

# EVALUACIÓN DEL MODELO DE NEYMAN-SCOTT PARA SIMULACIÓN DE LLUVIA EN UN PUNTO GEOGRÁFICO

Gerardo Adrián Riccardi

## Resumen:

Se presenta una aplicación y evaluación del modelo de simulación de series de lluvia de Neyman-Scott de pulsos rectangulares, en su formulación original. Dicho modelo está basado en la teoría de procesos de punteo, en el cual se simula la estructura celular de los campos reales de precipitación preservando los parámetros estadísticos relevantes, en un amplio rango de las escalas de agregación temporal. Esta característica vinculada a la agregación temporal hace de estos modelos herramientas útiles en estudios hidrológicos, tales como producción de escurrimiento, infiltración, recarga de acuíferos, predicción de crecidas y sequías, simulación hidrológica continua, etc.. El modelo se sustenta en la descripción de un proceso de Poisson que fija el origen de los eventos, un proceso que fija el número de celdas de lluvias generadas en cada evento y un proceso que fija el origen temporal de cada celda. Además, cada celda tiene una duración aleatoria y una intensidad aleatoria. La aplicación del modelo fue realizada en la serie de registros de la estación pluviográfica Rosario Aero (Rosario, Argentina). Las series fueron analizadas en agrupamientos mensuales con el fin de preservar la estacionariedad. Las escalas de agregaciones temporales consideradas para la determinación de los parámetros del modelo fueron 0.5, 1, 2, 3, 4, 6, 12 y 24 horas, en tanto que para la evaluación de los resultados generados se consideraron además escalas de 48, 72 y 168 horas. La generación de series sintéticas mostró un importante nivel de aproximación entre estadísticos y variables muestrales y generadas, tales como media, varianza, estructura de correlación, probabilidad de lluvia cero y valores extremos.

**Palabras clave:** modelos de lluvia; lluvias puntuales; simulación continua; pulsos rectangulares; clusters.

## INTRODUCCIÓN

La precipitación es uno de los procesos más relevantes en el sistema hidrológico, constituyendo como es bien sabido, una señal de entrada al sistema. Las mediciones puntuales han indicado la alta variabilidad temporal y espacial del proceso por lo que gran cantidad de trabajos han sido llevados a cabo para la representación de las secuencias temporales de lluvias y su distribución areal. En la década del 80 los trabajos Kavvas y Delleur (1981), Waymire y Gupta (1981), Rodríguez-Iturbe et al. (1984) y Valdes et al. (1985), entre otros, abordaron el modelado temporal de lluvias mediante el uso de teoría de procesos puntuales (Cox e Isham, 1980). Avances posteriores fueron presentados por Marien y Vandewiele (1986), Rodríguez-Iturbe (1986), Rodríguez-Iturbe et al. (1987a y b), Rodríguez-Iturbe et al. (1988), Cowpertwait, (1991),

Onof y Wheeler (1993), Salsón y García-Bartual (1998), Calenda y Napolitano (1999), Willems (1999 y 2001), Salsón (2001) y García-Bartual (2002). Dado que la información de lluvias está disponible a determinadas escalas de agregación, como lámina de lluvia en un determinado intervalo de tiempo (minutos, horas, días), los parámetros de los modelos continuos deben ser estimados a partir de datos agregados. Particular interés han cobrado los denominados modelos de Neyman-Scott (N-S) (Rodríguez-Iturbe et al., 1984) y Barlett-Lewis (B-L) (Rodríguez-Iturbe et al., 1987c), que simulan la estructura celular de los campos de lluvia y que son capaces de preservar, los parámetros estadísticos muestrales en un amplio rango de la escala de agregación temporal. En lo que concierne a grandes escalas temporales (días, semanas) compatibles con, entre otros procesos, las escalas de los procesos predominantes en el movimiento de flujo subterráneo,

Consejo de Investigaciones Universidad Nacional de Rosario. Centro Universitario Rosario de Investigaciones Hidroambientales. FCEIA. UNR. José Hernández 986. Telefax 0341 480 8541. 2000 Rosario. Argentina. e\_mail: riccardi@fceia.unr.edu.ar

Artículo recibido el 4 de Septiembre de 2003, recibido en forma revisada el 14 de diciembre de 2004 y aceptado para su publicación el 12 de abril de 2005. 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.

Entekhabi et al. (1989) introdujo una modificación al modelo N-S con un parámetro adicional con el objeto de mitigar la tendencia del modelo N-S original a sobreestimar la probabilidad de tiempo seco para grandes escalas. Posteriormente Velghe et al. (1994) argumentó, que aunque el modelo modificado produce mejores valores de probabilidad de altura de lluvia cero (o tiempo seco) introduciendo una complejidad adicional en la estimación de los parámetros del modelo, no se preserva en la misma forma que el modelo N-S original las propiedades de 2do. grado del proceso de lluvia.

En el presente trabajo se presenta aplicación del modelo N-S en la serie de lluvias registradas en la estación Rosario Aero (Argentina). En razón de haberse verificado una disparidad entre la función de varianza histórica y la generada por el modelo al utilizar en la toma de decisión de los parámetros del modelo, el clásico método de minimización de una función objetivo y una mejora posterior del ajuste, se empleó en este trabajo, un método de selección híbrido basado en una evaluación ponderada de: la minimización de una función objetivo, la mejora del ajuste aplicado a la comparación de estadísticos históricos y calculados y la mejora del ajuste aplicado a la función de varianza histórica y calculada.

### DEFINICIÓN DEL MODELO N-S

El proceso de agrupamiento (cluster) de Neyman-Scott (N-S) es un proceso de agrupamiento de Poisson de modo que un número aleatorio de puntos están distribuidos independientemente alrededor del centro del agrupamiento de acuerdo a una determinada función de distribución de la distancia (Waymire and Gupta, 1981). Según Waymire and Gupta (1981) este proceso de agrupamiento fue introducido por J. Neyman y E. Scott en el año 1958 en problemas de cosmología. En el proceso de N-S cada evento de lluvia es originado por un mecanismo disparador, denominado *origen del evento* y descrito por un proceso de Poisson (representando físicamente por ej. la llegada de un frente atmosférico o una tormenta convectiva). Cada tormenta es conceptualizada como un agrupamiento de celdas elementales de lluvia; el centro del agrupamiento se posiciona temporalmente en los puntos definidos por el proceso de Poisson que fija el comienzo del evento; el número de celdas elementales en cada agrupamiento es una variable aleatoria independiente e idénticamente distribuida. La cantidad de lluvia asociada en cada celda elemental es una variable aleatoria y la secuencia de tal variable es independiente e idénticamente distribuida. Cada celda elemental se ubica aleatoriamente en el tiempo a partir del

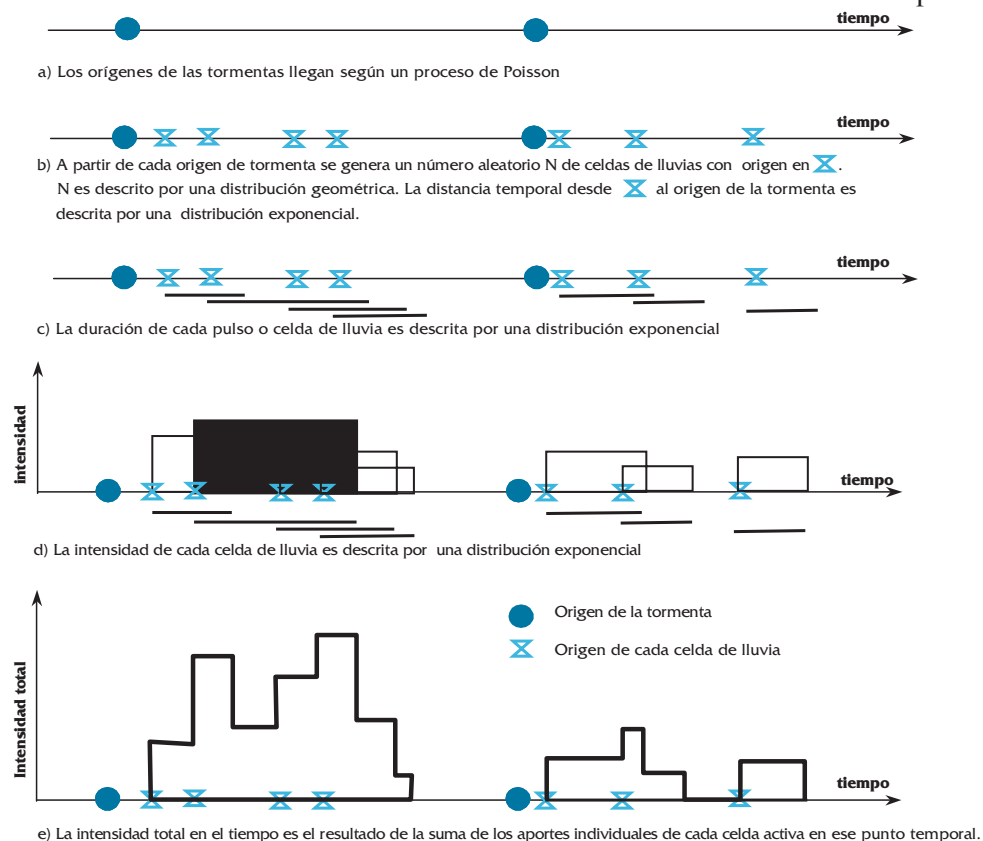


Figura 1. Esquema de representación de tormentas del Modelo de Neyman-Scott (Cowpertwait, 1991)

## EVALUACIÓN DEL MODELO DE NEYMAN-SCOTT PARA SIMULACIÓN DE LLUVIA EN UN PUNTO GEOGRÁFICO

correspondiente centro del agrupamiento, con una distancia temporal independiente e idénticamente distribuida. Originalmente el modelo N-S aplicado a generación de lluvias fue desarrollado para pulsos instantáneos (por ejemplo Rodríguez-Iturbe et al., 1984), vale decir, cada tormenta se representaba por medio de un grupo de celdas o pulsos con duración nula. En un posterior desarrollo Rodríguez-Iturbe (1986) introduce la duración del pulso por lo que la variable que fijaba la cantidad de lluvia es sustituida por dos variables aleatorias e idénticamente distribuidas, una variable describe la intensidad de lluvia y la segunda la duración de cada celda o pulso. Dentro de cada pulso la intensidad es considerada constante por lo cual el pulso adquiere la forma rectangular. El esquema de representación de tormentas del modelo de Neyman-Scott se ilustra en la Figura 1.

En el Apéndice se presenta la descripción del modelo de N-S, las ecuaciones constitutivas, el análisis de fluctuaciones y la estimación de los parámetros. Las ecuaciones del Apéndice son numeradas desde (1A) hasta (24A).

### APLICACIÓN DEL MODELO EN LA ESTACIÓN ROSARIO-AERO

#### Caracterización de la muestra

##### *Información pluviográfica empleada*

El modelo fue aplicado en una estación pluviográfica de Rosario (Santa Fe, Argentina) denominada Rosario-Aero, en la serie desde 1986 a 1998 (DGHyS, 1999). La serie comprendió 1036 eventos lluviosos. Si bien la estación tiene capacidad de registro de datos cada 5 minutos, con tal discretización sólo se registraron un 36% de las tormentas involucradas en la serie, con lo cual a los efectos de la aplicación del modelo se consideró un intervalo de agregación superior.

La elección del intervalo de agregación se fundamentó en : (a) el mínimo valor posible con información en la Estación Rosario-Aero en al menos el 95 % de las tormentas; (b) la existencia de información de distribuciones internas en series parciales de estaciones pluviográficas vecinas para las tormentas sin información en detalle en Estación Rosario-Aero y (c) confiabilidad del método de completamiento de la distribución interna verificado en base a tormentas registradas simultáneamente en todas las estaciones de la región. El mínimo intervalo de agregación adoptado de los registros de lluvia fue de 1/2 hora. Dicho intervalo comprende a más del 95% de las tormentas observadas. La estación Rosario-Aero tiene 5 estaciones vecinas en un radio no mayor a 35 km con registros parciales de la serie de lluvias en estudio (Arnold, Ibarlucea, DNC, Zavalla y Planetario) por lo que el completamiento de la distribución interna en intervalos de 1/2 hora en el 5% de las tormentas de la serie se llevó a cabo mediante la aplicación de un algoritmo de representación de campos paramétricos continuos desarrollado por Zimmermann y Silber (2002) que utiliza una formulación estadístico-matemática basada en estructuras espaciales de correlación.

#### *Estadísticos muestrales*

El modelo fue aplicado en series mensuales con el objeto de preservar la estacionariedad del proceso. Para cada mes fueron determinados estadísticos como: media, varianza, covarianza y coeficientes de autocorrelación de orden 1 y 2 y probabilidad de intervalos de tiempo seco para escalas de agregación de 0.5, 1, 2, 3, 4, 6, 12 y 24 horas. Los estadísticos correspondientes al mes de marzo se presentan en la Tabla 1. A los efectos ilustrativos se expone a marzo como mes testigo, siendo éste uno de los meses más lluviosos en la serie observada.

Tabla 1. Estación Rosario-Aero. Serie 1986-1998. Mes de Marzo. Estadísticos para distintas escalas de agregación.

| t (h) | Cant. Datos | E(h)<br>(mm) | Var(h)<br>(mm <sup>2</sup> ) | Covar <sub>1</sub> (h)<br>(mm <sup>2</sup> ) | r <sub>1</sub> (h) | Covar <sub>2</sub> (h)<br>(mm <sup>2</sup> ) | r <sub>2</sub> (h) | F[Y <sub>1</sub> =0] |
|-------|-------------|--------------|------------------------------|--|--------------------|--|--------------------|----------------------|
| 0.5   | 16368       | 0.0843       | 0.6397                       | 0.4069                                       | 0.6361             | 0.2665                                       | 0.4166             | 0.9600               |
| 1     | 8184        | 0.1687       | 2.0504                       | 1.2035                                       | 0.5870             | 0.7764                                       | 0.3787             | 0.9555               |
| 2     | 4092        | 0.3374       | 6.3234                       | 3.2413                                       | 0.5126             | 0.9312                                       | 0.1473             | 0.9450               |
| 3     | 2728        | 0.5061       | 12.4203                      | 4.5031                                       | 0.3626             | 0.9551                                       | 0.0769             | 0.9413               |
| 4     | 2046        | 0.6748       | 18.8716                      | 5.5910                                       | 0.2963             | 1.2677                                       | 0.0672             | 0.9272               |
| 6     | 1364        | 1.0121       | 28.7750                      | 9.9545                                       | 0.3459             | 2.2970                                       | 0.0798             | 0.9135               |
| 12    | 682         | 2.0243       | 70.5468                      | 19.7151                                      | 0.2795             | 6.0932                                       | 0.0864             | 0.8783               |
| 24    | 341         | 4.0486       | 202.7422                     | 19.7191                                      | 0.0973             | -8.1524                                      | -0.0402            | 0.8123               |

Covar<sub>k</sub> ; ρ<sub>k</sub> : covarianza y coeficiente de autocorrelación de orden k F[Y<sub>1</sub>=0]: Relación entre el número de datos ceros y el número total de datos.

### Análisis de Fluctuaciones

Para el análisis de fluctuaciones se estimaron las funciones de varianza a partir de la ec. (15A), utilizando datos agregados a 0.50 h. Posteriormente, mediante el uso de regresiones no-lineales se ajustaron en cada mes funciones de varianza como la propuesta por Vanmarke (1983) (ec. (20A)). En el caso del mes de marzo los valores resultaron  $m = 1.108$  y  $\theta = 3.212$ . Como se muestra en la Figura 3 la estimación de la función de varianza observada es ajustada de manera adecuada por la función propuesta por Vanmarke (1983), resultando un coeficiente de correlación de 0.999 entre ambas funciones. En los diferentes meses del año no se detectó una sustancial diferencia en la función de varianza  $\gamma(T)$ . Por otro lado en lo concerniente a la varianza  $\sigma^2$  del proceso continuo  $Y(t)$  se determinaron valores con un máximo de  $2.50 \text{ (mm/h)}^2$  en los meses de mayores lluvias (verano) en tanto que se determinaron valores con un mínimo de  $0.11 \text{ (mm/h)}^2$  en los meses de escasas lluvias (invierno).

### Determinación de los parámetros del modelo de N-S

Para la minimización de la función objetivo (ec. 23A) fue utilizado el método de convergencia cuadrática de Powell (Press et al., 1986). En vista a contemplar la influencia de los puntos iniciales de búsqueda de la solución se consideró un dominio espacial  $R^5$  de los valores de los parámetros  $\lambda(1/h)$ ,  $\eta(1/h)$ ,  $\mu$ ,  $\xi \text{ (mm/h)}$  y  $\beta(1/h)$  (ver ecuaciones 1A a 5A). Se establecieron valores mínimos y máximos (Tabla 2) de cada uno de los parámetros en función de consideraciones físicas y datos de la literatura (Calenda y Napolitano, 1999). Los intervalos de cada parámetro fueron divididos en  $n=11$  subintervalos, determinándose  $n^5$  puntos de inicio de búsqueda del mínimo. Con las 8 agregaciones utilizadas (0.5; 1; 2;

3; 4; 6; 12 y 24 horas) quedaron definidos 28 combinaciones de pares de escalas (0.5-1; 0.5-2; 0.5-3; 0.5-4; 0.5-6; 0.5-12; 0.5-24; 1-2; 1-3; 1-4; 1-6; 1-12; 1-24; 2-3; 2-4; 2-6; 2-12; 2-24; 3-4; 3-6; 3-12; 3-24; 4-6; 4-12; 4-24; 6-12; 6-24; 12-24) generando a su vez 28 sistemas de ecuaciones no lineales para cada mes. En cada sistema de ecuaciones se comenzó desde  $11^5 = 161051$  condiciones iniciales diferentes en búsqueda de valores mínimos de la función objetivo.

Tabla 2. Rango de los valores iniciales de los parámetros

| Parámetro       | $\lambda(1/h)$ | $\eta(1/h)$ | $\mu$ | $\xi \text{ (mm/h)}$ | $\beta(1/h)$ |
|-----------------|----------------|-------------|-------|----------------------|--------------|
| Límite inferior | 0.001          | 0.100       | 2     | 0.100                | 0.010        |
| Límite superior | 0.050          | 20          | 100   | 30                   | 3.000        |

A los efectos de caracterizar las diferentes combinaciones de escalas de agregación utilizadas en la resolución del sistema de ecuaciones se define el índice de escala:

$$\varepsilon_{i,j} = \frac{\tau_j}{\tau_i} \quad (1)$$

En la Figura 2 se presenta la relación entre el índice de escala  $\varepsilon_{i,j}$  y el valor de la función objetivo  $Z$  (mes de Marzo) para las mejores soluciones en cada escala de agregación, considerando como mejores a aquellas soluciones con valores de la función objetivo  $Z < 1 \times 10^{-2}$ . En la Figura 2 puede observarse claramente que las combinaciones de escalas con menores valores de la función objetivo  $Z$  son aquellas en que la relación entre las escalas es mayor o igual a 6. Del análisis de todos los meses surge que las mejores soluciones se encuentran para relaciones de escalas superiores o iguales a 3. Calenda y Napolitano (1999) determinaron valores de las relaciones de escalas con menores valores de  $Z$ , aún mayores a los aquí expuestos y ensayan una respuesta a esa tendencia analizando la relación entre parámetros en las distintas escalas.

Tabla 3. Parámetros resultantes de la mejora del ajuste.

| Mes | $\tau_i \text{ (h)}$ | $\tau_j \text{ (h)}$ | $\lambda(1/h)$ | $\eta(1/h)$ | $\mu$    | $\xi \text{ (mm/h)}$ | $\beta(1/h)$ | $Z$      | $S \text{ (%)}$ | $\theta$ | $S(\gamma) \text{ (%)}$ |
|-----|----------------------|----------------------|----------------|-------------|----------|----------------------|--------------|----------|-----------------|----------|-------------------------|
| Ene | 0.5                  | 4                    | 0.010174       | 4.128838    | 10.06710 | 6.594752             | 0.456184     | 0.40E-23 | 17.04           | 1.341    | 20.50                   |
| Feb | 1                    | 24                   | 0.009951       | 1.438589    | 2.416999 | 7.018521             | 0.177796     | 0.74E-24 | 1.85            | 2.085    | 5.18                    |
| Mar | 3                    | 24                   | 0.007110       | 1.082894    | 3.779467 | 6.79748              | 0.162049     | 3.78E-24 | 1.35            | 3.173    | 1.36                    |
| Abr | 1                    | 12                   | 0.006882       | 1.885393    | 8.426320 | 5.341779             | 0.163642     | 2.09E-32 | 1.47            | 2.797    | 8.97                    |
| May | 0.5                  | 6                    | 0.005138       | 2.875743    | 7.726374 | 5.042089             | 0.414621     | 3.64E-24 | 3.65            | 1.611    | 13.86                   |
| Jun | 0.5                  | 24                   | 0.005102       | 1.159011    | 2.819853 | 3.521109             | 0.283863     | 1.65E-06 | 1.28            | 2.551    | 4.61                    |
| Jul | 0.5                  | 3                    | 0.003646       | 2.291419    | 5.489877 | 2.456174             | 0.583397     | 9.14E-08 | 26.80           | 1.568    | 11.58                   |
| Ago | 0.5                  | 12                   | 0.004076       | 2.251276    | 6.891461 | 2.057554             | 0.294686     | 8.45E-17 | 1.64            | 1.989    | 13.23                   |
| Set | 1                    | 24                   | 0.005394       | 1.089781    | 2.921599 | 2.410589             | 0.242612     | 3.02E-05 | 9.76            | 2.771    | 0.43                    |
| Oct | 0.5                  | 6                    | 0.008183       | 2.118620    | 9.801529 | 3.583404             | 0.283702     | 1.09E-23 | 12.65           | 2.473    | 11.64                   |
| Nov | 1                    | 12                   | 0.009903       | 1.251981    | 2.950291 | 6.456836             | 0.277236     | 7.22E-18 | 1.58            | 2.421    | 3.11                    |
| Dic | 3                    | 24                   | 0.014110       | 1.393174    | 3.557947 | 5.050017             | 0.254766     | 1.36E-23 | 11.01           | 2.360    | 1.62                    |

## EVALUACIÓN DEL MODELO DE NEYMAN-SCOTT PARA SIMULACIÓN DE LLUVIA EN UN PUNTO GEOGRÁFICO

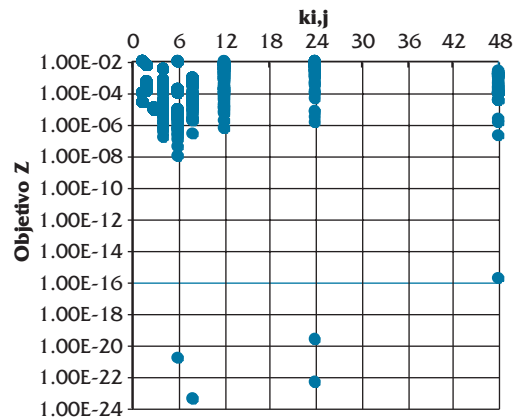


Figura 2. Función Objetivo Z vs Índice de Escala  $\epsilon_{i,j}$ , Mes de Marzo.

Dichos autores concluyen que cuando las escalas son cercanas, el comportamiento estadístico de los procesos es semejante con lo cual los parámetros no serían unívocamente determinados. Por el contrario si las escalas son bien diferenciadas, los valores de los parámetros no cambian de acuerdo al punto de inicio existiendo una única solución.

### Mejora del ajuste

La estimación de la mejor solución para el conjunto de las 8 escalas temporales (0.5, 1, 2, 3, 4, 6, 12 y 24 h) se lleva a cabo mediante un proceso de mejora del ajuste. A partir de la evidencia observada en cálculos preliminares acerca de que la mejor solución no necesariamente implica el valor mínimo absoluto de Z en la combinación de escalas respectiva, se evaluaron para la mejora, todas los grupos de soluciones con valor de  $Z \leq 0.01$ , considerando los parámetros obtenidos al resolver las 28 combinaciones de escalas.

El proceso de mejora del ajuste propiamente dicho consistió en una calificación ponderada simultánea de:

- a) Valor de la función Z (60% de participación en ponderación)
- b) Valor de la función S (ec. 24A) (25% de participación en ponderación)
- c) Valor de la función S aplicada a la diferencia entre la función varianza  $\gamma(T)$  de la muestra (ec. 20A) y  $\gamma(T)$  del modelo N-S (ec. 21A) (15% de participación en ponderación).

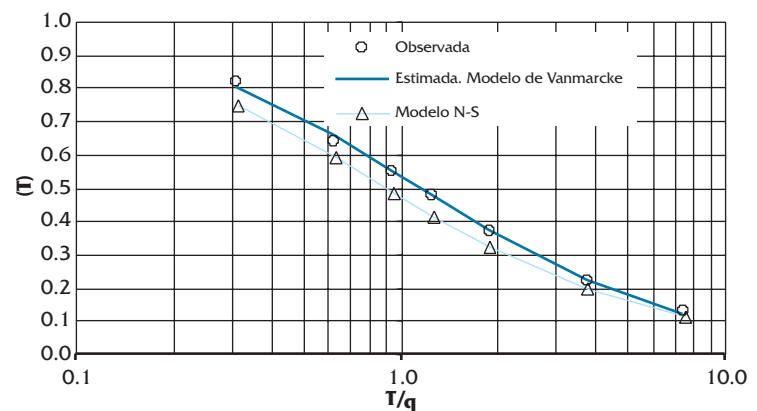


Figura 3. Función de varianza  $\gamma(T)$  Mes de Marzo.

El valor óptimo de la mejora del ajuste para el mes de marzo correspondió a un grupo de parámetros calculados en la combinación de escalas de 3 - 24 h, con un valor de función objetivo  $Z = 3.78E-24$ , función de mejora  $S = 1.35\%$  y  $S(\gamma(T)) = 1.36\%$ . En la Figura 3 se presentan la función de varianza  $\gamma(T)$  del modelo de pulsos rectangulares de Neyman-Scot, conjuntamente con los valores originados a partir de la muestra, establecidos a partir de la ec.(15A) y el modelo de función de varianza ajustado por ec. (21A). Los parámetros del modelo en todos los meses para la solución óptima se presentan en la Tabla 3. De acuerdo a valores de S hallados en la bibliografía (Calenda y Napolitano (1999), Velghe et al. (1994)) puede considerarse el nivel de ajuste como aceptable.

En los meses de enero y julio no se hallaron soluciones con valores de la función S inferiores a 17% y 26.8% respectivamente, lo que indica una menor bondad del modelo para la reproducción de lluvias en esos meses. Analizando término a término la sumatoria planteada por la ec. (24A) se comprobó que el alto valor de S se debe al fallo en el ajuste del lag-1 en la escala de 24 h. En los meses aludidos se produce en las lluvias observadas el mínimo valor positivo de lag-1 en la escala de 24 h (mes de enero) y el único valor negativo del lag-1 en la misma escala (mes de julio), ambos cercanos a 0. Cuando se intentó buscar soluciones para esos meses condicionando la mejora planteada en la ec. (24A) a las soluciones donde la diferencia entre los valores teóricos y observados no superaran el 50% del valor observado y no se encontró una solución que satisfaga esa condición para  $Z \leq 0.01$ .

**Generación de Series Sintéticas**

Una vez seleccionados los parámetros para cada mes, se procedió a generar las series sintéticas de lluvias aplicando el método de Monte Carlo a las distribuciones establecidas en las ecuaciones (1A) a (5A). Se determinaron 10 series de 12 años

de longitud temporal, definiéndose en cada caso la serie máxima, media y mínima en función del acumulado en la escala de 24 horas. Se calcularon los principales estadísticos para todas las escalas de agregación utilizadas en la determinación de los parámetros (0.5 a 24 h) y en las escalas adicionales de 48, 72 y 168 h.

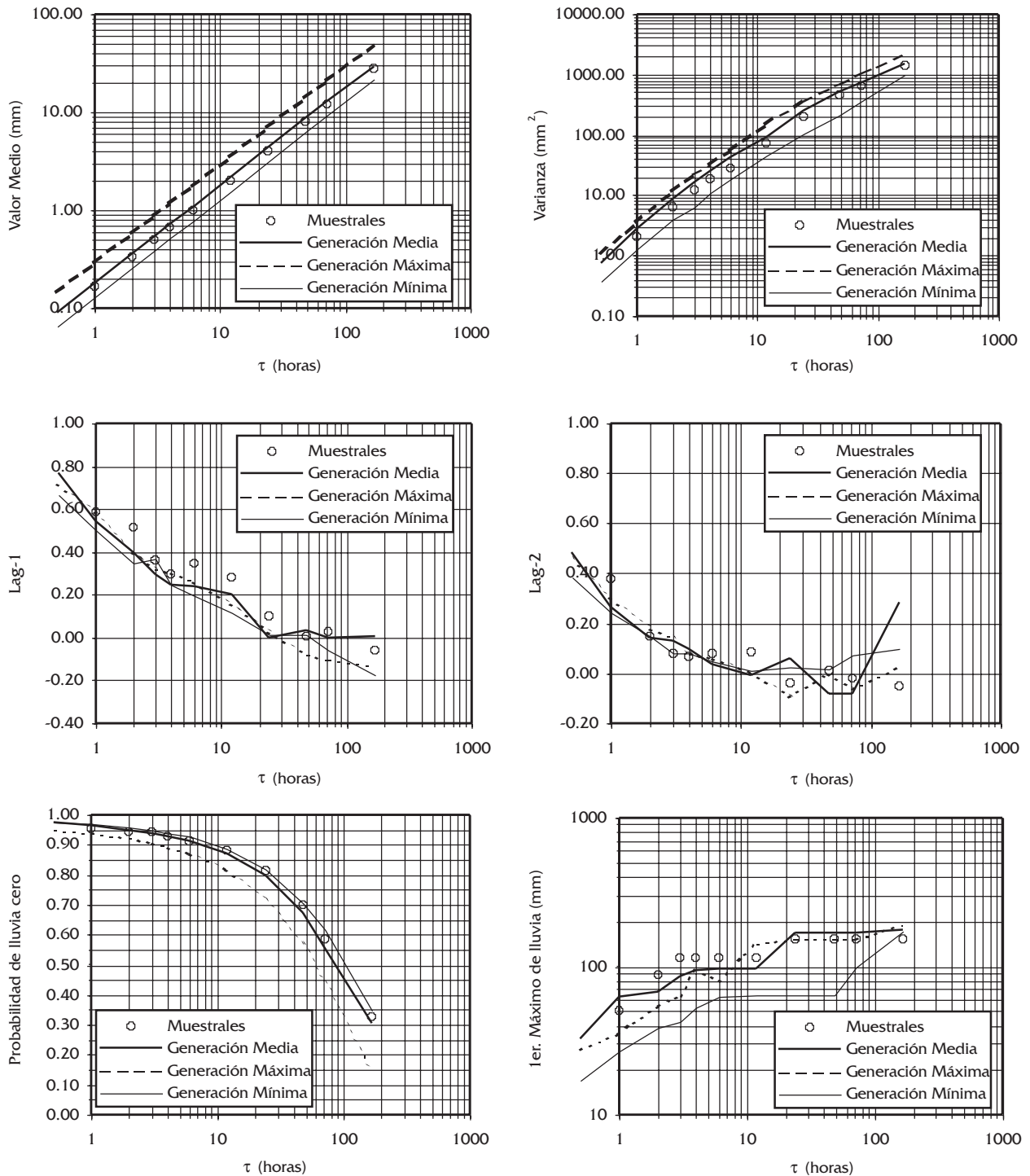


Figura 4. Comparación de los valores calculados y observados como función de la escala de agregación para el mes de Marzo. (a) Valor Medio; (b) Varianza; (c) Coeficiente de autocorrelación de orden 1; (d) Coeficientes de autocorrelación de orden 2; (e) Probabilidad de tiempo seco y (f) Primer máximo de lluvia

## EVALUACIÓN DEL MODELO DE NEYMAN-SCOTT PARA SIMULACIÓN DE LLUVIA EN UN PUNTO GEOGRÁFICO

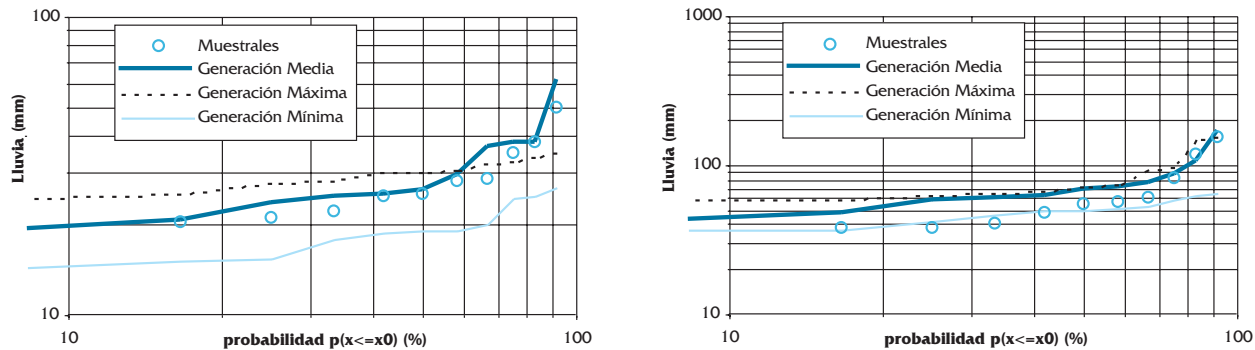


Figura 5. Probabilidad de no excedencia de los valores máximos calculados para el mes de Marzo. (a) Máximos en 1 hora y (b) Máximos en 24 horas.

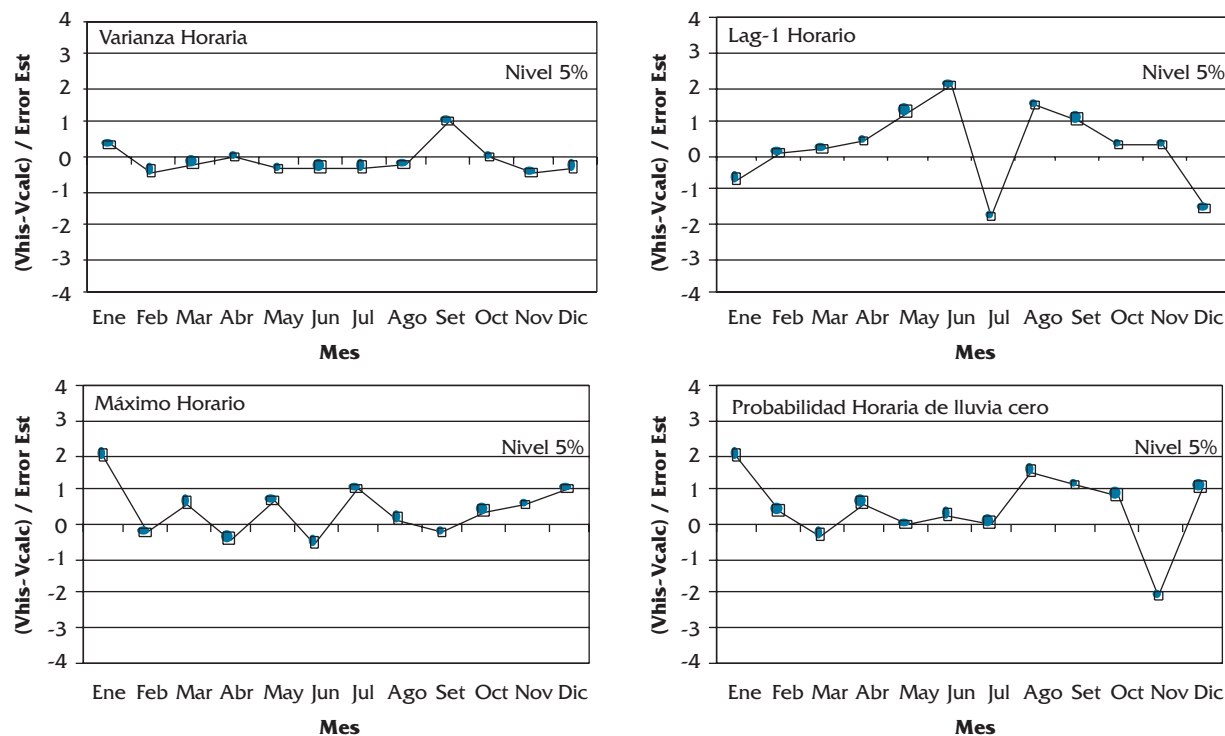


Figura 6. Test T para una agregación horaria de las series de tiempos. (a) Varianza; (b) Lag-1; (c) Valor máximo y (d) Probabilidad de lluvia cero. Vhis: valor histórico; Vcalc: valor calculado y Error Est: error estandar

Para cuantificar la calidad de los resultados calculados por el modelo se procedió a comparar estadísticos y variables características con los valores obtenidos desde las series observadas. La comparación incluyó valor medio, varianza, coeficientes de autocorrelación lag-1, lag-2 y lag-3 y probabilidad de lluvia cero. Asimismo se compararon variables características como máximos (1° a 11°) en cada escala de agregación y probabilidades de no excedencia de lluvia. Para evaluar globalmente la capacidad de simulación en los 12 meses se aplicó un Test T a la varianza, lag-1, probabilidad de lluvia cero y máximo horario, diario y semanal; en tanto que también se aplicó el test a valores medios y máximos mensuales.

Las Figuras 4 y 5 ilustran las comparaciones de valores calculados y observados para el mes de marzo. Las Figs. 4a y 4b corresponden a valor medio y varianza; las Figs. 4c y 4d corresponden a los coeficientes de autocorrelación de orden 1 y 2; en la Fig. 4e se representa la probabilidad de lluvia cero en tanto que el 1er. máximo de la serie generada se representa en la Fig. 4f. Las 5a y 5b corresponden a las probabilidades de no excedencia para lluvias de 1 y 24 horas.

En el caso del valor medio, la varianza y la probabilidad de lluvia cero se lograron resultados altamente satisfactorios computándose una correlación entre curvas observadas y calculadas (simulaciones mínima, media y máxima)  $R^2 > 0.99$ .

Es menos satisfactorio, sin dejar de ser aceptable, el ajuste para la estructura de correlación (lag-k). El lag-1 mostró una correlación superior  $R^2 > 0.90$  en tanto que el lag-2 de la simulación media fue mas pobre  $R^2 > 0.80$  ( $R^2 > 0.89$  para simulación máxima y  $R^2 > 0.77$  para simulación mínima).

Analizando los máximos de las series desde el 1er al 11mo máximo la correlación es superior al 0.85. En lo referente a la probabilidad de no excedencia las correlaciones entre valores calculados y muestrales fueron  $R^2 > 0.94$  para 1y 24 horas en tanto que desciende a 0.82 y 0.72 para 72 y 168 horas respectivamente.

En lo que concierne a la evaluación global en todos los meses del año se utilizó el Test T aplicado a variables horarias, diarias, semanales y mensuales respectivamente, evaluando la relación entre la diferencia entre el valor histórico y calculado respecto al error estandar del valor calculado. La mayor parte de los valores del test varían alrededor de la línea del cero indicando una capacidad satisfactoria del modelo en todos los meses del año. Se observan algunas diferencias por encima del nivel de significancia del 5% en el caso de la probabilidad diaria de lluvia cero en algunos meses de año. Sin embargo, las diferencias no se propagan a variables características a escala semanal y mensual. La Figura 6 ilustra el Test T aplicado a variables horarias.

## CONCLUSIONES

El modelo de agrupamiento de pulsos rectangulares de Neyman-Scott (N-S) ha demostrado una capacidad de generación de series sintéticas de lluvias puntuales con semejanza estadística a las series observadas. El modelo ha permitido aproximar muy aceptablemente propiedades de 2do orden y probabilidad de lluvia cero de las series generadas a las observadas en un rango de escalas temporales agregadas desde 0.50 hasta 168 horas (1 semana). En tanto que la aproximación es de menor calidad, sin dejar de ser aceptable, en lo que respecta a estructura de correlación (autocorrelación de orden 1 a 3) valores extremos y probabilidad de no excedencia de valores máximos

No se observó la tendencia reportada por otros autores (Entekhaby et al., 1989) del modelo N-S original, a sobreestimar la probabilidad de tiempo seco para grandes escalas temporales.

Las soluciones con menores valores de la función objetivo Z se encontraron para relaciones de escalas temporales superiores o iguales a 3.

Las soluciones con menor valor de la función de mejora S no necesariamente implicaron el mínimo absoluto de la función Z en la combinación de escalas respectiva. En los casos de soluciones con valores mínimos cercanos de Z y S resultó adecuado para la selección de la menor solución para cada mes, la utilización de un criterio adicional basado en la comparación entre las funciones de varianza observada y del modelo de N-S. Este criterio adicional aseguró la no selección de soluciones que si bien permitían valores claramente aceptables de Z y S, por otro lado establecían valores de los parámetros marcadamente sin significado físico. En los meses de enero y julio no se hallaron soluciones con valores de la función S inferiores a 17% y 26.8% respectivamente, lo que indica una menor bondad del modelo para la reproducción de lluvias en esos meses. Pudo comprobarse que el alto valor de S se debe a un pobre ajuste de los lag-1 teóricos en la escala de 24 h, cuando los valores observados son cercanos a 0.

En la evaluación del modelo a lo largo de todos los meses del año se observan una mayoría de valores del test T aplicado a diversas variables a escala horaria, diaria, semanal y mensual, por debajo del 5% de nivel de significancia.

Una de las potencialidades del presente modelo radica en la descripción matemática del proceso físico de lluvia y la posibilidad de su agregación temporal a tiempos superiores a los de la duración de la tormenta, lo que lo hace apto no sólo para estudios concernientes a transformación lluvia caudal, sino para su utilización en estudios de procesos del ciclo hidrológico como infiltración, balance, recarga de acuíferos, etc.

## APÉNDICE

### EL MODELO DE NEYMANN-SCOTT

#### Descripción y Ecuaciones Constitutivas

La formulación básica del proceso de N-S es :

- el tiempo entre orígenes de tormentas sucesivas es una variable independiente e idénticamente distribuida, según una distribución exponencial:



$$P_{Te}(te) = 1 - e^{-\lambda te} \quad (1A)$$

- el número de celdas o pulsos rectangulares de lluvia  $N$  asociado puede ser representado por una distribución de Poisson o por una distribución geométrica. En el caso de la distribución geométrica, para asegurar que al menos una celda es generada en cada evento lluvioso se asume un transformación de variable del tipo  $C = N-1$  con  $C$  con una distribución geométrica con valor medio  $E(C) = \mu-1$ , de modo que  $E(N) = \mu$ . La función densidad de probabilidad de  $C$  resulta:

$$p_C(c) = \frac{1}{\mu} \left(1 - \frac{1}{\mu}\right)^c \quad (2A)$$

- el tiempo de comienzo de cada celda de lluvia  $T_c$ , medida desde el origen del evento se considera distribuido exponencialmente con un parámetro  $\beta$ :

$$P_{Tc}(\tau_c) = 1 - e^{-\beta \tau_c} \quad (3A)$$

- el hietograma de cada celda es rectangular con duración  $L$  e intensidad  $X$ , exponencialmente distribuidas con parámetros  $\eta$  y  $\xi$  respectivamente:

$$P_L(l) = 1 - e^{-\eta l} \quad (4A)$$

$$P_X(x) = 1 - e^{-1/\xi x} \quad (5A)$$

La intensidad de precipitación para un cierto tiempo  $t$ ,  $Y(t)$  esta dada por la suma de los aportes individuales de cada celda activa en el tiempo  $t$ :

$$Y(t) = \int_{u=0}^{\infty} X_{t-u}(u) dN(t-u) \quad (6A)$$

donde  $X_{t-u}$  es la intensidad de los pulsos rectangulares disparados a partir del tiempo  $u$  y  $N(t)$  representa el proceso de conteo de las presentaciones de las celdas individuales.

Las expresiones analíticas de los estadísticos de 2do. orden (media, varianza y estructura de correlación) del proceso continuo  $Y(t)$ , fueron presentadas por Rodríguez-Iturbe et al. (1987a). Ya que los datos de lluvias están disponibles en intervalos de tiempo discretos (minutos, horas, días) es necesario determinar las propiedades agregadas para la estimación de los parámetros del modelo. El proceso agregado a una escala de tiempo  $\tau$  se define como:

$$Y_i^\tau = \int_{(i-1)\tau}^{i\tau} Y(t) dt \quad (7A)$$

donde  $i$  es el orden del intervalo de tiempo. Las propiedades de 2do. orden del proceso agregado, derivadas por Rodríguez-Iturbe et al. (1987a) son:

$$E[Y_i^{(\tau)}] = \frac{\lambda \mu \tau}{\eta \xi} \quad (8A)$$

$$\text{Var}[Y_i^{(\tau)}] = \frac{\lambda (\eta \tau - 1 + e^{-\eta \tau})}{\eta^3 \xi^2} \left[ 4\mu + 2(\mu^2 - \mu) \frac{\beta^2}{(\beta^2 - \eta^2)} \right] - \lambda(\beta \tau - 1 + e^{-\beta \tau}) \frac{2(\mu^2 - \mu)}{\xi^2 \beta (\beta^2 - \eta^2)} \quad (9A)$$

y para lags (o distancias temporales)  $k \geq 1$

$$\text{Cov}[Y_i^{(\tau)}, Y_{i+k}^{(\tau)}] = \frac{\lambda (1 - e^{-\eta \tau}) e^{-\eta(k-1)\tau}}{\eta^3 \xi^2} \left[ 2\mu + (\mu^2 - \mu) \frac{\beta^2}{\beta^2 - \eta^2} \right] - \lambda(1 - e^{-\beta \tau}) e^{-\beta(k-1)\tau} \frac{(\mu^2 - \mu)}{\xi^2 \beta (\beta^2 - \eta^2)} \quad (10A)$$

### Análisis de Fluctuaciones

La lluvia puntual es un proceso altamente fluctuante a través del tiempo. Por otro lado los hidrólogos se interesan mayormente por las características de las lluvias agregadas a determinados intervalos de tiempo  $Y_T(t)$  (intensidad media de la lluvia en un intervalo de tiempo  $T$ ), o en forma equivalente, por el proceso integrado  $T Y_T(t)$  (lámina de lluvia en un intervalo de tiempo  $T$ ).

$$Y_T(t) = \frac{1}{T} \int_{t-(T/2)}^{t+(T/2)} Y(t) dt \quad (11A)$$

$$I_T(t) = T Y_T(t) = \int_{t-(T/2)}^{t+(T/2)} Y(t) dt \quad (12A)$$

Surge de aquí, la cuestión de la magnitud del intervalo de tiempo  $T$  requerido para obtener estimaciones estables (de baja varianza) de los valores fluctuantes relacionados a  $Y_T(t)$ . Vanmarcke (1983) ha presentado una generalización al respecto aplicado a un caso multidimensional.

Una medida de la variabilidad de  $Y_T(t)$  es dada por su varianza:

$$\text{Var}[Y_T(t)] = \sigma_T^2 = \gamma(T) \sigma^2 \quad (13A)$$

donde  $\sigma^2$  es la varianza de  $Y(t)$  y  $\gamma(T)$  es la función de varianza de  $Y(t)$  definida por:

$$\gamma(T) = \frac{1}{T^2} \int_0^T \int_0^T \rho(t_1 - t_2) dt_1 dt_2 \quad (14A)$$

$$\gamma(T) = \frac{2}{T} \int_0^T \left(1 - \frac{t}{T}\right) \rho(t) dt \quad (15A)$$

donde  $\rho(t_1 - t_2)$  es la función de correlación de  $Y(t)$ . Las funciones aleatorias  $Y_T(t)$  e  $I_T(t)$  difieren solamente por  $T$ , y sus respectivas varianzas por  $T^2$  de modo que:

$$\text{Var}[I_T] = T^2 \sigma_T^2 = T^2 \gamma(T) \sigma^2 \quad (16A)$$

Vanmarcke (1983) define a la escala de fluctuación de  $Y(t)$  como:

$$\theta = \lim_{T \rightarrow \infty} [T \gamma(T)] \quad (17A)$$

por lo tanto cuando  $T \rightarrow \infty$

$$\gamma(T) = \frac{\theta}{T} \quad (18A)$$

Cuando la ec.(17) tiene un valor finito es demostrable que una definición equivalente para la escala de fluctuación es (Rodríguez-Iturbe, 1986):

$$\theta = \int_{-\infty}^{\infty} \rho(t) dt \quad (19A)$$

Debe tenerse en cuenta la importancia de  $\theta$  como una escala de tiempos que caracteriza la estructura de correlación de  $Y(t)$ .

Según Vanmarcke (1983) una familia útil de modelos para representar la función varianza de la muestra es descrita por:

$$\gamma(T) = \left[ 1 + \left( \frac{T}{\theta} \right)^m \right]^{-1/m} \quad (20A)$$

La escala de fluctuación  $\theta$  y el parámetro  $m$  pueden ser estimados por ajuste de una regresión no lineal a los valores obtenidos desde la ec.(15A) para las distintas escalas de agregación disponibles en la muestra.

Por otra parte, la función de varianza  $\gamma(T)$  y la escala de fluctuaciones  $\theta$  del modelo de pulsos rectangulares de Neyman-Scot pueden escribirse en función de los parámetros del modelo de la manera (Rodríguez- Iturbe, 1986):

$$\gamma(T) = \frac{(8\beta(\eta^2 - \beta^2)(\eta T - 1 + e^{-\eta T}) + 2\mu\eta^3(\beta T - 1 + e^{-\beta T}))}{(4\eta^2\beta(\eta^2 - \beta^2)T^2 + \mu\beta^2\eta^3T^2)} \quad (21A)$$

$$\theta = \frac{8(\eta^2 - \beta^2) + 2\mu\eta^2}{4\eta(\eta^2 - \beta^2) + \mu\beta\eta^2} \quad (22A)$$

## Estimación de parámetros

El procedimiento más habitual para la estimación de los parámetros es el método de los momentos (Rodríguez-Iturbe et al., 1987a, b). El modelo de pulsos rectangulares de N-S depende de los cinco parámetros  $\lambda$ ,  $\mu$ ,  $\beta$ ,  $\eta$  y  $\xi$ , por lo que siguiendo al método de los momentos deben ser calculados los cinco estadísticos de la serie observada e igualados a sus expresiones teóricas (ecuaciones (8A) a (10A)). El sistema de ecuaciones resultantes debe ser resuelto en función de los parámetros. El procedimiento aquí utilizado es el que propusiera originalmente Rodríguez-Iturbe et al. (1987a, b), en donde la serie histórica es agregada a diferentes escalas temporales, pudiendo comprender desde minutos a días. Del conjunto de escalas consideradas se van tomando pares de escalas  $\tau_i$  y  $\tau_j$  calculándose media, varianza y covarianza por medio de las ec. (8A) a (10A) en una escala y varianza y covarianza utilizando las ec. (9A) y (10A) para la otra. Cada par de escalas originará un sistema de ecuaciones de 5 x 5 no lineal, el cual se resuelve mediante la minimización de una función objetivo  $Z(x)$ , definida por la suma de los residuos normalizados entre los momentos teóricos  $\theta_k(x, \tau)$  y los momentos observados  $\theta_k^*(x, \tau)$  para una determinada escala de agregación seleccionada.

$$Z(x) = \sum_{k, \tau} \left[ \frac{\theta_k(x, \tau)}{\theta_k^*(\tau)} - 1 \right]^2 \quad (23A)$$

El uso de la relación (23A) asegura que los grandes valores numéricos no dominen el ajuste. Resulta evidente que en cada combinación de 2 escalas de agregación surgirá un conjunto de grupos de valores de los parámetros  $\lambda$ ,  $\mu$ ,  $\beta$ ,  $\eta$  y  $\xi$ , que minimizan la función  $Z(x)$  por debajo de un cierto umbral. Por ello, es necesaria una mejora del ajuste, donde se evalúe el ajuste global de cada grupo de parámetros en todas las escalas temporales consideradas. Una función de mejora del ajuste, utilizable para la evaluación de los estadísticos calculados a partir de cada grupo de parámetro en todas las escalas es (Calenda y Napolitano, 1999; Velghe et al., 1994):

$$S = \frac{1}{n} \left[ \sum_{l=1}^n \left( 1 - \frac{X_{\text{cal},l}}{X_{\text{his},l}} \right)^2 \right] \times 100 \quad (24A)$$

donde el índice  $l$  indica la media, varianza y lag-1 de todas las escalas,  $X_{\text{cal},l}$  es el valor de los estadísticos  $l$  obtenido usando los parámetros calculados, en tanto que  $X_{\text{his},l}$  es el valor histórico del estadístico.

## EVALUACIÓN DEL MODELO DE NEYMAN-SCOTT PARA SIMULACIÓN DE LLUVIA EN UN PUNTO GEOGRÁFICO

### LISTA DE SÍMBOLOS

|                         |   |
|-------------------------|---|
| $T_e$ :                 | tiempo entre orígenes de tormentas  |
| $P_{T_e}(te)$ :         | función distribución de probabilidad (fdp) de variable $T_e$                                    |
| $\lambda$ :             | parámetro de distribución exponencial de fdp de variable $T_e$                                  |
| $T_c$ :                 | tiempo de comienzo de cada celda de lluvia, medida desde el origen de la tormenta               |
| $P_{T_c}(te)$ :         | fdp de variable $T_c$   |
| $\beta$ :               | parámetro de distribución exponencial de fdp de variable $T_c$                                  |
| $L$ :                   | duración de cada celda de lluvia  |
| $P_L(l)$ :              | fdp de variable $L$   |
| $\eta$ :                | parámetro de distribución exponencial de fdp de variable $L$                                    |
| $X$ :                   | intensidad de cada celda de lluvia  |
| $P_X(x)$ :              | fdp de variable $X$   |
| $1/\xi$ :               | parámetro de distribución exponencial de fdp de variable $X$                                    |
| $N$ :                   | número de celdas de lluvia asociada a cada tormenta   |
| $C$ :                   | transformada de variable $N$ ( $C = N-1$ )  |
| $p_C(c)$ :              | función densidad de probabilidad de variable $C$  |
| $\mu$ :                 | parámetro de distribución geométrica de función de densidad de probabilidad de variable $C$     |
| $t$ :                   | coordenada temporal   |
| $Y(t)$ :                | intensidad de precipitación   |
| $X_{t-u}$ :             | intensidad de lluvia de los pulsos a partir del tiempo $u$                                      |
| $M(t)$ :                | conteo estocástico de las presentaciones de celdas individuales de lluvia                       |
| $Y_i^\tau$ :            | intensidad agregada a una escala de tiempo $\tau$ en el intervalo temporal $[(i-1)\tau, i\tau]$ |
| $E[Y_i^{(\tau)}]$ :     | valor medio de la intensidad agregada a un intervalo de tiempo $\tau$                           |
| $Var[Y_i^{(\tau)}]$ :   | varianza de la intensidad agregada a un intervalo de tiempo $\tau$                              |
| $Covar[Y_i^{(\tau)}]$ : | covarianza de la intensidad agregada a un intervalo de tiempo $\tau$                            |
| $Y_T(t)$ :              | función intensidad media agregada en un intervalo de tiempo $T$                                 |
| $I_T(t)$ :              | función lámina de lluvia agregada en un intervalo de tiempo $T$                                 |
| $\sigma^2$ :            | varianza de la función $Y(t)$   |
| $\sigma_T^2$ :          | varianza de la función $Y_T(t)$   |
| $\gamma(T)$ :           | función de varianza de $Y(t)$   |
| $\theta$ :              | escala de fluctuación de $Y(t)$   |
| $\rho(t)$ :             | función de correlación de $Y(t)$  |
| $m$ :                   | parámetro de regresión no lineal de modelo de función de varianza                               |
| $Z(x)$ :                | función objetivo  |
| $\theta_k(x, \tau)$ :   | momentos teóricos   |
| $\theta_k^*(x, \tau)$ : | momentos observados   |

|                    |   |
|--------------------|---|
| $S$ :              | función de evaluación de ajuste entre valores calculados y observados |
| $X_{cal,i}$ :      | valor calculado   |
| $X_{his,i}$ :      | valor histórico   |
| $\epsilon_{i,j}$ : | índice de escala  |
| $R^2$ :            | correlación entre curvas observadas y calculadas                      |

### REFERENCIAS

- DGHyS (1999), Serie Temporal de lluvias en Rosario-Aero 1986-1998. Dirección General de Hidráulica y Saneamiento de la Municipalidad de Rosario, Rosario, Argentina.
- Calenda, G. and F. Napolitano (1999), Parameter estimation of Neyman-Scott processes for temporal point rainfall simulation, *Journal of Hydrology*, Vol. 225, pp. 45-66.
- Cowpervait, P.S.P.(1991), Further Developments of the Neyman-Scott clustered point process for modeling rainfall, *Water Resources Research*, Vol. 27 (7), pp. 1431-1438.
- Cox, D.R. and V. Isham, V. (1980), *Point Processes*, Chapman and Hall, London.
- García-Bartual, R. (2002), Aplicaciones en Gestión de Crecidas de la Modelación Estocástica de la Precipitación, *Jornadas sobre Sistemas de Ayuda a la Decisión ante Problemas Hidráulicos e Hidrológicos en Tiempo Real*, CEDEX, Madrid, pp. 183-206.
- Entekhabi, D., Rodriguez-Iturbe, I. and P.S. Eagleson (1989), Probabilistic Representation of the Temporal Rainfall Process by a Modified Neyman-Scott Rectangular Pulses Model: Parameter Estimation and Validation, *Water Resources Research*, Vol. 25 (2), pp. 295-302.
- Kavvas, M.L. and J. W. Delleur (1981), A stochastic model of daily rainfall sequences, *Water Resources Research*, Vol. 17 (4), pp. 1151-1160.
- Marien, J.L. and G.L. Vandewiele (1986), A Point Rainfall Generator with Internal Storm Structure, *Water Resources Research*, Vol. 22 (4), pp. 475-482.
- Onof, C. and H.S. Wheater (1993), Modelling of British rainfall using a random parameter Barlet-Lewis rectangular pulse model, *Journal of Hydrology*, Vol. 149, pp. 67-95.
- Press W.H., B. Flannery, S. Teukolsky and W. Vetterling (1986), *Numerical Recipes*, Cambridge University Press, New York, pp. 297-300.
- Rodriguez-Iturbe, I. (1986), Scale of fluctuations of rainfall models, *Water Resources Research*, Vol. 22 (9), 15S-37S.
- Rodriguez-Iturbe, I., V.K. Gupta, and E. Waymire (1984), Scale considerations in the modelling of temporal rainfall, *Water Resources Research*, Vol. 20 (11), pp. 1611-1619.

- Rodriguez-Iturbe, I. and P.S. Eagleson (1987a), Mathematical models of rainstorm events in space and time, *Water Resources Research*, Vol. 23 (1), pp. 181-190.
- Rodriguez-Iturbe, I., D.R. Cox and V. Isham (1987b), Some models for rainfall based on stochastic point processes, *Proc. R. Soc. Lond., Series A* (410), pp. 269-288.
- Rodriguez-Iturbe, I., B. Febres de Power and J. B. Valdés (1987c), Rectangular Pulses Point Process Models for Rainfall: Analysis of Empirical Data, *Journal of Geophysical Research*, Vol. 92 (D8), pp. 9645-9656.
- Rodriguez-Iturbe, I., D.R. Cox and V. Isham (1988) A Point Process for Rainfall: further developments. *Proc. R. Soc. Lond., Series A* (417), pp. 283-298.
- Salsón, S. y B. Garcia-Bartual (1998), Desagregación de lluvias para aplicaciones en simulación de sistemas de recursos hidráulicos, *Revista de Obras Públicas*, Jul-Ago1998, n° 3378, pp. 25-35.
- Salsón, S. (2001), Tesis Doctoral Modelo Matemático para Simulación Numérica Espacio-Temporal de Intensidades de Lluvia en Episodios Torrenciales de carácter Convectivo, Universidad Politécnica de Valencia.
- Valdes, J.B., I. Rodriguez-Iturbe and V.J. Gupta (1985) Approximations of temporal rainfall model from a multidimensional model. *Water Resources Research*, Vol. 21 (8), pp. 1259-1270.
- Vanmarcke, E. (1983) *Random Fields: Analysis and Synthesis*, MIT Press, Cambridge, MA, 382 pp.
- Velghe T., P. Troch, F. De Troch and J. Van de Velde (1994), Evaluation of cluster-based rectangular pulses point process models for rainfall, *Water Resources Research*, Vol. 30 (10), pp. 2847-2857.
- Waymire, E. and V.K. Gupta (1981), The mathematical structure of rainfall representations: 1. A review of the stochastic rainfall models. 2. A review of the theory of point processes. 3. Some applications of the point process theory to rainfall processes, *Water Resources Research*, Vol. 17 (5), pp.1261-1294.
- Willems, P. (1999), Stochastic generation of spatial rainfall for urban areas, *Water Science and Technology*, Vol. 39(9), pp. 23-30.
- Willems, P. (2001), A spatial rainfall generator for small spatial scales, *Journal of Hydrology*, Vol. 252, pp. 126-144.
- Zimmermann E., y Silber M. (2002), Algoritmo para la Representación de Campos de Precipitación”, XIX Congreso Nacional del Agua, T54, Villa Carlos Paz, Córdoba, Argentina.