



**XVI CONGRESO LATINOAMERICANO DE HIDRÁULICA
SANTIAGO, CHILE.**

7 al 11 de Noviembre, 1994

**REGIONALIZACION DEL URUGUAY SEGUN EL CICLO ANUAL DE
PRECIPITACIONES MEDIANTE "CLUSTER ANALYSIS"**

Rafael Terra y Gabriel Pisciotano

Instituto de Mecánica de los Fluidos e Ingeniería Ambiental

(IMFIA) "Prof. Ing. Oscar J. Maggiolo".

Universidad de la República, Montevideo, Uruguay

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DEPARTAMENTO DE
DOCUMENTACION Y BIBLIOTECA
MONTEVIDEO - URUGUAY

Resumen. El propósito de este trabajo es regionalizar el Uruguay en zonas homogéneas respecto del régimen anual de precipitaciones. Para realizar la partición deseada, se caracterizó cada estación por su ciclo anual medio en el tiempo. Se usó la técnica de *Cluster Analysis* previa descomposición en Componentes Principales (PCA/CA). Basado en el buen desempeño documentado en aplicaciones anteriores, se seleccionó al *Nucleated Agglomerative clustering* (NA), como método. También se ensayaron otros algoritmos con el fin de señalar la consistencia de los resultados. Se compararon los ciclos anuales de las regiones obtenidas. En particular, se encontró que el régimen con doble estación lluviosa, que se atenúa de oeste a este, predomina en todo el país.

Abstract. The purpose of this study is to divide Uruguay in regions having homogeneous rainfall annual patterns. In order to achieve such partition, the annual cycle averaged over the period of study was considered for each station of the network. Principal Component Analysis and Cluster Analysis (PCA/CA) techniques were used. Based on previously published researches, Nucleated Agglomerative clustering (NA) algorithm was selected. Other clustering methods were used to test the consistency of the outcomes. The annual cycles of the resulting regions were compared. In particular, it is found that even though the double rainy season regime weakens from west to east, it prevails throughout the country.

1. Introducción

La partición espacial de un conjunto de estaciones en grupos sustancialmente homogéneos con respecto a un campo físico, facilita estudios posteriores de variabilidad en el tiempo o de correlación con alguna otra variable externa.

En particular, la clasificación de las estaciones pluviométricas del Uruguay en grupos homogéneos, según su ciclo anual, conduce a la obtención de una regionalización del país según el régimen de lluvias. Se tiene así una elección racional de pluviómetros representativos que cubren el país. Esta regionalización objetiva puede ser útil en la búsqueda de relaciones con agentes físicos externos a la hora de ajustar pronósticos; también es necesaria en múltiples aplicaciones hidrológicas.

El Uruguay se encuentra dentro de la región climática de transición entre el régimen continental de lluvias estivales, y uno marítimo con máximos en invierno (World Survey of Climatology Vol.12, Prohaska 1976). La influencia de múltiples agentes climáticos en la región, la gran variabilidad interanual, y su pequeña superficie, conspiran contra una clara partición del mismo según el patrón anual de precipitaciones.

El objetivo del trabajo es lograr la clasificación de las estaciones y en consecuencia la regionalización anteriormente mencionada de forma sistemática. Para ello, se aplicó la técnica de Análisis de Componentes Principales y Análisis de Cluster (PCA/CA), basado en el trabajo de Gong y Richman (1992). A cada estación se la caracterizó por su ciclo anual promediado en el tiempo.

En la sección 2, se especifica el origen de los datos usados y su procesamiento preliminar; la sección 3 trata sobre el PCA y los resultados primarios obtenidos; en la sección 4 se muestran los métodos y resultados del CA; finalmente, se hace una breve presentación y discusión de resultados generales en la sección 5.

2. Datos

Se usaron para este trabajo datos de precipitación mensual en milímetros reportados por 100 estaciones pluviométricas. Este conjunto de datos fue compilado por el IMFIA a partir de datos obtenidos por la Dirección Nacional de Meteorología del Uruguay. Es el mismo conjunto de datos que fue usado en el estudio de Pisciotto et al. (1993), donde está detalladamente documentada la ubicación, período de datos disponibles, y otras características de cada estación.

Se caracterizó a cada estación por su ciclo anual promedio en el período 1933 - 1978 (un vector de R^{12} por estación). Para dicho período, se tenían registros completos en 99 de las

estaciones, las que fueron usadas en Pisciotano et al. (1993, 1994). Para la estación 2826 (Mataojo), faltan algunos datos en la década del 50, por lo cual el promedio se realizó con los datos disponibles.

Las 100 estaciones, representada cada una por su ciclo anual promedio, caracterizan la variación espacial del campo de precipitaciones del Uruguay tal como se dio en el periodo 1933 - 1978. Se evitó el uso de registros más extensos en estaciones donde eran disponibles para que no apareciera como variabilidad espacial parte de la información debida a la variabilidad temporal.

Se realizó una partición geográfica del Uruguay según el área de influencia de cada estación para representar los resultados. Para determinar dicha área de influencia, se partió de una división geográfica básica preexistente en cuadrados de aproximadamente 9.5 km de lado (mapa base de la Red Pluviométrica Nacional, Dirección General de Meteorología del Uruguay). A partir de esta división fina (su paso espacial es de menor orden que la distancia típica entre dos estaciones), se asignó cada cuadrado a la zona de influencia de la estación más cercana, obteniendo la partición de la Figura 1 (adaptación del método de Thiessen).

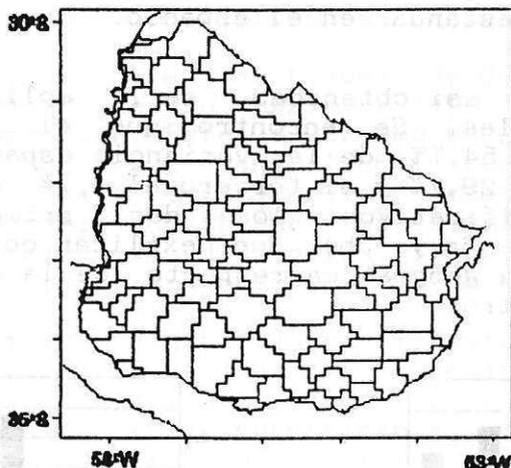


FIG. 1. Zona de influencia de cada una de las 100 estaciones integrantes del conjunto de datos utilizado.

3. Análisis de Componentes Principales

Al ciclo anual medio de cada estación, se le restó el ciclo anual medio promediado espacialmente a través de las 100 estaciones en todo el Uruguay (Fig.2). El vector resultante

caracteriza cuanto se desvía cada estación del promedio espacial uruguayo, y consiste en doce valores cuya unidad son los milímetros. No se normalizaron estos vectores para así distinguir estaciones con regímenes de lluvia de igual estructura a través de los doce meses y distinta intensidad absoluta, lo cual es de interés en este estudio.

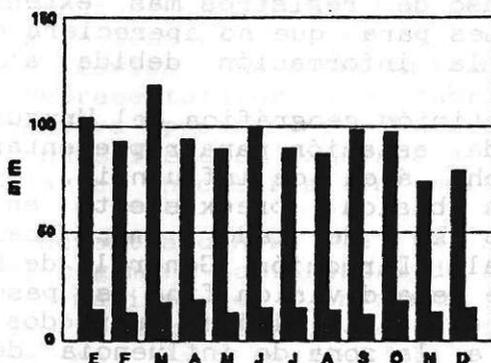


FIG. 2. Ciclo anual medio del Uruguay, y su desviación estándar en el espacio.

A las desviaciones así obtenidas, se le aplicó Análisis de Componentes Principales. Se encontró que el primer vector principal explica el 54,7% de la variancia espacial del ciclo anual, el segundo el 29,7% y el tercero el 4,7% de la misma. Se tomaron como significativos los dos primeros vectores principales (Figuras 3a y 3b), que explican conjuntamente el 84,4% de la variancia geográfica respecto de la media uruguayu; se despreciaron los otros.

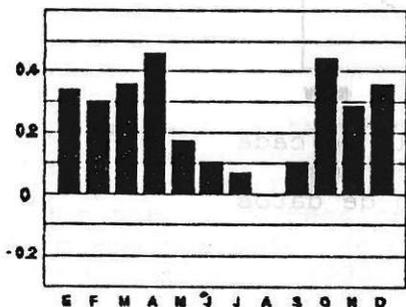


FIG. 3a. Primer vector principal que explica el 54,7% de la variancia geográfica en el ciclo anual de lluvias.



FIG. 3b. Segundo vector principal que explica el 29,7% de la variancia geográfica en el ciclo anual de lluvias.

Estos vectores propios satisfacen el criterio de estabilidad de North et al. (1982); la suma de las estimaciones de los errores de muestreo para los respectivos valores propios, vale menos de la mitad de su diferencia.

El primer vector principal es un ciclo esencialmente semianual, por lo que la proyección sobre él indica la tendencia hacia regímenes con doble estación lluviosa. El segundo vector es un ciclo anual y señala la tendencia hacia patrones con máximo en invierno.

4. Cluster Analysis

Para agrupar las 100 estaciones de la red en grupos homogéneos se aplicó la técnica multivariada de *Cluster Analysis* (CA). El CA es una técnica algorítmica con la habilidad para definir aglomerados o grupos con cohesión interna y aislamiento externo. Una descripción sintética del método y sus variaciones, se puede encontrar en Gong y Richman (1992).

Se caracterizó cada estación por los dos primeros componentes principales; éstos forman un vector en milímetros de dos coordenadas que resulta de proyectar las desviaciones anteriormente halladas sobre los vectores principales (considerados como los patrones anuales estadísticamente relevantes).

Por su mejor desempeño en el trabajo de Gong y Richman (1992), se seleccionó al *Nucleated Agglomerative clustering* (NA) como método de aglomeración. Este algoritmo no jerárquico, descrito con detalle en Ronberg y Wang (1986), surge de la aplicación sucesiva y aglomerativa del algoritmo *K-means* (KM); en esencia, consiste en una serie de KM que seleccionan las semillas para una última aplicación de KM. También Milligan (1980) recomienda KM como de buena performance si se hace una adecuada elección de semillas.

Un primer elemento a determinar al aplicar CA es el número final de *clusters*. La pequeña desviación estándar de la muestra, el hecho que sea explicada en más de un 84% por solo dos patrones, lo pequeño de la superficie a regionalizar, y sobre todo el objetivo del trabajo, llevaron a que se eligieran a priori en 4 el número final de *clusters*. Recordamos que se pretende una regionalización adecuada y utilizable en potenciales estudios de variabilidad climática regional, para los cuales no tiene sentido considerar en demasiado detalle variaciones espaciales locales.

Por ser no jerárquico, NA requiere de un segundo dato: la elección de semillas (o conjunto inicial de *clusters*) para el primer paso. Como se sugiere en Ronberg y Wang (1986), de tomarse un número suficientemente grande de semillas, se soluciona este problema, pues la clasificación final resulta independiente de las mismas. En este caso, se obtuvo la misma

división en 4 regiones partiendo de tres diferentes conjuntos de 50 semillas: las primeras 50 estaciones, las últimas 50, y las estaciones impares. La regionalización resultante se muestra en la Figura 4.

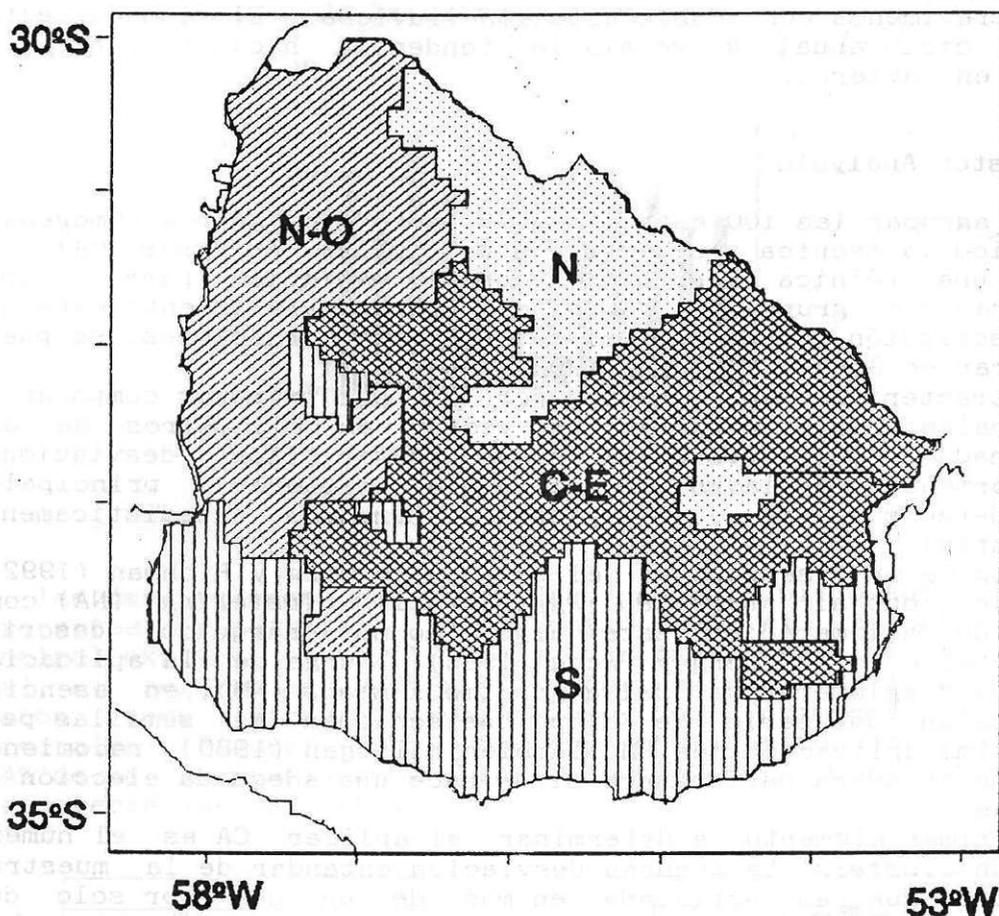


FIG. 4. Partición del Uruguay en 4 *clusters* según su ciclo anual de precipitaciones usando *Nucleated Agglomerative clustering* (NA4).

Para cada una de las 4 regiones identificadas se construyó el ciclo anual medio, promediando el ciclo anual entre todas las estaciones de cada *cluster*. Ver Figura 6.

Se repitió el trabajo con dos métodos jerárquicos: *Average Linkage within the New Group* (AL2) y *Ward's Method* (W), los cuales están explicados en Gong y Richman (1992). Se podrá comprobar así cuales regiones son robustas al cambio de algoritmo y cuales no. En las Figs. 5a y 5b, se muestran los

resultados para los métodos AL2 y W respectivamente, donde, por ser jerárquicos, se pueden distinguir 4 y 6 *clusters*. En las Figs. 5c y 5d, se muestran las particiones con el método NA de 5 y 6 regiones respectivamente.

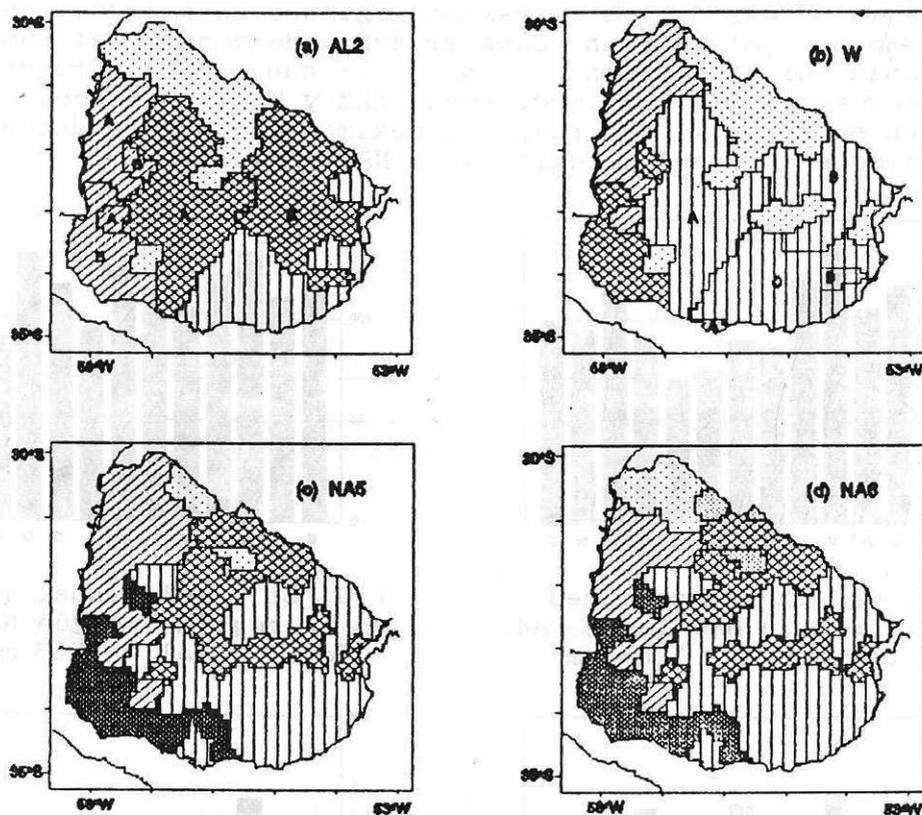


FIG. 5. Partición del Uruguay según su ciclo anual de precipitaciones usando diferentes algoritmos. (En las Fig. a y b, la división en 6 regiones se señala mediante las letras A, B y C.)

5. Discusión

Salvo 4 estaciones aisladas, NA4 regionaliza en regiones conexas y de tamaños comparables: una ubicada al noroeste del país (N-O) con 23 estaciones, otra al norte (N) con 12, otra en el centro-este (C-E) con 32 y otra al sur (S) con 33 estaciones (Fig.4).

Comparando con las restantes regionalizaciones obtenidas (Fig.5), se constata que las regiones N-O, N y C-E, se repiten con pequeñas variaciones.

La diferencia más notoria se produce en la unificación de la región S, que con los otros métodos aparece dividida. Este hecho relativiza la mayor influencia marítima en el sudeste del país, en donde, según el World Survey of Climatology (Vol.12, Prohaska 1976), predomina el régimen con máximos en invierno. Si bien hay estaciones en las cuales su máximo registro se produce en junio, no llegan a conformar un *cluster* definido cuando el número de divisiones es del orden de 4 a 6. Aún en el promedio (no mostrado) en el *cluster* que, según AL2 y W (con 6 regiones) se ubica en el sudeste del país, el máximo se sigue produciendo en marzo como en todas las regiones de NA4 (Figura 6).

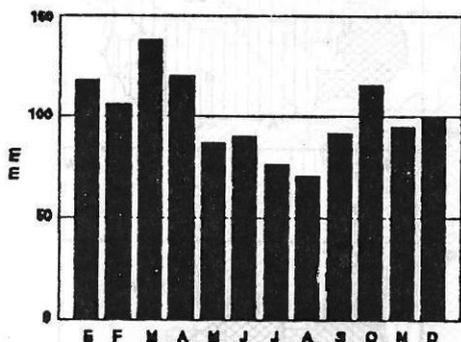


FIG. 6a. Ciclo anual medio en la región N-O según NA4. Total anual = 1208 mm

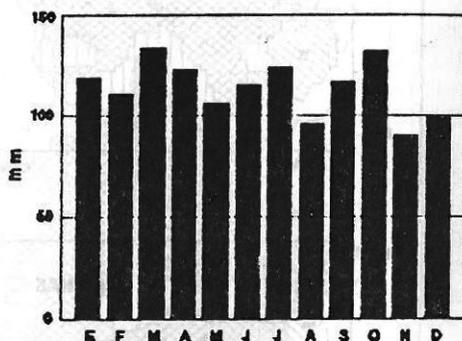


FIG. 6b. Ciclo anual medio en la región N según NA4. Total anual = 1366 mm

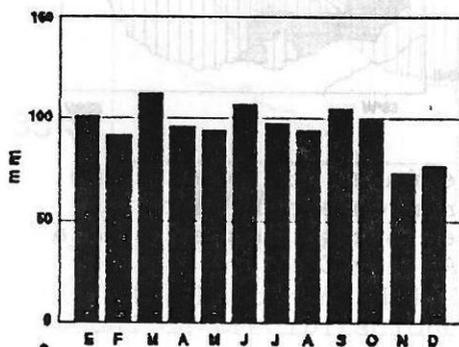


FIG. 6c. Ciclo anual medio en la región C-E según NA4. Total anual = 1148 mm

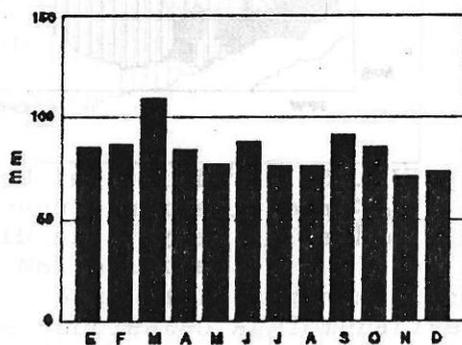


FIG. 6d. Ciclo anual medio en la región S según NA4. Total anual = 1002 mm

La inspección de la Fig.6 y la Tabla 1 motiva las siguientes reflexiones. La región N-O se distingue por un marcado ciclo semianual (con máximos en otoño), y una intensidad de precipitaciones mayor en verano que en invierno. Estas

características se atenúan, y en algunos casos se revierten, en las restantes regiones definidas por NA4, pero siempre conservándose el máximo en Marzo. El mínimo valor de la Tabla 1 se alcanza entre las regiones N-O y C-E, esta última con mayores registros invernales que estivales, todo lo cual confirma la transición entre un régimen continental y uno marítimo, produciéndose dicha transición de oeste a este.

La diferencia más notoria entre los regimenes de las zonas N, C-E, y S, se da en el total anual, que como es sabido, aumenta de sur a norte.

	N	S	C-E	N-O
N	1	0.72	0.79	0.59
S	-	1	0.80	0.66
C-E	-	-	1	0.28

Tabla 1. Correlación entre los ciclos anuales medios (Fig.6) de las regiones definidas por NA4.

NOTA: La lista de los números de las estaciones correspondientes a cada región está disponible en documentación técnica del IMFIA.

AGRADECIMIENTOS

Los autores quieren agradecer a R.Guarga y J.L.Genta del IMFIA el continuo entusiasmo transmitido durante el transcurso de este trabajo, y a A.Díaz varios valiosos comentarios. G.Pisciottano agradece a C.R.Mechoso la oportunidad de haber conocido las posibilidades de la técnica de *Cluster Analysis* durante su estadía en la UCLA. R.Terra agradece a F.Charbonier el continuo asesoramiento computacional brindado. Este trabajo fue parte de proyectos financiados por Usinas y Transmisiones Eléctricas-Uruguay, Comisión Sectorial de Investigación Científica-UR, y Consejo Nacional de Investigación Científica y Técnica-Uruguay.

REFERENCIAS

Gong, X.-F. and M.B. Richman, 1992: An examination of methodological issues in clustering North American precipitation. (pp J103-J108). Proceedings of the 12th Conference on Probability and Statistics in the Atmospheric Sciences, Toronto, Canada. American Meteorological Society.

Milligan, G.M., 1980: An examination of the effect of six types of error perturbation of 15 clustering algorithms. *Psychometrika*, 45, 325-342.

North, G.R., T.L. Bell, and R.F. Cahalan, 1982: Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.*, 110, 699-706.

Pisciottano, G., A. Diaz, G. Cazes, and C.R. Mechoso, 1993: Rainfall anomalies in Uruguay associated with the extreme phases of El Niño Southern Oscillation phenomenon. Versión disponible en el IMFIA, Fac. de Ingeniería, Universidad de la República, J.Herrera y Reissig 565, Montevideo, Uruguay.

Pisciottano, G., A. Diaz, G. Cazes, and C.R. Mechoso, 1994: El Niño Southern Oscillation impact on rainfall in Uruguay. *Journal of Climate*, en prensa.

Prohaska, F., 1976: The Climate of Argentina, Paraguay and Uruguay. (pp.13-112), W. Schwerdtfeger (Editor), *Climates of Central and South America*. World Survey of Climatology, Vol.12, Elsevier, 532pp.

Ronberg, B. and W.-C. Wang, 1987: Climate patterns derived from Chinese proxy precipitation records: An evaluation of the station networks and statistical techniques. *J.Climatol.*, 7, 391-416.