	2548
E3 (Pto. Suarez)	-18.983	-57.817	134

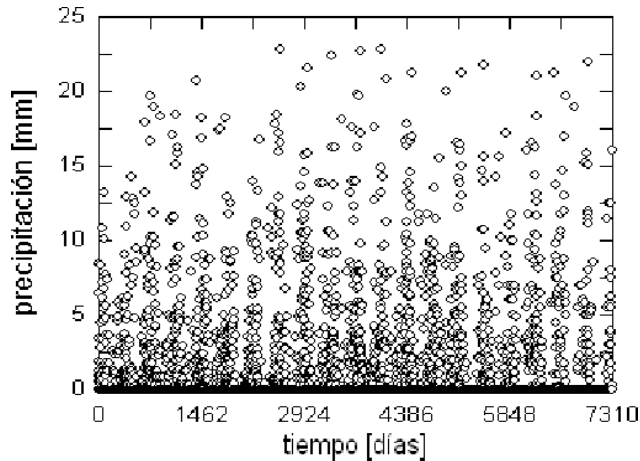


FIG. 1.— Datos de precipitación en E1 (Ayo Ayo) desde el 1 de Enero de 1989 hasta el 31 de Diciembre de 2008. Es notoria la periodicidad anual afectada por dispersión en valores pluviales altos.

riodicidad significativa es la anual; (ii) que existe dispersión adicional de los datos entre los valores medios y los superiores y (iii) que el número y la distribución de días secos son eventualidades independientes de las de los días húmedos.

La propiedad (iii) sigue de que, no obstante ser -en buena parte- responsables de la periodicidad anual (zonas secas centrales), los valores nulos de precipitación se distribuyen también en las zonas húmedas de tal ciclo (puntos aparentemente continuos sobre las abscisas de la Fig. 1.).

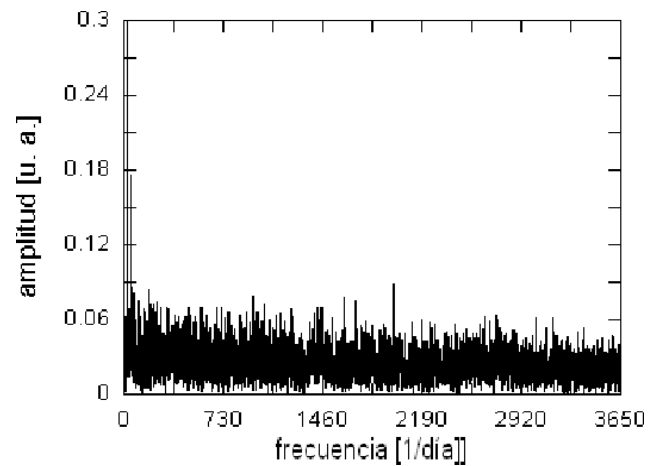
Análogamente, las figuras 9, 10(a), 10(b), 16, 17(a) y 17(b) permiten reconocer las características (i), (ii) y (iii) también en los registros de las estaciones E2 y E3. En general, como se verá más adelante, el comportamiento de la precipitación pluvial es estadísticamente similar para las tres estaciones y, consecuentemente, al menos en toda la región media del territorio boliviano.

Antes de discutir la naturaleza de ese comportamiento y, más bien, para permitir la discusión, conviene primero explicar las motivaciones y detalles del proceso de simulación empleado para imitar numéricamente a las series temporales empíricas.

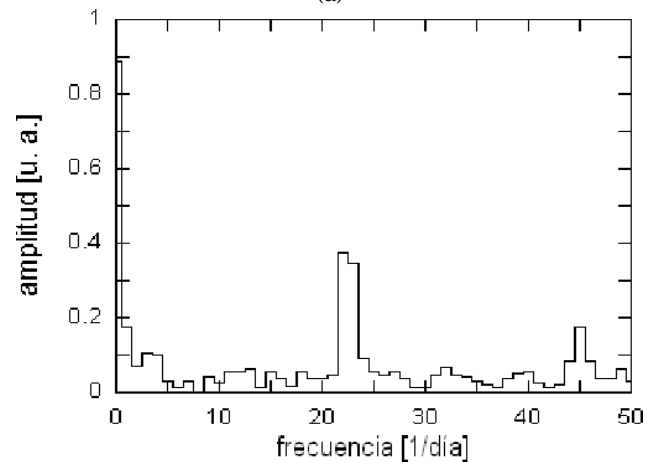
### 3. MODELO MONTE CARLO

La figura 3 es un ejemplo (entre 365 posibles) de frecuencia (normalizada) de precipitación para un mismo día durante 20 años. El ajuste de la distribución alfa

$$E(t) = ae^{-\alpha t}. \quad (1)$$



(a)



(b)

FIG. 2.— (a) Transformada rápida de Fourier (TRF) sobre los datos de la Fig. 1. Excepto la región de períodos grandes, el patrón es notoriamente característico al de ruido aleatorio. (b) Primeros 50 puntos de la TRF de la Fig. 2a.; el pico corresponde al obvio período anual  $8192/23 = 356 < T < 8192/22 = 372.36$ . El segundo pico, más pequeño, es sólo una resonancia en  $8192/45 = 182.04$  (medio año) y no es significativo.

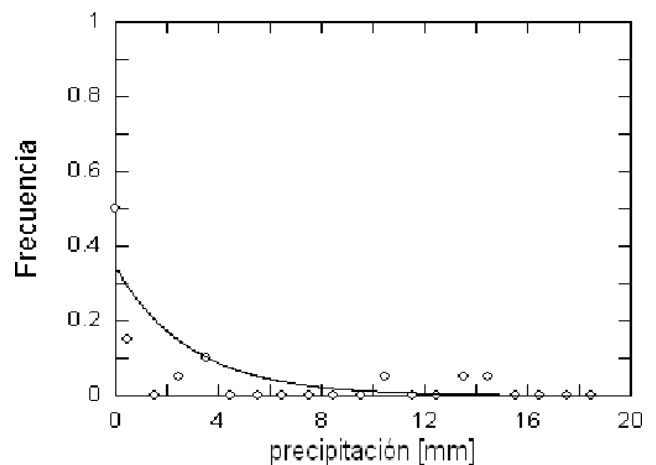


FIG. 3.— Ajuste de la distribución alfa, por mínimos cuadrados, a las frecuencias relativas de precipitación correspondientes a un mismo día del año durante los 20 años del registro total.

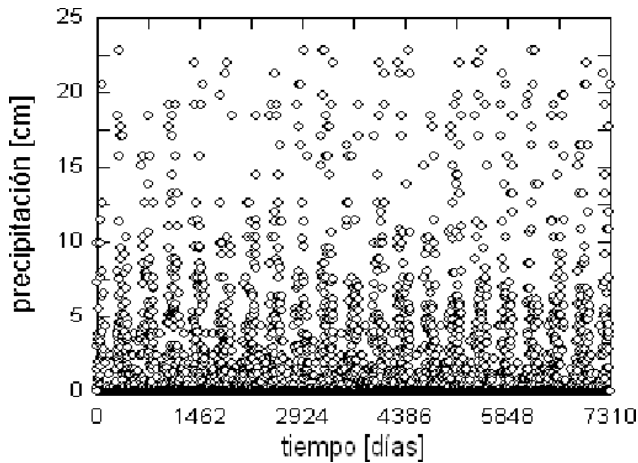


FIG. 4.— Datos simulados para la estación E1 donde se aprecia la aproximación, en valores máximos y dispersión relativa, al patrón formado por los datos empíricos de la Fig. 1.

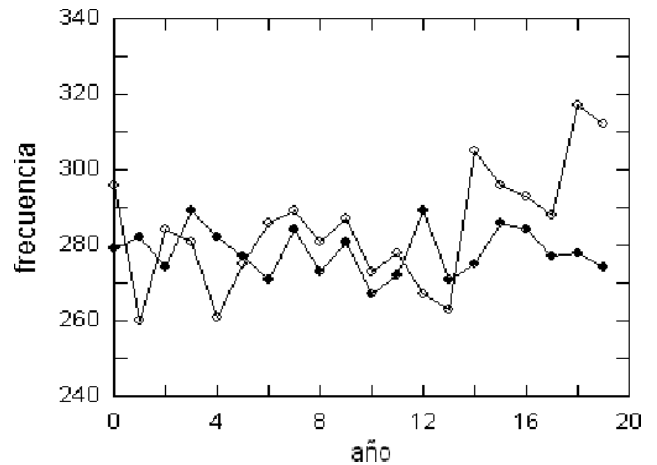
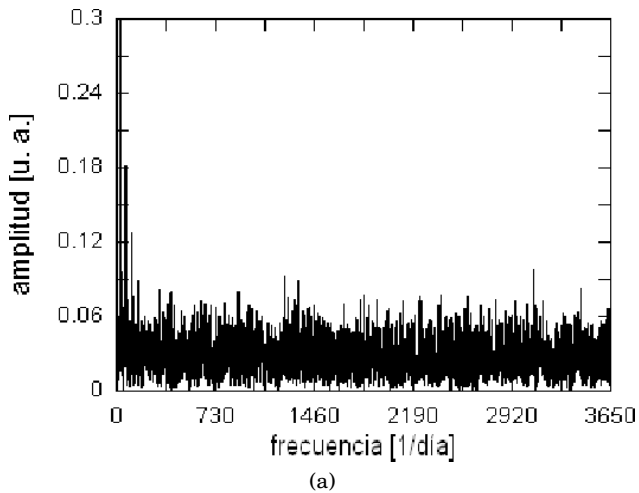
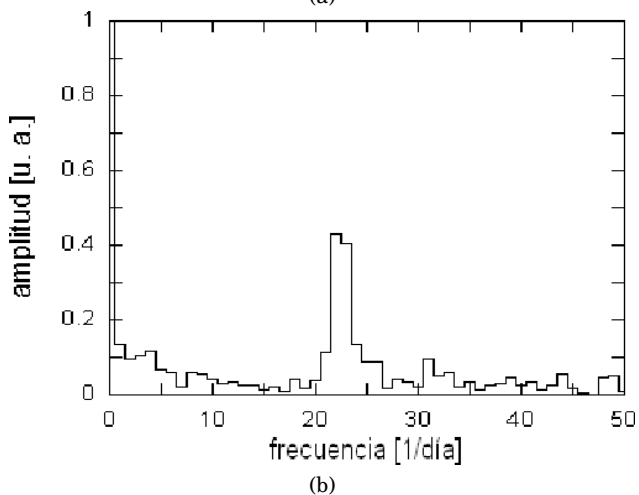


FIG. 6.— Comparación entre frecuencias anuales de días secos empíricas y simuladas (puntos llenos). La aproximación es aceptable.



(a)



(b)

FIG. 5.— (a) TRF a los datos de la Fig 3. La características cualitativas son completamente análogas a las observadas en la Fig. 2a. (b) Histograma de los primeros 50 puntos de TRF de la Fig. 2a. El único pico es el esperado para un período anual  $356.17 < T < 372.36$ .

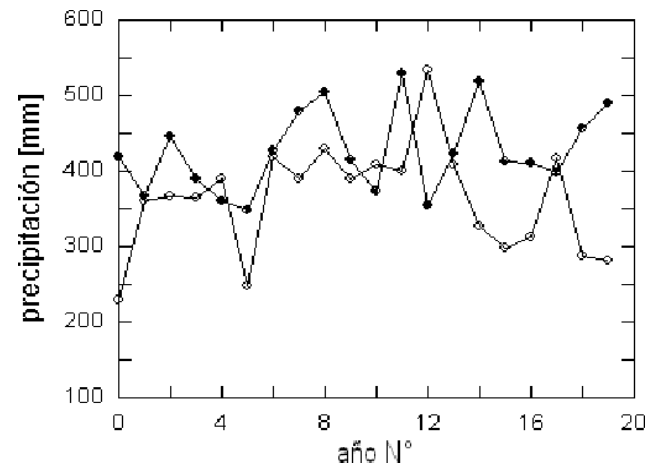


FIG. 7.— Comparación de las precipitaciones anuales empírica y simulada (puntos llenos). El comportamiento de las secuencias son muy similares aunque los valores numéricos, claro, no coinciden.

es pertinente sólo por las tendencias de probabilidad (alta para precipitación baja y viceversa). Sin embargo, en varios otros casos se tiene días con 20 valores cero de precipitación y aún grupos de días secos para los cuales, por supuesto, ninguna distribución es apropiada.

Para producir numéricamente este efecto y también la dispersión mencionada en la sección anterior, en vez de considerar la matriz de probabilidades de transición  $p_{ij}$  para un proceso de Markov, se ha preferido asignar una probabilidad fija  $\beta_s$  a la porción seca ( $90 < i < 270$ ) y otra  $\beta_h$  a la zona húmeda ( $270 < i < 90$ ) como umbrales encima de los cuales el día se considera seco; y un parámetro  $\alpha_0$  (que determina los valores máximos de precipitación) para la distribución exponencial (1). El proceso de simulación Monte Carlo, entonces, procede como sigue:

Para cada valor del índice  $0 \leq i \leq 364$  se sortean dos números aleatorios,  $\xi_1$  y  $\xi_2$ , el primero con 100 valores equidistantes en el intervalo  $[0, 1]$  y el segundo con  $k$  valores enteros  $1 \leq \xi_2 \leq k$ . Ahora, si  $x$  representa a la precipitación se tiene, para la zona

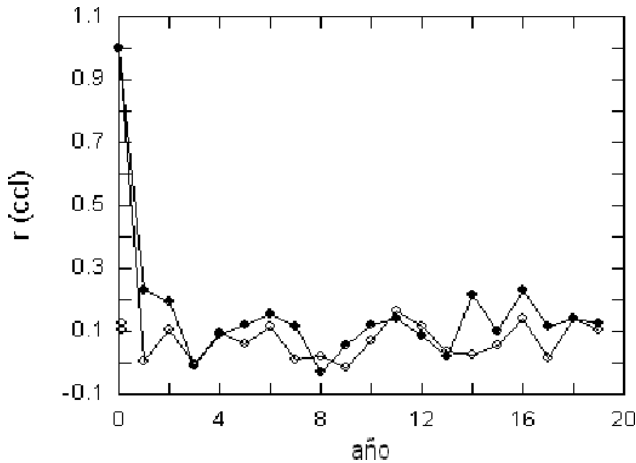


FIG. 8.— Correlaciones lineales entre los primeros registros anuales y los 19 siguientes para los datos empíricos y los simulados (puntos llenos). La similitud de los esquemas es probabilísticamente notable.

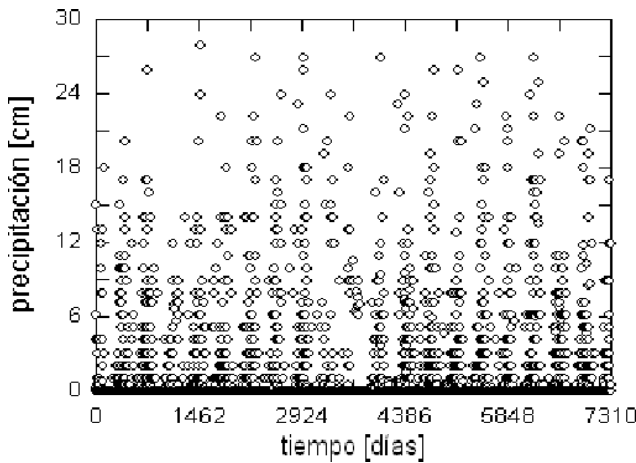


FIG. 9.— Datos de precipitación en E2 (Cochabamba) desde el 1 de Enero de 1989 hasta el 31 de Diciembre de 2008. La estructura básica de periodicidad anual y "ruido" dispersivo están presentes (como es el caso para E1).

seca,

$$x = y \text{ si } \xi_1 < \beta_s$$

$$x = 0 \text{ si } \xi_1 \geq \beta_s$$

donde

$$y = -\frac{1}{\alpha} \ln(1 - \xi_1)$$

$$\alpha = \xi_2 \alpha_0 .$$

Análogamente, para la región húmeda,

$$x = y \text{ si } \xi_1 < \beta_h$$

$$x = 0 \text{ si } \xi_1 \geq \beta_h$$

con las mismas expresiones para  $y$  y  $\alpha$ .

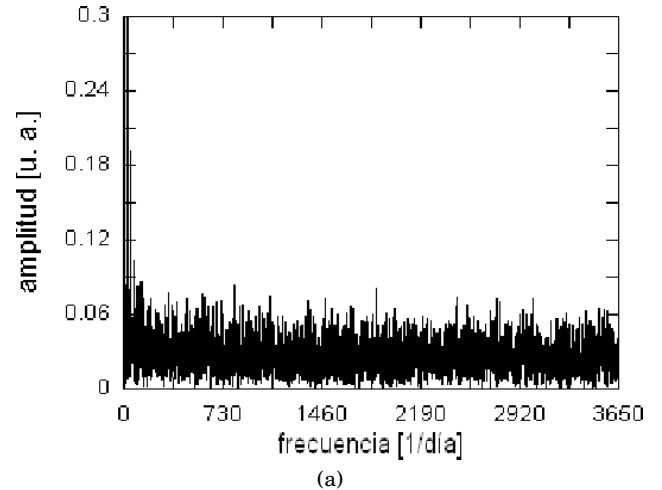


FIG. 10.— (a) TRF a los datos de la Fig 9. La características cualitativas son muy semejantes a las observadas en la Fig. 2a. (b) Histograma de los primeros 50 puntos de TRF de la Fig. 10a. con el único pico esperado para un período anual  $356.17 < T < 372.36$ .

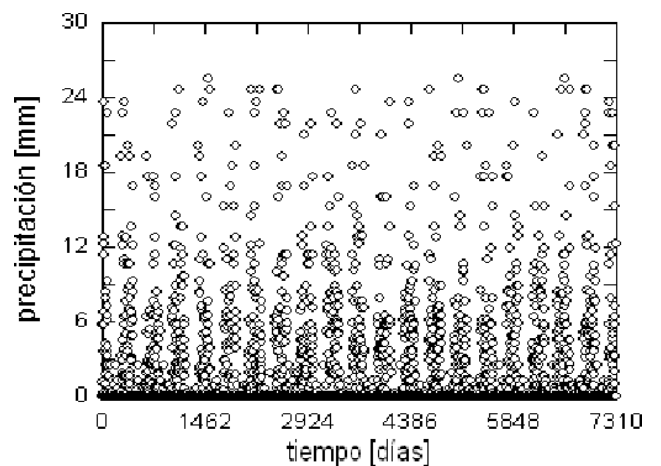


FIG. 11.— Datos simulados para la estación E2 con un patrón, en valores máximos y dispersión relativa, análogo al de los datos empíricos de la Fig. 9.

El sorteo de los números  $\xi$  para las variables  $\beta$  y  $k$  se hizo usando una rutina adaptada del "Mersenne Twister" Matsumoto & Nishimura (1998) para el ensamblador de 32 bites.

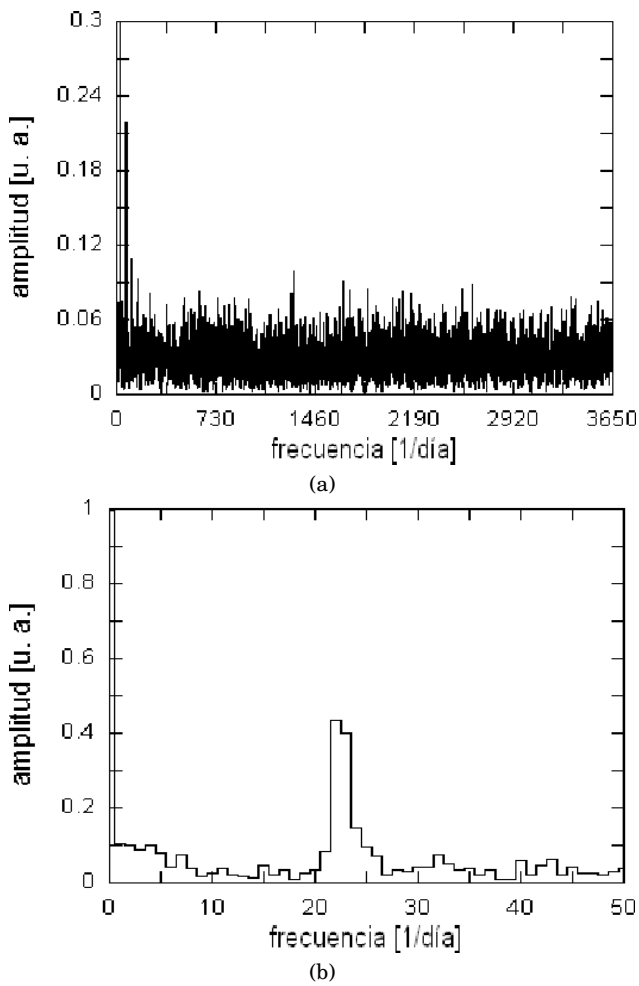


FIG. 12.— (a) TRF de la serie en la Fig 11. El comportamiento del espectro es parecido al de la Fig. 10a. (b) Primeros 50 puntos de TRF de la Fig. 12a. El pico es también el esperado para un periodo anual  $356.17 < T < 372.36$ .

Los parámetros  $\alpha_0$ ,  $\beta_s$ ,  $\beta_h$ , y  $k$  deben ser ajustados “manualmente” hasta que la altura de las ordenadas y la dispersión de los puntos se correspondan, tan aproximadamente como sea posible, con los registros empíricos. Por ejemplo, las figuras 4, 5(a) y 5(b) ilustran el resultado de la simulación, con los valores que figuran en la tabla 2, para la estación E1. La semejanza entre los registros empírico y simulado, así como entre sus respectivas transformadas de Fourier, son muy aceptables, más aún considerando las comparaciones cualitativas de otras propiedades que se discute en la sección siguiente.

#### 4. REGISTROS SIMULADOS

Como se ha supuesto que el comportamiento probabilístico de la ocurrencia de días secos es independiente de la distribución para los niveles de precipitación, se esperaría que la evolución del número total anual de días secos y la de precipitación total anual se comportaran de manera independiente; por otra parte, es obvio que un mayor número de días secos debe tender a disminuir la cantidad total de precipitación. Estos efectos contrarios se manifiestan realmente en el respectivo coeficiente de correlación.

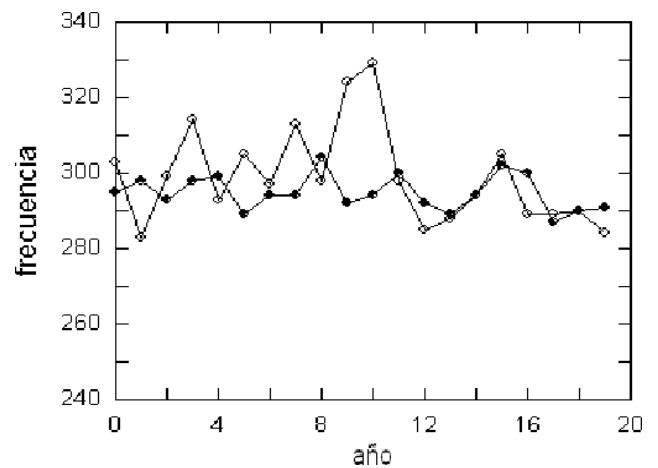


FIG. 13.— Frecuencias anuales empíricas y simuladas (puntos llenos) de días secos correspondientes a la estación E2.

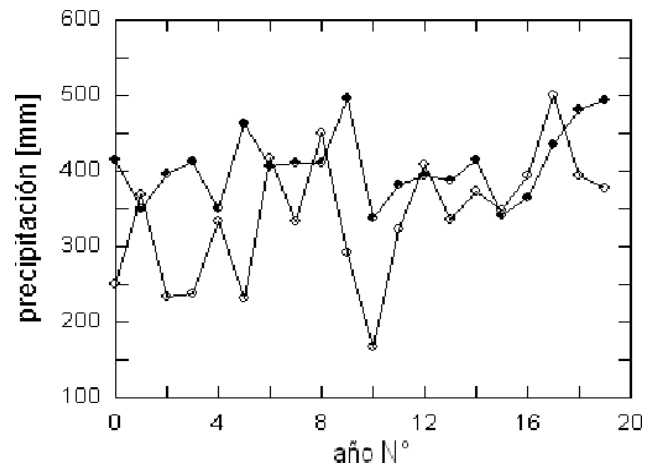


FIG. 14.— Precipitaciones anuales empíricas y simuladas (puntos llenos) de la estación E2.

Por ejemplo, para la estación E1,  $r_e = -0.555$  y  $r_s = -0.387$ , esto es, el registro empírico presenta algo más anticorrelación que el registro simulado. Por lo demás, como se aprecia en las figuras 6 y 7, el comportamiento estadístico es bastante similar para ambos registros.

Es también interesante la comparación entre las secuencias de coeficientes de correlación de los datos del primer año con los 19 restantes para los dos registros (Fig. 8). Estas correlaciones que oscilan alrededor de cero permiten explicar el hecho de que todas las comparaciones sólo tienen significado estadístico: las series empíricas, al igual que las simuladas, no muestran regularidades consistentes (excepto la anual) y admiten la interpretación de que cada patrón anual es sólo una instancia (un ensayo o muestra individual) aleatoria sobre un número - en principio- infinito de posibilidades (mientras las condiciones físicas causantes del fenómeno cíclico de precipitación pluvial se mantengan).

Estos mismos resultados pueden ser observados en el caso de las estaciones E2 y E3. Las figuras 9, 10(a), 10(b), 16, 17(a) y 17(b) muestran que las características estructurales son completamente simi-

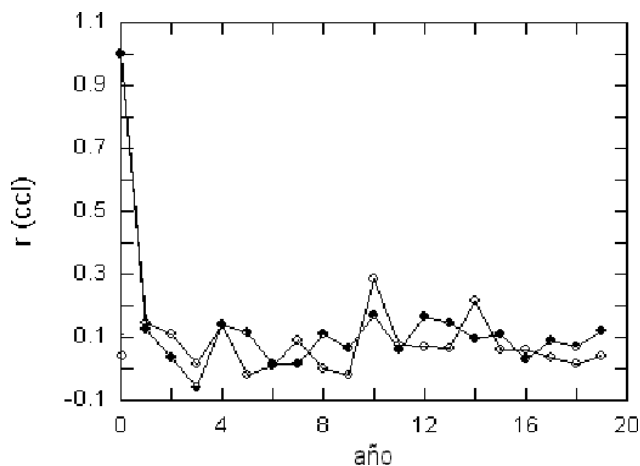


FIG. 15.— Coeficientes de correlación lineal entre los primeros registros anuales y los 19 siguientes para los datos empíricos y los simulados (puntos llenos) de la estación E2.

TABLA 2  
PARÁMETROS DE LAS SIMULACIONES.

Estación	$\beta_s$	$\beta_h$	$\alpha_0$	$k$
E1	0.10	0.40	0.22	4
E2	0.05	0.35	0.0165	4
E3	0.05	0.30	0.005	6

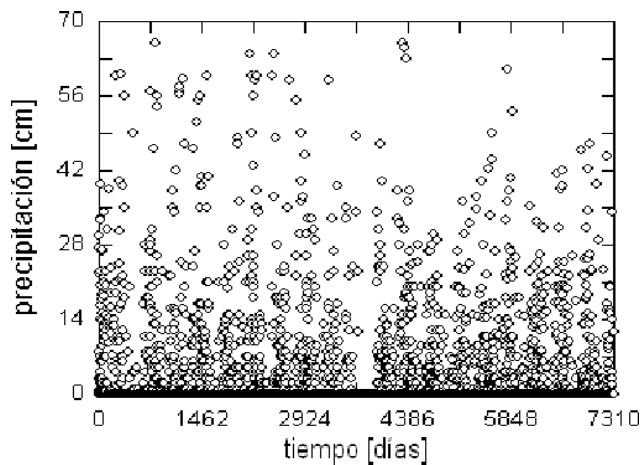


FIG. 16.— Datos de precipitación en E3 (Puerto Suarez) desde el 1 de Enero de 1989 hasta el 31 de Diciembre de 2008. Hay periodicidad anual y dispersión, como en E1.

lares, independientemente de la posición y altura sobre el nivel del mar. La misma configuración aparece en las figuras 11, 12(a), 12(b), 18, 19(a) y 19(b) que corresponden, respectivamente, a los registros simulados para E2 y E3.

En todos los casos el pico de periodicidad se encuentra entre los índices 22 y 23 los cuales, para un registro de 8192 (7304 datos originales más un relleno de 888 ceros), indica que T se encuentra en el intervalo [356.17; 372.36]. Veinte años de datos son todavía poco para observar periodicidades mayores a un año y las menores carecen de significado físico; la comparación entre espectros de potencias indica que el resto de los periodogramas contiene solamente

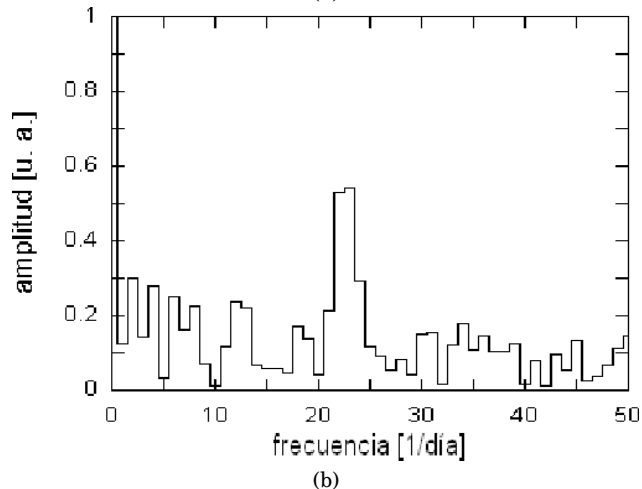
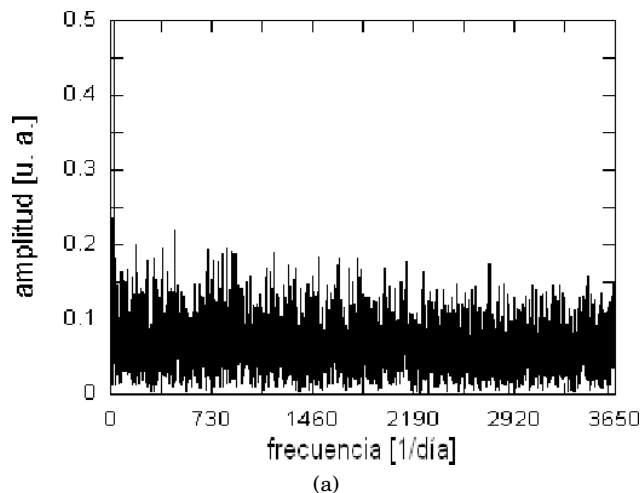


FIG. 17.— (a) TRF sobre los datos de la Fig. 16. El espectro contiene, también, ruido aleatorio excepto en la región de períodos grandes. (b) Detalle de la TRF de la Fig. 17a. El único período significativo es el anual.

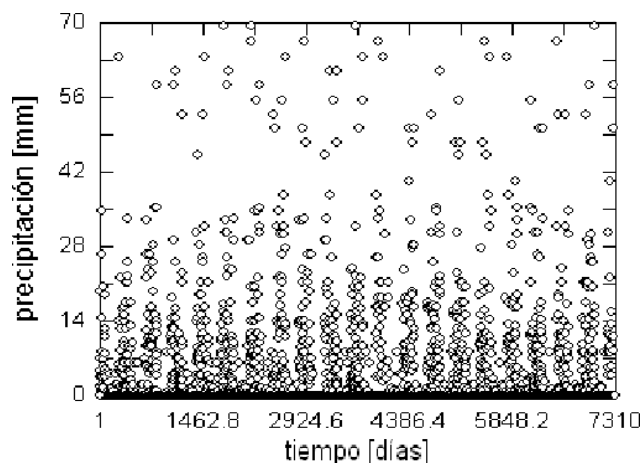


FIG. 18.— Serie simulada para la estación E3; el patrón es similar al de la Fig. 16.

ruido aleatorio.

El efecto que es menos obvio y más difícil de explicar es el de la dispersión de puntos, especialmente para valores medios y superiores de precipitación. Al abandonar el modelo de probabilidades de tran-

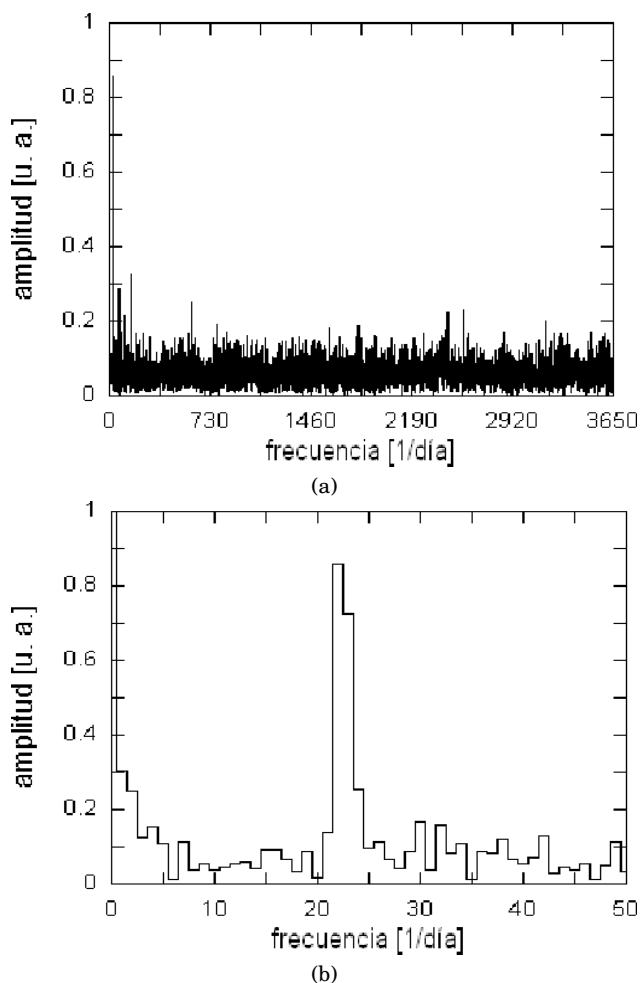


FIG. 19.— (a) Espectro de potencias del registro simulado para E3. Las características no difieren de lo obtenido para E1 y E2. (b) Primeros 50 puntos de la TRF de la Fig. 19a. Igualmente,  $356.17 < T < 372.36$ .

sición, como ya se vio, se adopta la idea que hay una probabilidad umbral que determina si el día puede ser considerado como seco o húmedo y un valor diario *aleatorio* para el parámetro de la distribución exponencial, con la cual se sorteaba la magnitud de la precipitación (hay una distribución exponencial *aleatoria* diaria).

La tabla 2. resume la elección de parámetros de simulación para las tres estaciones. Éstos valores fueron elegidos estrictamente para imitar el comportamiento de los registros experimentales;  $\alpha_0$  para fijar los valores máximos de precipitación y  $\beta_s$ ;  $\beta_h$ ;  $k$  para controlar la dispersión.

5. RESULTADOS

Una buena estimación de los efectos locales está dada por los coeficientes de correlación entre las series empíricas de cantidad de días secos entre las tres estaciones, éstos son  $r_{s12} = -0.031$ ,  $r_{s13} = -0.4354$ ,  $r_{s23} = 0.327$ . Análogamente, los coeficientes de correlación entre las series de precipitación acumulada anual,  $r_{h12} = 0.300$ ,  $r_{h13} = -0.326$ ,  $r_{h23} = -0.114$ , indican que las tres estaciones tienen comportamiento pluvial diario independiente (estocásticamente); sin

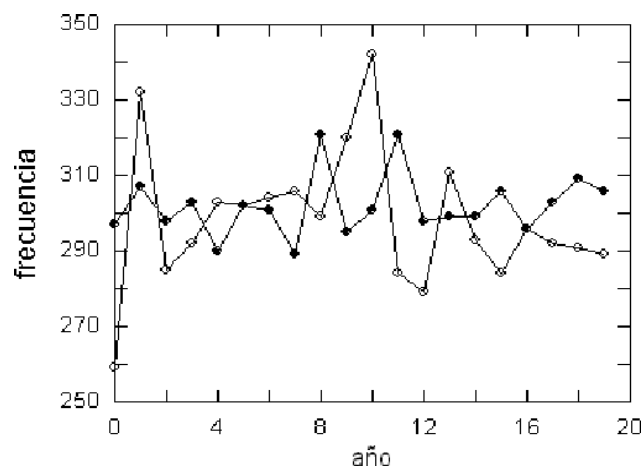


FIG. 20.— Frecuencias anuales de días secos, empíricas y simuladas (puntos llenos) para la estación E3.

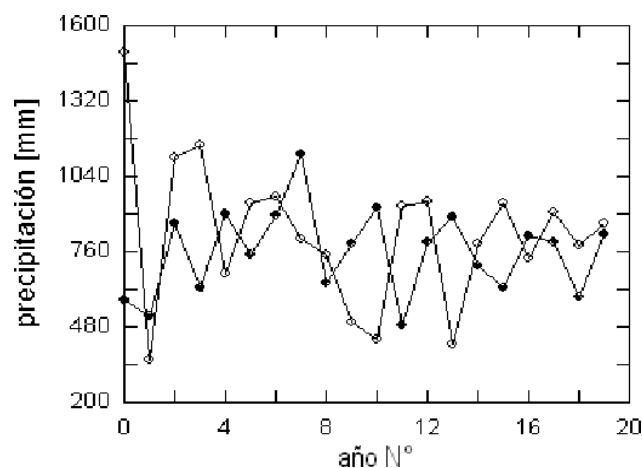


FIG. 21.— Precipitaciones anuales empírica y simulada (puntos llenos) de la estación E3.

embargo, las figuras 13, 14, 20 y 21, como las 6 y 7, permiten -por la comparación con los resultados simulados- apreciar la analogía de los tres procesos aunque, naturalmente, los valores numéricos pueden no coincidir. También es notorio el mayor grado de aleatoriedad en los registros empíricos, los cuales presentan fluctuaciones más acentuadas que sus simulaciones respectivas.

Finalmente, las figuras 15 y 22, como la 8, ilustran las correlaciones entre la serie para el primer año consigo mismo (punto en el origen) y con los 19 restantes; como se ve, los coeficientes de correlación oscilan alrededor de cero en todos los casos, apenas con pequeñas fluctuaciones, en ciertos intervalos, para los valores relativos a los datos empíricos.

6. CONCLUSIONES.

Para otros lugares y registros más largos, ciertos análisis se orientan a detectar oscilaciones significativas en los datos de precipitación (v. g. Rodríguez-Puebla et al. (1998)). Los registros disponibles para este trabajo, ya descritos, definitivamente no manifiestan otro ciclo que el anual.

Las propiedades estadísticas son simulables según lo descrito en la sección 3 y las comparaciones con

TABLA 3  
VALORES EXTREMOS

<i>Est.</i>	$P_{max}$	$C_{max}$	$C_{min}$	$A_{max}$	$A_{min}$
<i>E1(e)</i>	22.80	317	260	533.40	228.44
<i>E1(s)</i>	22.76	289	271	529.48	348.52
<i>E2(e)</i>	27.90	329	283	500.44	165.95
<i>E2(s)</i>	25.50	304	287	496.54	337.14
<i>E3(e)</i>	66.00	342	259	1502.46	356.50
<i>E3(s)</i>	69.32	321	289	1123.67	483.27

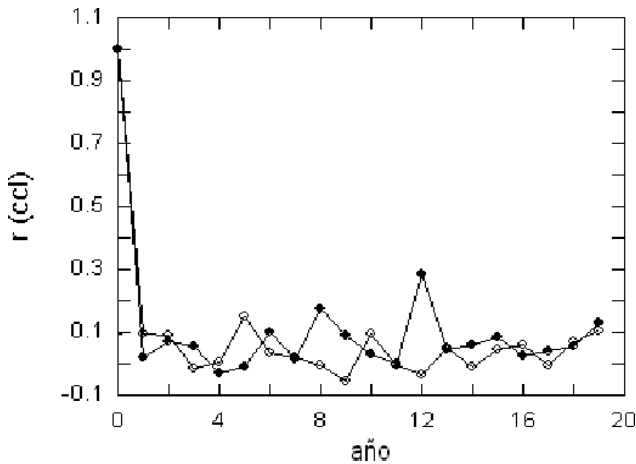


FIG. 22.— Estación E3: correlaciones lineales entre los primeros registros anuales y los 19 siguientes para los datos empíricos y los simulados (puntos llenos).

las propiedades de las series de datos, expuestas en la sección 4, por una parte, indican que el modelo adoptado es muy aceptable para producir versiones subrogadas de los datos y, por otra, que los registros empíricos contienen grados de aleatoriedad mayores que sus versiones simuladas (la simulación no incluye fluctuaciones climáticas esporádicas).

Como consecuencia, es posible considerar a los registros anuales de precipitación diaria como instancias independientes de un patrón probabilístico determinado por la superposición de condiciones físicas, un esquema similar a los *estados* aleatorios de la física estadística.

Admitido esto, para registros mayores, se podría emplear las simulaciones para detectar cambios notables en tales condiciones físicas que darían paso a cambios perceptibles en las probabilidades que determinan las distribuciones.

Se concluye que, para los registros de 20 años empleados aquí, el comportamiento estadístico es “estable” en las tres estaciones consideradas, aunque las fluctuaciones estadísticas de los registros simulados son más conservadoras. En la tabla 3, para valores de precipitación máxima  $P_{max}$  muy próximos, se advierte que el número de días secos máximo  $C_{max}$  y mínimo  $C_{min}$  están siempre dentro del intervalo definido por los extremos empíricos y lo mismo acontece para la precipitación acumulada anual máxima  $A_{max}$  y mínima  $A_{min}$ . Esto hace conjeturar la posibilidad de producir datos subrogados por otras vías y, también, elimina la de “predecir” las configuraciones anuales de precipitación.

#### REFERENCIAS

- Deni S. M. & Jemain A. A. (2009), *Met. and Atmosph. Physics* **104**, 13  
 Hayhoe H. N. (2000), *Clim. Res.* **14**, 75  
 Katz R. W. (1974), *J. Appl. Meteor.* **13**, 953  
 —. (1977)a, *J. Appl. Meteor.* **16**, 671  
 —. (1977)b, *J. Appl. Probability* **14**, 598  
 Matsumoto M. & Nishimura T. (1998), *ACM Transactions on Modeling and Computer Simulation* **8**, 3  
 Rodríguez-Puebla C., Encinas A. H., Nieto S. & Germendia J. (1998), *Int. J. Climatol.* **18**, 299  
 Servicio Nacional de Meteorología e Hidrología “SENAMHI” (2010), (Comunicación Privada)