

# APLICACIÓN DE AUTÓMATAS CELULARES PARA LA DESCRIPCIÓN DE EXTREMOS DE LA PRECIPITACIÓN

A.H. ENCINAS\*; A. MARTÍN DEL REY\*\*; C. RODRÍGUEZ-PUEBLA\*\*\*;  
O. GARCÍA DELGADO\* y G. RODRÍGUEZ SÁNCHEZ\*\*\*\*

\* *Departamento de Matemática Aplicada, E.T.S.I.I. de Béjar*

\*\* *Departamento de Matemática Aplicada, E.P.S. de Ávila*

\*\*\* *Departamento de Física General y de la Atmósfera, Facultad de Ciencias Físicas*

\*\*\*\* *Departamento de Matemática Aplicada, E.P.S. de Zamora  
Universidad de Salamanca*

## RESUMEN

El objetivo de este estudio es comprobar la utilidad de la técnica de Autómatas Celulares (CA) para describir la ocurrencia de extremos de precipitación. Las series temporales de precipitación presentan algunos grados de periodicidad o fluctuaciones debidas a variabilidad climática natural. Una vez filtradas estas componentes mediante técnicas espectrales, aplicamos el modelo propuesto de CA a la serie residual con el fin de caracterizar las rachas secas y húmedas. Este método va a simular dos estados para excesos de precipitación, el estado 0 se corresponde con precipitaciones inferiores al percentil 75, mientras que el estado 1 para las superiores al percentil 75. De forma análoga procedemos para caracterizar rachas de precipitación escasa, consideraremos estado 0 para precipitaciones superiores o igual al percentil 25 y estado 1 para precipitaciones inferiores a dicho percentil. El método lo aplicamos a datos mensuales de precipitación en Salamanca en el periodo 1949 a 2003. Los primeros resultados indican que se consigue aproximar en un 60% la ocurrencia de extremos de precipitación, lo cuál sería de gran interés por su aplicación en planificaciones agrícolas.

**Palabras clave:** Autómata celular, precipitación, variabilidad climática.

## ABSTRACT

*The objective of this study is to show the utility of using Cellular Automata (CA) to describe the occurrence of the precipitation extremes. Since the precipitation time series contain some quasi-periodic fluctuations, it is required to filter these components before applying the CA model to characterize the dry and wet spells. The CA used to describe the wet spells consists of two states 0 and 1, which were established according to the threshold of 75 percentile. Similarly, to describe the dry spells we propose a CA with two states 0 and 1 for values higher and lower than 25 percentile respectively. The method is applied to monthly precipitation data of Salamanca (Spain) for the period 1949 to 2003. The first results indicate that the model can capture the occurrence of the precipitation extremes with an 60% of accuracy, therefore this characterization could be useful for agricultural planning.*

**Key words:** Cellular automata, climatic variability, precipitation.

## 1. INTRODUCCIÓN

Las observaciones de precipitación muestran una gran variabilidad, con periodos de sequía y de inundaciones que conllevan consecuencias muy perjudiciales para la agricultura. Numerosos trabajos tratan de modelizar y predecir el comportamiento de la precipitación con resultados no

satisfactorios para la ocurrencia de los extremos pluviométricos. Por otra parte, la distribución frecuencial de los datos de precipitación nos indica que además de cambiar los valores medios se está produciendo un cambio de variabilidad (PALMER y RAISANEN, 2002) relacionan estos cambios con la intensificación de los gases de efecto invernadero.

En este trabajo analizamos la ocurrencia de extremos de precipitación mediante la técnica estadística de Autómatas Celulares (CA). Se obtiene un modelo de dos autómatas que considera la influencia de los dos años anteriores y la influencia de los dos meses anteriores sobre unos datos bimodales clasificados de acuerdo a determinados niveles o percentiles. Los estudios SAUNDERS y GHIL (2001), CUADRA y FERNÁNDEZ (2001), nos ha sugerido la posibilidad de aplicar esta técnica para la caracterización de los extremos climáticos.

Las series de precipitación contienen variaciones estacionales e interanuales cuyo origen está en la propia naturaleza del clima y que debemos filtrar previamente a la aplicación del método CA, las fluctuaciones cuasi-periódicas pueden considerarse como componentes conservativas del clima. Las series residuales deben incluir los extremos climáticos cuyas causas pudieran ser los forzamientos climáticos internos y se pueden asociar con componentes disipativas del clima.

Vamos a aplicar el método a las series de precipitación mensual de Salamanca desde 1949 a 2003. El tratamiento previo exige la identificación de oscilaciones cuasi-estacionarias mediante métodos espectrales (RODRÍGUEZ-PUEBLA *et al.*, 1998, 2001) así como la filtración de las componentes significativas mediante procedimientos no lineales.

## 2. DATOS

Las series temporales utilizadas son las observaciones instrumentales de precipitación acumulada mensualmente en el periodo 1949 a 2003 de Salamanca, estos datos han sido proporcionados por el Instituto Nacional de Meteorología de España.

Como hay 55 años de estudio y 12 meses para cada año, se dispone de una matriz real de dimensión 12 x 55.

Para poder comparar los datos de los distintos meses se normalizan, calculando el valor medio (VM) y la desviación típica ( $\sigma$ ) de cada mes en el periodo de estudio.

Tabla 1. Valores medios y desviaciones típicas de los distintos meses de precipitación en Salamanca para el periodo 1949-2003

Precipitación	En	Fb	Mz	Ab	My	Jn	Jl	Ag	Sp	Oc	Nv	Dc
VM	35,34	31,48	29,61	33,04	41,91	31,55	17,12	10,91	33,19	41,13	45,13	38,81
$\sigma$	26,68	25,67	20,91	21,05	22,11	22,77	19,23	12,38	24,44	33,51	38,09	31,22

Posteriormente, a los datos iniciales les restamos el valor medio y los dividimos por la desviación típica de ese mes. De este modo, obtenemos la matriz con los valores normalizados (RE).

A las series mensuales normalizadas se les aplica el análisis espectral. En general hemos identificado oscilaciones significativas entorno a: cuasi-bienal, cuasi-cuadrinal, decadal y tendencia. Las cuasi-oscilaciones se filtran para cada mes y posteriormente se unen las series residuales de todos los meses para obtener una matriz de trabajo (*RS*).

Los datos de la matriz residual (*RS*) se clasifican de acuerdo con el nivel del percentil 75 (también llamado tercer cuartil) para identificar aquellos meses cuyo valor de precipitación es abundante. También se aplica el estudio a los extremos secos, clasificando los datos de acuerdo con el nivel del percentil 25 (también llamado primer cuartil), para identificar aquellos meses cuyo valor de precipitación es escasa.

La matriz residual (*RS*) se transforma en dos matrices *D* de ceros y unos para valores superiores, inferiores o iguales a los percentiles 75 y 25 según el siguiente criterio: para exceso de precipitación se asigna el valor 0 si el dato es inferior o igual al percentil 75 y 1 si es superior a dicho percentil. Sea  $a_i^{(t)}$  el estado del mes *i*-ésimo en el instante *t* (v.g. precipitación en marzo para 1954), si  $d_i^{(t)} = 0$ , entonces en dicho mes hay precipitación normal o baja y si  $d_i^{(t)} = 1$  entonces la precipitación es alta; para el defecto de precipitación se asigna el valor 0 si el dato es superior al percentil 25 y 1 si es inferior o igual a dicho percentil. Sea  $d_i^{(t)}$  el estado del mes *i*-ésimo en el instante *t* (v.g. precipitación en marzo para 1954), si  $d_i^{(t)} = 0$ , entonces en dicho mes hay precipitación normal o alta y si  $d_i^{(t)} = 1$  entonces la precipitación es baja.

Si consideramos el percentil 75, la matriz *D75* asociada a las precipitaciones abundantes es:

$$d_i^{(t)} = \begin{cases} 0 & \text{si } a_i^{(t)} \leq P_i^{75} \\ 1 & \text{si } a_i^{(t)} > P_i^{75} \end{cases}$$

Donde  $a_i^{(t)}$  son los elementos de la matriz *RS* y  $d_i^{(t)}$  los de la matriz *D75* (Fig. 1)

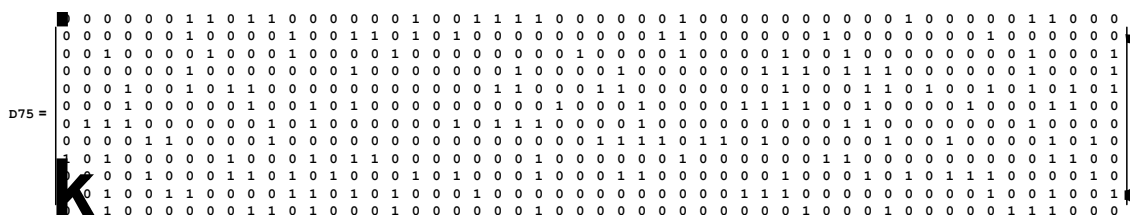


Fig. 1. Matriz booleana de datos de precipitación abundante

De forma análoga, si consideramos el percentil 25, la matriz *D25* (Fig. 2) asociada a las precipitaciones escasas es:

$$d_i^{(t)} = \begin{cases} 0 & \text{si } a_i^{(t)} \geq P_i^{25} \\ 1 & \text{si } a_i^{(t)} < P_i^{25} \end{cases}$$

Y la matriz de datos que se obtiene es la siguiente:

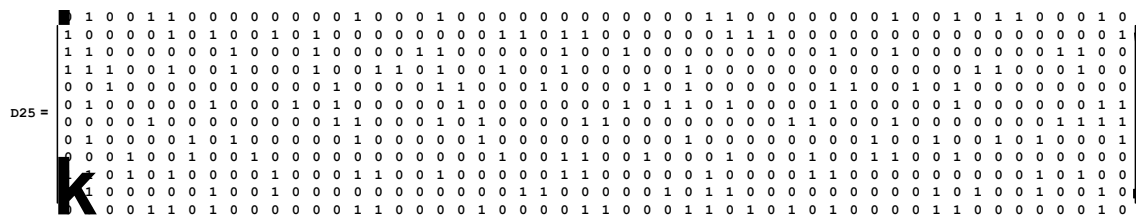


Fig. 2. Matriz booleana de datos de precipitación escasa

Sobre estas dos matrices se aplican los autómatas celulares.

### 3. MÉTODO

Podemos diferenciar dos etapas: la primera consiste en un pre-proceso de los datos mediante la identificación y posterior filtración de las variaciones naturales por métodos espectrales; la segunda consiste en la modelización de los extremos de precipitación mediante la aplicación del método de autómatas celulares.

#### 3.1. Análisis espectral

La mayoría de las irregularidades observadas en las series temporales podrían ser generadas mediante la iteración no lineal de algunas componentes cuasi-periódicas (CHATFIELD, 1980). Así la serie temporal  $X(t)$ , se puede considerar como la suma de una función sinusoidal de frecuencia  $f$  y amplitud  $B$ , más un término debido al ruido ( $\mu(t)$ ):

$$X(t) = Be^{2\pi ift} + \mu(t)$$

De esta forma, una serie temporal se puede descomponer o caracterizar por su densidad espectral. Existen numerosas técnicas de obtener espectros (GHIL and YIOU, 1996). En este estudio aplicamos la transformada rápida de Fourier de la función de autocorrelación. Las fluctuaciones significativas se identifican mediante la comparación con el espectro de la serie autorregresiva para desfase 1 adaptado a los datos.

Mostramos un ejemplo de aplicación del método espectral con las series de precipitación acumuladas anualmente. La figura 3 muestra la serie inicial y la reconstruida con las cuasi-oscilaciones de 6 y 3,3 años.

La figura 4 muestra la densidad espectral de la serie de precipitación anual, en la cual identificamos las fluctuaciones significativas por comparación con el ruido espectral o nivel de significación considerado (RODRÍGUEZ *et al.*, 1992).

Hemos aplicado el método a las series mensuales de precipitación. A partir de la reconstrucción y posterior filtración obtenemos las series residuales que supuestamente contienen las anomalías extremas de precipitación. Transformamos estas series en las dos nuevas matrices  $D75$  y  $D25$  mencionadas en la sección 2, sobre las que se aplica el autómata celular.

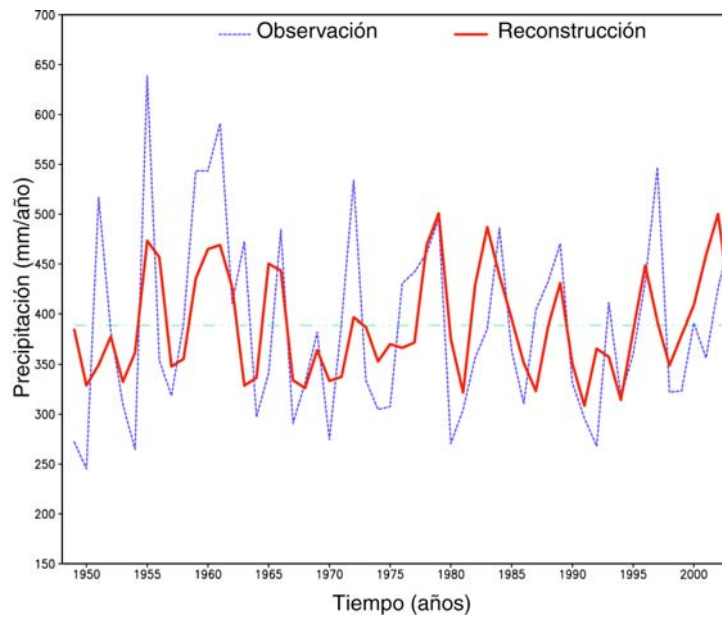


Fig. 3. Evolución de la precipitación anual de Salamanca

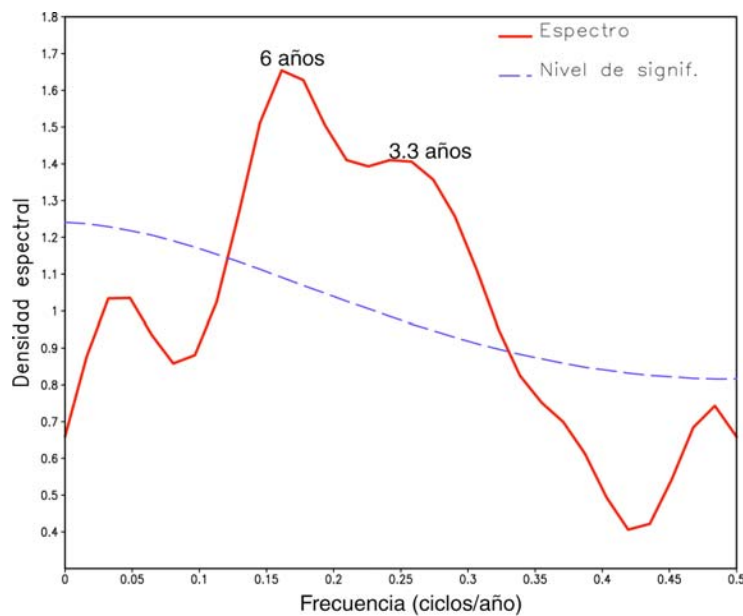


Fig. 4. Espectro de la serie anual de precipitación de Salamanca

### 3.2. Autómatas celulares

Se propone un modelo basado en autómatas celulares unidimensionales híbridos para el estudio y predicción de las precipitaciones extremas de un determinado año, mes a mes.

En general un autómata celular (CA) unidimensional o lineal, es un conjunto de 4 elementos  $A = (I, S, V, f)$ , donde  $I$  es el espacio celular formado por una cadena lineal de  $n$  componentes

denominadas células o celdas, que se denotan por  $\langle i \rangle$ ,  $0 \leq i \leq n-1$ .  $S$  es el conjunto finito de estados y sus elementos son todos los posibles valores que pueden tomar las células, generalmente se considera  $S = \wedge_k V \subset \wedge$  es un conjunto ordenado finito, denominado el conjunto de índices de  $I$ , de tal forma que para cada célula  $\langle i \rangle \in I$ , su vecindad  $V_{\langle i \rangle}$  es un conjunto ordenado de  $m$  células definidas por:  $V_{\langle i \rangle} = \{\langle i+\alpha_1 \rangle, \dots, \langle i+\alpha_m \rangle : \alpha_i \in V\}$ . Si el valor de la célula  $\langle i \rangle$  para el instante de tiempo  $t$  es  $a_i^{(t)}$ , entonces la evolución en el tiempo de esta célula viene dada por la función de transición local  $f: S^m \rightarrow S$ , que depende de los valores de la células que aparezcan en  $V_{\langle i \rangle}$ , esto es  $a_i^{(t+1)} = f(a_{i+\alpha_1}^{(t)}, \dots, a_{i+\alpha_m}^{(t)})$ . Finalmente, como el espacio celular es finito, se deben establecer condiciones de contorno. Por otra parte, se denomina configuración en el instante  $t$  a  $C^{(t)} = (a_0^{(t)}, \dots, a_{n-1}^{(t)}) \in S \times S \times \dots \times S$ , de tal forma que para  $t = 0$ ,  $C^{(0)}$  es la configuración inicial.

En nuestro estudio en particular se han utilizado dos autómatas, uno probabilístico y otro determinístico, y tendrán las siguientes características: el espacio celular  $I$  estará formado por 12 células representando cada una de ellas un mes del año. El conjunto de estados  $S$  será  $\wedge_2 = \{0, 1\}$ . La vecindad  $V_{\langle i \rangle}$  de cada célula  $\langle i \rangle$ , es diferente en cada uno de los autómatas utilizados como se verá más adelante. La función de transición es  $F = (f_i)$ , con  $0 \leq i \leq n-1$ . Donde  $f_i$  es la función de transición local asociada a cada célula  $\langle i \rangle$  y que consta de dos partes: una de carácter probabilístico, obtenida a partir de los datos reales y otra determinista que depende de las vecindades de cada célula. La regla de transición que rige el CA está formada por dos pasos que pasamos a detallar.

Sea  $D = (d_i^{(t)})$  con  $1 \leq i \leq 12$ ,  $1 \leq t \leq 55$ , y  $d_{ij} \in \{0, 1\}$ , la matriz de datos obtenida experimentalmente, como se indica en el apartado 2.

***Fase probabilista:*** Se construye basándose en las probabilidades de transición de estados, encontradas empíricamente con los datos de las series reales. Estas probabilidades dependen de la configuración de las vecindades en el tiempo  $t_n$ . En esta etapa, el estado de un mes determinado estará influenciado por los estados del mismo mes en los dos años inmediatamente anteriores. De forma gráfica podríamos expresarlo de la siguiente manera, donde  $P_{(1,0)}^0$  es la probabilidad de que se pase de 1 a 0 cuando la suma de vecindades es 0.

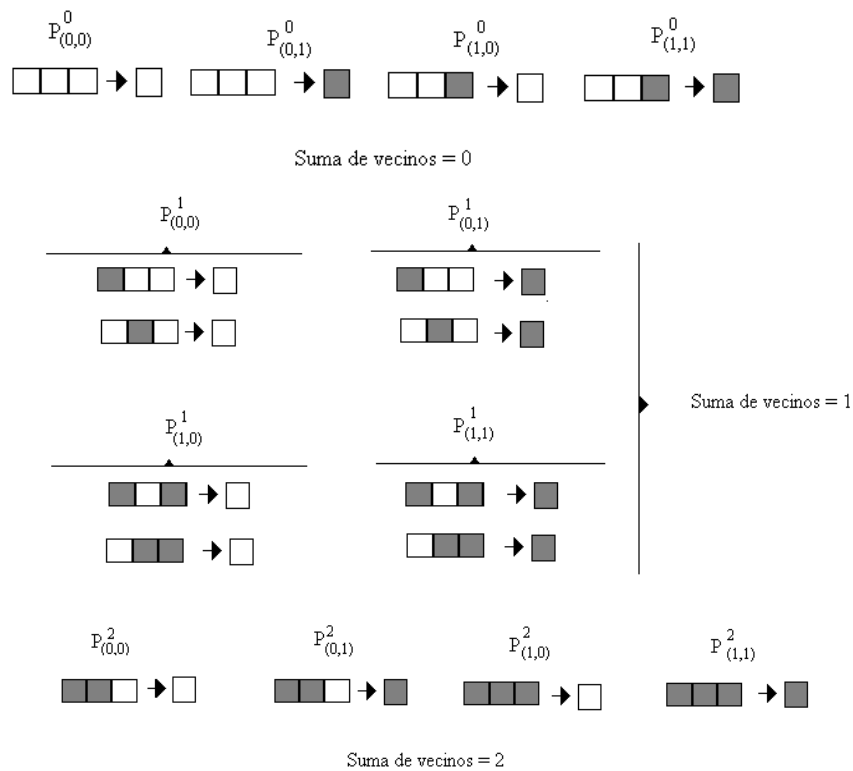


Fig. 5. Configuraciones de las celdas centrales para el cálculo de la matriz de transición (0 se representa en blanco y 1 en gris)

**Fase determinista:** De la fase anterior obtenemos como resultado una matriz que llamaremos  $(\beta_i^{(t)})$ , con  $1 \leq i \leq 12, 1 \leq t \leq 55$ , que es la simulación efectuada si sólo se aplicara dicha fase (la probabilista). En esta segunda fase el estado de un mes determinado estará influenciado por los estados de los dos meses anteriores dentro del mismo año. Así construimos una matriz  $(a_i^{(t)})$  donde:

$$a_i^{(t)} = \begin{cases} 0, & \text{si } \beta_i^{(t)} = 0 \text{ y } \beta_{i-1}^{(t)} + \beta_{i-2}^{(t)} \neq 2 \\ 1, & \text{en otro caso} \end{cases}$$

Las condiciones de contorno consideradas son: delante de la célula 0 y detrás de la  $n-1$  tendremos en cuenta los valores de las filas  $n-1, n-2$  y 0, 1 respectivamente, con lo que, para enero de 1949 tendremos en cuenta la media de diciembre y noviembre de todos los años. La configuración inicial considerada es la obtenida de los datos reales.

Una vez aplicados los dos autómatas obtenemos dos matrices  $RC$ , una para cada percentil:  $RC75$  y  $RC25$ . Comparamos estas matrices con las matrices  $D75$  y  $D25$  respectivamente y se realiza el recuento de coincidencias.

#### 4. RESULTADOS

La determinación de las secuencias del autómata considerado se ha llevado a cabo con el programa Mathematica<sup>TM</sup>.

La configuración inicial para la utilización del autómata celular en cada caso ha sido la obtenida a partir de los datos reales. A partir de ella se van generando los periodos restantes mediante la regla de transición de dos pasos del CA explicada en el apartado 3.2. En cada iteración se realiza el primer paso, que depende del valor de la matriz de transición de estados respectiva a todas las celdas del CA y luego un segundo paso que depende de la suma de vecindades generadas en el primero.

Se usa un valor percentílico: el 75 para determinar aquellas células cuyo valor es de precipitación alta, y un valor percentílico de 25 para determinar aquellas células cuyo valor es de precipitación baja.

Para cada una de estas dos simulaciones obtenemos las matrices calculadas RC75 y RC25 que se contrastan con las de partida D75 y D25. Obtenemos las discrepancias entre valores simulados y observados para cada año lo cual se muestra en las figuras 6 y 7 para rachas húmedas y secas respectivamente. Los resultados proporcionan un acuerdo superior al 60%, por lo que la utilización de autómatas celulares proporcionan una representación aceptable del comportamiento disipativo de las series temporales.

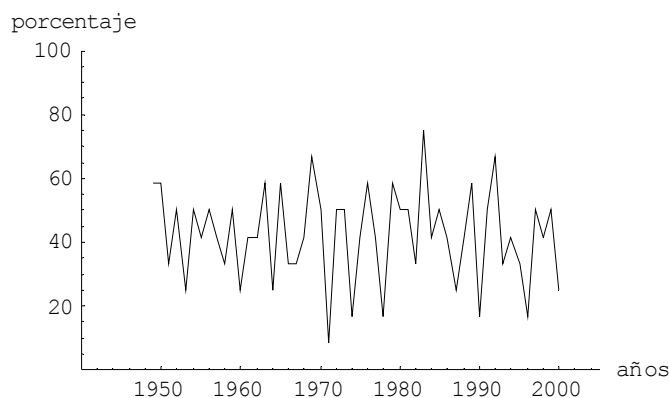


Fig. 6. Porcentajes de discrepancias anuales entre los extremos observados y simulados para el percentil 75



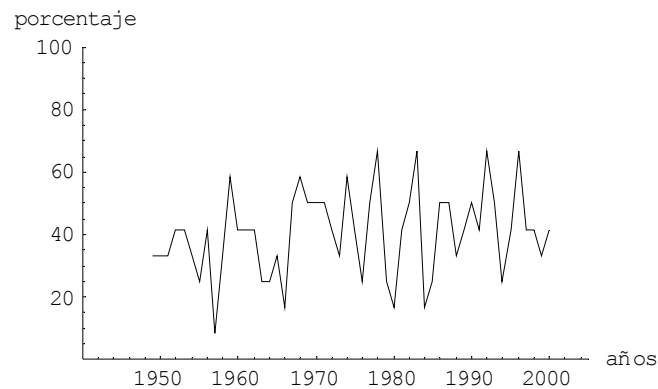


Fig. 7. Porcentajes de discrepancias anuales entre los extremos observados y simulados para el percentil 25

## 5. CONCLUSIONES

Realizamos un estudio sobre la utilidad de aplicar la técnica CA para modelizar la ocurrencia de extremos de precipitación en Salamanca. El modelo discreto que proponemos describe las anomalías húmedas y secas de acuerdo con dos estados establecidos a las series de precipitación que contienen la variabilidad extrema.

En todas las iteraciones del autómata celular, el tanto por ciento de aciertos supera el 60%, por lo que se puede suponer que el algoritmo da una aproximación aceptable de los datos reales.

Es necesario continuar esta investigación en las siguientes direcciones: por una parte, hay que combinar la descripción de los extremos proporcionada por los CA a la descripción obtenida con los métodos espectrales para su aplicación en estudios de predicción estacional; además, debemos proporcionar justificaciones físicas de las variaciones encontradas en las series de precipitación.

## 6. AGRADECIMIENTOS

Este trabajo se ha realizado con la subvención del proyecto “Fundación Memoria de D. Manuel Solórzano Barruso” y del proyecto “REN2002-04584-CO-02/CLI” del MCYT. Agradecemos al Instituto de Meteorología de España por los datos climáticos.

## 7. REFERENCIAS

- CUADRA, M. y FERNÁNDEZ B. (2001). Modelación espacio-temporal de sequías regionales mediante autómatas celulares. En: *Actas del XV Congreso Chileno de Hidráulica*. Concepción (Chile).
- CHATFIELD, C. (1980). *The analysis of time series: an introduction*. Chapman and Hall, London.
- GHIL, M. and YIOU P. (1996). *Spectral methods: what they can and cannot do for climatic time series*. NATO ASI Series, vol. I 44, Decadal climate variability. Dynamics and predictability. ED. D.L.T. Anderson and J. Willebrand
- PALMER, T.N. and RAISANEN, (2002). “Quantifying the risk of extreme seasonal precipitation events in a changing climate”. *Nature*, 415, pp. 512-514.

- RODRIGUEZ, C. *et al.* (1992). "Statistical method of precipitation prediction". *Atmos. Research*, 28, pp. 299-309
- RODRÍGUEZ-PUEBLA, C. *et al.* (1998). Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula. *Int. J. Climatol.*, 18, 3, pp. 299-316.
- RODRÍGUEZ-PUEBLA, C.; ENCINAS, A.H. and SÁENZ, J. (2001). "Winter precipitation over the Iberian peninsula and its relationships to circulation indices". *Hydrology and Earth System Sciences*, 5, pp. 233-244
- SAUNDERS, A. and GHIL, M. (2001). "A Boolean delay equation model of ENSO variability". *Physica D*, 160, pp. 54-78