

## DISPONIBILIDAD POTENCIAL DE RECURSOS HÍDRICOS EN COLOMBIA SEGÚN EL CRITERIO DE LA ENTROPÍA INFORMACIONAL

*Oscar David Álvarez V. & Germán Poveda J.*  
*Posgrado en Aprovechamiento de Recursos Hidráulicos, Facultad de Minas*  
*Universidad Nacional de Colombia, Medellín*  
*odalvare@unal.edu.co, gpoveda@unal.edu.co*

Recibido para evaluación: 26 de Mayo de 2006 / Aceptación: 30 de Junio de 2006 / Versión final: 07 de Julio de 2006

### RESUMEN

Evaluamos la disponibilidad potencial del recurso hídrico en Colombia mediante la estimación del desorden en las series de intensidad de la precipitación y de su distribución durante el año. Para cuantificar el grado de desorden se utiliza el concepto de entropía a partir de dos enfoques representados por la entropía de intensidad (*EI*) y la entropía de repartición (*ER*). La primera mide el desorden para el registro completo de precipitación, y la segunda mide el desorden en la distribución mensual de la cantidad de agua disponible por precipitación durante todo el año. Tales medidas de la entropía (*EI* y *ER*) permiten establecer criterios para clasificar la disponibilidad potencial del recurso hídrico, usando algoritmos de clasificación simple y de agrupamiento por centroides (“*k-Means clustering*”). Para el análisis se ha usado información de registros mensuales de precipitación en 108 estaciones de Colombia. Los resultados permiten clasificar las zonas húmedas y las zonas áridas para el territorio Colombiano.

**PALABRAS CLAVES:** Entropía, Disponibilidad de Agua, Recursos Hídricos, Precipitación, Colombia.

### ABSTRACT

The availability of water resources in Colombia is quantified through estimation of the disorder degree of precipitation records and their distribution during the annual cycle. Disorder is quantified by using two different measures of entropy, the so-called Intensity entropy (*IE*) and apportionment entropy (*AE*). The first one measures disorder of the complete precipitation record, whereas the second one measures disorder of the monthly distribution of water availability during the seasonal cycle. Such entropy measures (*IE* and *AE*) allow defining criteria to classify the potential availability of water resources, by using simple and k-Means clustering classification algorithms. Our analysis is performed to monthly records of precipitation in 108 raingauges located in Colombia. Our results allow the classification of wet and dry regions for Colombia.

**KEY WORDS:** Entropy, Water Availability, Water Resources, Precipitation, Colombia.

## 1. INTRODUCCIÓN

Para el desarrollo de estrategias locales y nacionales del manejo de los recursos hídricos es importante conocer la disponibilidad potencial de agua y sus demandas futuras con el fin de predecir si podrían presentarse épocas de déficit en la disponibilidad efectiva del recurso hídrico. En particular, la estimación de la disponibilidad de agua en una zona dada debe tener en cuenta el hecho de que la precipitación constituye un insumo principal de agua dentro del ciclo hidrológico, por lo cual la estimación de su desorden temporal y espacial puede surgir como una medida preliminar de oferta hídrica disponible y puede ser útil para establecer zonas en las cuales la disponibilidad del agua sea alta o baja y establecerse las zonas de escasez en las cuales se debe tomar medidas de regulación para asegurar la disponibilidad del recurso durante todo el año (Kawachi, et al., 2001).

Una medida del desorden de una variable aleatoria es la entropía, la cual puede calcularse a partir de una función de distribución de probabilidad para una variable aleatoria dada, tanto de manera discreta como de manera continua, usando la entropía definida por Shannon. De acuerdo con su propuesta, la entropía es sinónimo de incertidumbre (Kapur y Kasevan, 1992). Shannon afirma que la incertidumbre o entropía de una variable aleatoria está asociada con el negativo del valor esperado del logaritmo de su función de distribución de probabilidad. En general, se ha encontrado que mientras aumenta la aleatoriedad y el desorden de la variable, disminuye la asimetría de su distribución de probabilidad, lo cual contribuye a que aumente el valor de la entropía. El valor máximo de la entropía para una variable aleatoria se logra cuando se asume que su función de distribución de probabilidad es uniforme, mientras que el valor se reduce a cero cuando se tiene una variable determinista, es decir, cuando la probabilidad de ocurrencia del valor de la variable es uno (Goovaerts, 1999).

Por otro lado, si el concepto de entropía se compara con el concepto de varianza, puede observarse que la varianza cuantifica solamente la desviación de las observaciones con respecto a la media, mientras que la entropía puede llegar a ser una medida más general de desviación (Maruyama, et al 2005).

Ahora bien, la aplicación de la entropía a la precipitación observada puede cuantificar su desorden. Para tal fin se utiliza dicho concepto a partir de dos enfoques diferentes, representados mediante la entropía de intensidad ( $EI$ ) y la entropía de repartición ( $ER$ ). En el primer caso se considera que la precipitación mensual es una variable aleatoria y que la relación entre su valor y su probabilidad de ocurrencia dentro de un período dado puede ser capturada mediante la entropía. En el segundo caso se ha asumido que la relación entre la lluvia mensual y la lluvia acumulada anual es una variable aleatoria que refleja el carácter probabilístico de la variación de la precipitación dentro del año, y para la cual, análogamente al caso anterior, puede estimarse su desorden a partir del concepto de entropía. Finalmente, se emplean dos algoritmos de clasificación: clasificación simple y agrupamiento proximal con base en centroides ("*k-Means Clustering*") con el fin de obtener categorías representativas con respecto a la disponibilidad del recurso para el territorio colombiano.

## 2. DATOS Y METODOLOGÍA

### 2.1. Selección de estaciones para el análisis

Para la estimación de la disponibilidad potencial de recursos hidráulicos para el territorio colombiano, se utilizaron registros de precipitación a resoluciones mensual, diaria y horaria de 157 estaciones extraídas de dos fuentes: (i) 51 estaciones con registro en resolución horaria utilizadas en trabajos anteriores para caracterizar el ciclo diario de la hidrología colombiana (Viera y Moreno, 2002) y (ii) 106 estaciones a resolución diaria contenidas dentro de la base de datos de HidroSIG. A partir de dicho conjunto, se seleccionaron 108 estaciones definitivas que cumplieron con una longitud mínima de registro de 20 años (ver Figura 1). Éste criterio se estableció con el fin de asegurar cierto grado de confiabilidad en la estimación de la entropía, considerando el hecho de que la confiabilidad de la estimación aumenta a medida que se incrementa la longitud en el registro de la serie. Los faltantes dentro del registro de precipitación utilizados para el análisis se completaron mediante la aplicación del método del vecino más cercano. Para el caso de las estaciones a resolución horaria y diaria, sus valores se agregaron para obtener series de precipitación a resolución mensual.

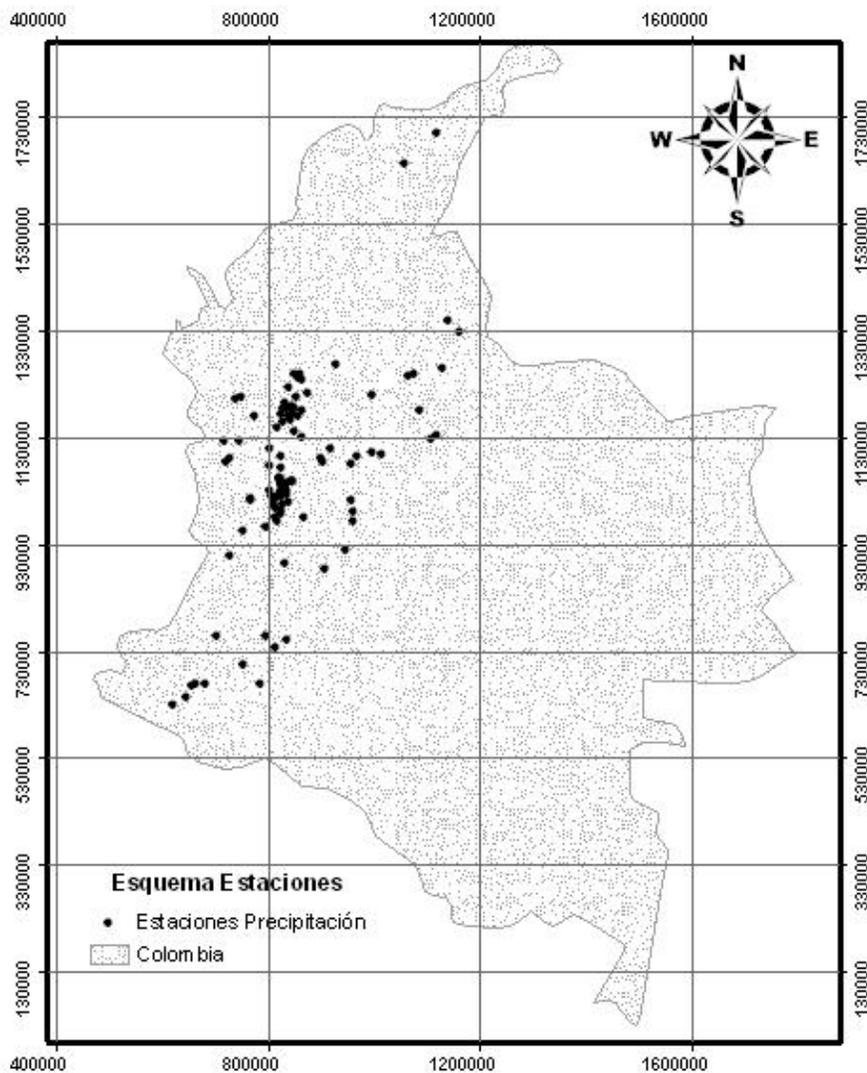


Figura 1. Esquema de la localización de las estaciones de precipitación

## 2.2. Cálculo de la Entropía de Intensidad (EI)

La entropía de intensidad para un conjunto discreto de probabilidades  $p_1, p_2, \dots, p_n$ ; se define como:

$$EI = - \sum_{i=1}^n p_i \log(p_i) \quad (1)$$

En la cual  $EI$  es la entropía de intensidad y  $p_i$  es la probabilidad de ocurrencia para un valor de precipitación mensual, cuando ésta se considera como una variable aleatoria. Las probabilidades  $p_i$  para cada estación se calculan de manera discreta tomando en cuenta todos los valores disponibles de precipitación mensual en cada estación y su probabilidad de ocurrencia, a partir del siguiente esquema metodológico:

- 1) Se debe disponer de series completas de precipitación mensual para cada uno de los  $N$  años disponibles dentro del registro.
- 2) Se divide todo el intervalo de precipitación en  $n$  clases del mismo tamaño.
- 3) Se cuentan el número de observaciones por clase con el fin de estimar la frecuencia de ocurrencia  $f_i$  para cada clase y obtener una tabla de distribución de frecuencias.
- 4) Se calcula la frecuencia relativa ( $f_i/N$ ) para cada clase, de esta forma se obtiene una función de densidad de probabilidad de forma discreta para todo el rango de lluvia mensual.
- 5) Se estima la entropía de intensidad  $IE$  usando la Ecuación 2, la cual se encuentra expresada en función

de las frecuencias relativas:

$$EI = -\sum_{i=1}^n \left(\frac{f_i}{N}\right) \log_2 \left(\frac{f_i}{N}\right) = -\sum_{i=1}^n \left(\frac{f_i}{N}\right) \frac{\ln\left(\frac{f_i}{N}\right)}{\ln(2)} \quad (2)$$

En donde:  $N$  es el número de años de registro,  $n$  es el número de clases y  $f_i$  es el número de observaciones dentro de la clase  $i$ . En la Ecuación 2, la unidad de  $EI$  es un “bit” con una base logarítmica de 2. De la Ecuación 2 puede verse que la entropía de intensidad está definida dentro del intervalo semi-infinito  $[0, \infty]$ , y es una medida para estimar el desorden de la precipitación mensual. Valores bajos de entropía de intensidad son indicio de que hay poco desorden en la función de distribución de probabilidad, lo cual, en este caso, se traduce en que la

distribución de frecuencias relativas de precipitación mensual está sesgada. Por el contrario, valores elevados en la entropía de intensidad indican que hay mucho desorden en una función de distribución de probabilidad, lo cual se traduce en que los valores de precipitación mensual se distribuyen en un intervalo grande dentro de la distribución de frecuencias. Las probabilidades de la función de densidad de la precipitación mensual se definen sobre el dominio positivo de las abscisas, incluyendo un origen en cero debido a la no negatividad de la lluvia. A medida que aumenta el valor de la entropía de intensidad, aumenta el valor esperado de la precipitación mensual, aplanando la gráfica de la función de frecuencias relativas (ver Figura 2). Lo anterior sugiere que la  $EI$  se relaciona directamente con la cantidad de lluvia mensual que se precipita sobre un sitio (Maruyama et al, 2005).

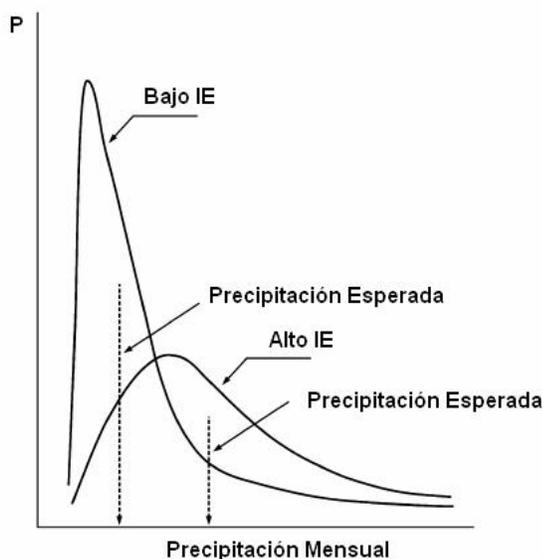


Figura 2. Aumento esperado de la precipitación mensual a medida que aumenta  $EI$  (Adaptada de Maruyama et al., 2005)

### 2.3. Cálculo de la Entropía de Repartición ( $ER$ )

A partir de la definición de entropía (Ecuación 1), puede definirse la entropía de repartición ( $ER$ ) para la precipitación mensual. Considerando que  $r_i$  representa la precipitación agregada durante el mes  $i$  en un año cualquiera dentro de la longitud del registro, la precipitación agregada  $R$  durante un año puede expresarse como la suma de los  $r_i$  dentro del año:

$$R = \sum_{i=1}^{12} r_i \quad (3)$$

Ahora bien, la proporción  $r_i/R$  representa la cantidad de meses en los cuales hay disponibilidad de agua dentro de un año cualquiera, lo cual puede interpretarse como una probabilidad de ocurrencia de eventos de precipitación dentro del año. Entonces, siendo  $p_i$  la relación entre la precipitación mensual y la precipitación anual agregada, y empleando la definición de entropía de Shannon, la entropía de repartición, puede escribirse como:

$$ER = -\sum_{i=1}^{12} \left(\frac{r_i}{R}\right) \log_2 \left(\frac{r_i}{R}\right) = -\sum_{i=1}^{12} \left(\frac{r_i}{R}\right) \frac{\ln\left(\frac{r_i}{R}\right)}{\ln(2)} \quad (4)$$

En donde la unidad de  $ER$  es un “bit” del logaritmo de base 2. La entropía de intensidad se define para cada estación sin importar el año, mientras que la entropía de distribución se define para cada año y para cada estación. Con los valores de entropía de repartición para cada año se define una entropía de repartición promedio, la cual es representativa para cada estación. La entropía de repartición mide el desorden de la distribución de la precipitación mensual dentro de un año. Por definición, la Ecuación 4 muestra cuando la precipitación mensual está distribuida dentro de los doce meses del año  $ER$  alcanza un máximo valor de  $\log_2(12)$  y un valor mínimo de cero. Valores grandes de  $ER$  indican buena disponibilidad potencial de recursos hidráulicos con menor variabilidad estacional en la precipitación mensual.

### 2.4. Estandarización

Con base en las Ecuaciones 1 y 4 pueden calcularse la entropía de intensidad y la entropía de repartición, respectivamente para cada estación, con lo cual se tienen dos conjuntos de entropías. Para representar éstos conjuntos en los mismos términos deben ajustarse las propiedades de sus distribuciones de probabilidad, de tal forma que su media sea cero y su desviación estándar sea unitaria, lo cual se logra estandarizando los valores de cada una de las componentes del conjunto de la siguiente manera:

$$EX^* = \frac{EX - m_{EX}}{\sigma_{EX}} \quad (5)$$

Donde  $EX^*$  es el valor estandarizado de la entropía,  $EX$  es el valor estimado de la entropía no estandarizada en unidades de bits ( $EI$  o  $ER$  según Ecuaciones 1 y 4, respectivamente),  $m_{EX}$  es la media de la entropía no estandarizada ( $EI$  o  $ER$ ) y  $\sigma_x$  es la desviación estándar para la media de la entropía no estandarizada ( $EI$  o  $ER$ ). Mediante la estandarización puede dibujarse un diagrama de dispersión para visualizar la dependencia de la localización para la relación entre la entropía de repartición y la entropía de intensidad. Entonces, puede estimarse la disponibilidad potencial de recursos hidráulicos dentro de un área de interés.

### 2.5. Agrupamiento simple

Las parejas de valores para la entropía de intensidad y repartición para cada estación se dibujan en un diagrama de dispersión, en el cual las coordenadas de un punto caracterizan la disponibilidad potencial de recursos hidráulicos en la localización correspondiente a cada estación, mientras que los ejes del diagrama de dispersión representan la media de los valores estandarizados de las entropías calculadas. El agrupamiento de los puntos disponible en categorías puede ayudar a delinear zonas características con respecto a la disponibilidad del recurso hídrico. La forma más sencilla para la clasificación es usar líneas hipotéticas que pasan por el origen del sistema de coordenadas del diagrama. Mientras más líneas se usen, más categorías se incluyen con el propósito de clasificar la disponibilidad. Los esquemas empleados en este análisis se muestran en la Figura 3.

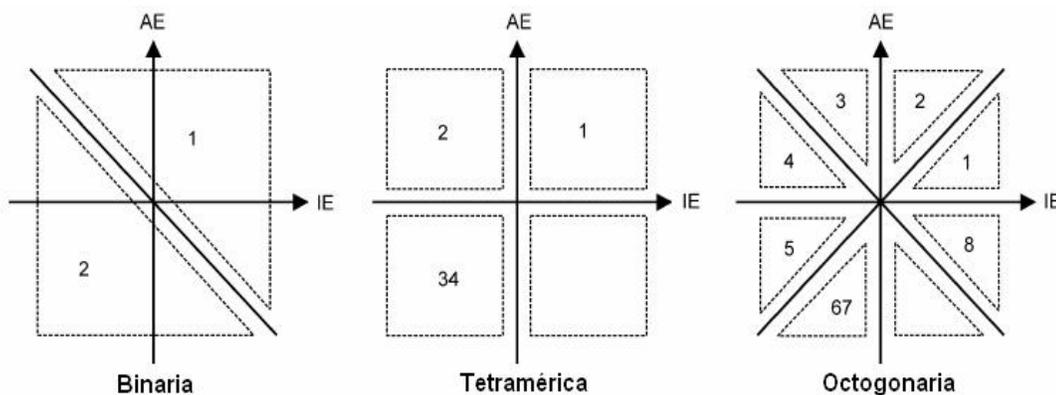


Figura 3. Clasificación simple

## 2.6. Agrupamiento “*k-Means*”

Otra manera de agrupar es mediante algoritmos de agrupamiento con base en centroides (“*k-Means clustering*”), el cual es utilizable cuando se tiene un conjunto de puntos que deben clasificarse, como el presente caso de estudio. Éste algoritmo es uno de los más sencillos para realizar clasificación no supervisada, y propende por clasificar de manera rápida y sencilla un conjunto de datos dentro de un número de categorías determinadas a priori. La idea principal es determinar los centroides correspondiente a cada clase y posteriormente reclasificar por proximidad. Inicialmente se eligen *k* centroides de manera subjetiva, pero de tal manera que se encuentran lo más separados posible entre sí. El siguiente paso es clasificar inicialmente los puntos de acuerdo con el centroide que se encuentre más cercano. En este punto se recalcula la posición del centroide como el resultante de los puntos dentro de cada clase y se reclasifican los puntos. La operación se repite hasta que los centroides no se mueven más. La finalidad del algoritmo es minimizar la función objetivo del error cuadrático (McQueen, 1967):

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (6)$$

Donde:  $\|x_i^{(j)} - c_j\|^2$  es una distancia entre el dato puntual de entropía  $x_i^{(j)}$  y el centro de la clase a la cual pertenece,  $c_j$ . El algoritmo que se implementó tiene la siguiente forma:

- 1) Se ubican *K* puntos que representaron los centroides iniciales para las clases. Sus coordenadas se establecieron de acuerdo con la clasificación simple análoga, es decir, se tomaron como centroides iniciales puntos ubicados dentro de las clases obtenidas en la clasificación simple con el mismo número de categorías.
- 2) Cada par de entropías (*EI*, *ER*) se asignan a la clase

- correspondiente al centroide más cercano.
- 3) Se recalcula la posición de los *K* centroides.
- 4) Se repiten los pasos 2 y 3 hasta que las coordenadas calculadas para los centroides no varíen.

## 3. CLASIFICACIÓN DE LA DISPONIBILIDAD POTENCIAL DE RECURSOS HÍDRICOS EN COLOMBIA

### 3.1. Clasificación y Elaboración de mapas

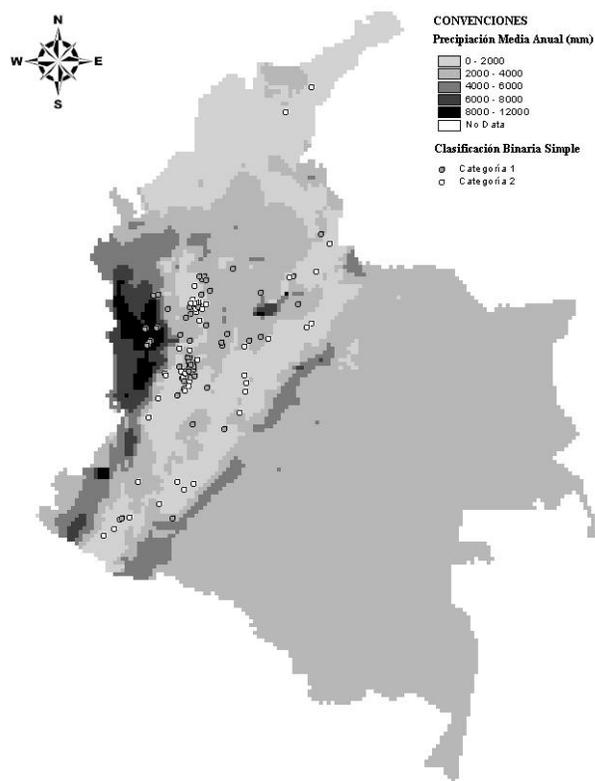
En cada estación se calcularon los valores para la entropía de intensidad y la entropía de repartición. Se encontró que el cálculo de la entropía de intensidad es sensible a cambios del tamaño del intervalo de clase del histograma, por lo cual se calcularon y se clasificaron diversos valores de entropía obtenidos mediante histogramas de frecuencia relativa con varios tamaños de clase. En la Tabla 1 se muestran los estadísticos descriptivos para los diferentes conjuntos obtenidos de entropía de intensidad y entropía de repartición. A pesar de que los valores para la entropía de repartición cambian, las clasificaciones obtenidas son parecidas, con excepción de zonas para las cuales los valores de *EI* y *ER* son cercanos a la media. Los valores de entropía de intensidad y entropía de distribución se estandarizan mediante el uso de la Ecuación 5 y los valores para la media y la desviación estándar consignados en la Tabla 1. Con el fin de realizar una comparación visual se clasificó la disponibilidad potencial de recursos hidráulicos mediante agrupamiento simple y agrupamiento por centroides (“*k-Means clustering*”) para los dos tipos de entropía utilizados. Los resultados se representan mediante el uso de diagramas de dispersión y la localización de los resultados según la estación correspondiente sobre mapas del territorio colombiano. En ambos esquemas, cada punto representa la pertenencia de cada estación incluida dentro del análisis a cada uno de los grupos de disponibilidad potencial de recursos hídricos previamente establecidos.

**Tabla 1.** Estadísticos para las muestras de entropía sin estandarizar *IE\** y *AE\**

Número de Clases	10 (200mm)	20 (100mm)	40 (50mm)
Media <i>IE*</i>	1.258	2.342	3.660
Media <i>AE*</i>	3.363	3.363	3.363
D. E. <i>IE*</i>	0.438	0.441	0.459
D. E. <i>AE*</i>	0.084	0.084	0.084
Varianza <i>IE*</i>	0.191	0.007	0.211
Varianza <i>AE*</i>	0.007	0.195	0.007

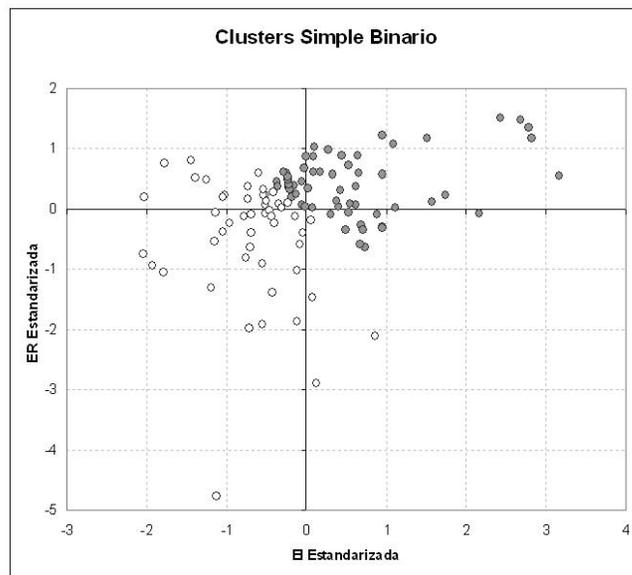
### 3.2. Resultados y Análisis

Las Figuras 4 y 5 muestran las clasificaciones simple, binaria y tetramérica para la disponibilidad potencial de recursos hidráulicos (precipitación), para un tamaño de clase de histograma de frecuencia de 50 mm. En forma similar, en las Figuras 6 y 7, se muestran las clasificaciones obtenidas para la disponibilidad potencial de recursos hidráulicos mediante el uso del algoritmo de los k-Centroides (“*k-Means clustering*”) para tamaño de clase de histograma de frecuencia de 50 mm. La distribución espacial de la clasificación de disponibilidad potencial se ha desplegado sobre el mapa de precipitación media multianual de Colombia con el fin de tener una idea de cómo se relaciona la clasificación obtenida con la precipitación anual estimada. Para tal fin se ha usado el mapa de precipitación obtenido mediante kriging con deriva del ENA, el cual se encuentra contenido dentro de la base de datos de HidroSIG (ver Poveda et al., 2006).

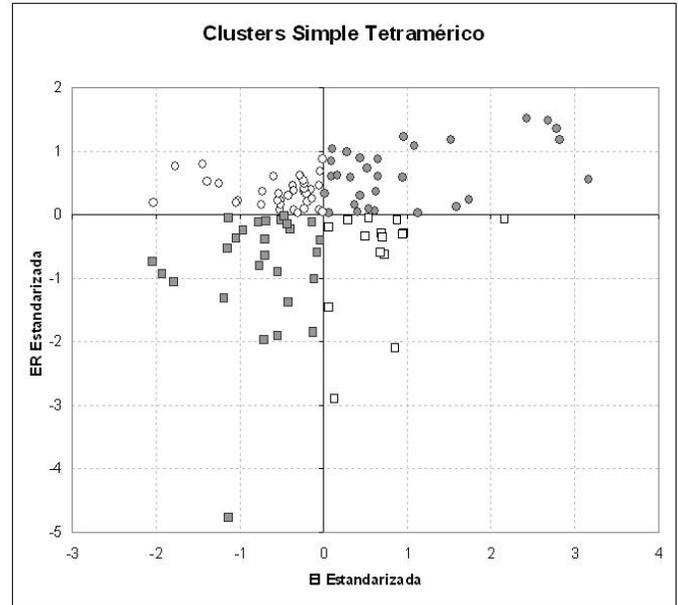
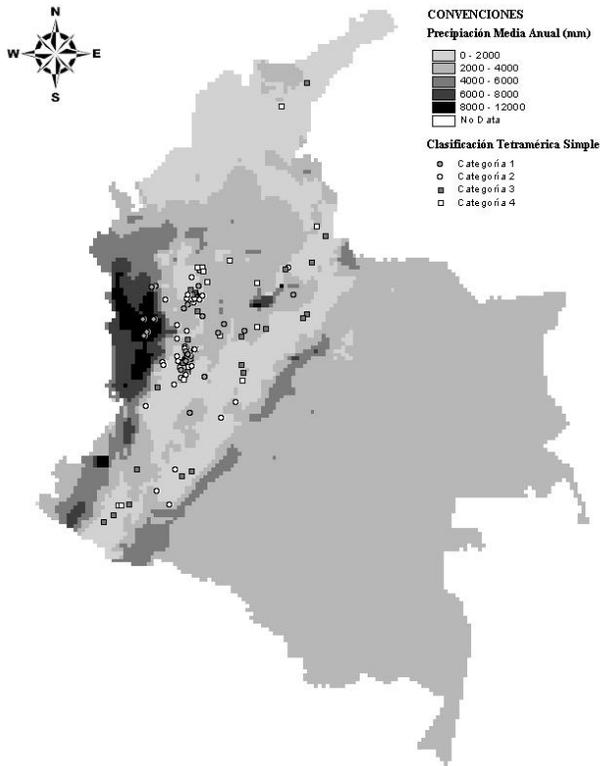


#### 3.2.1. Clasificación binaria

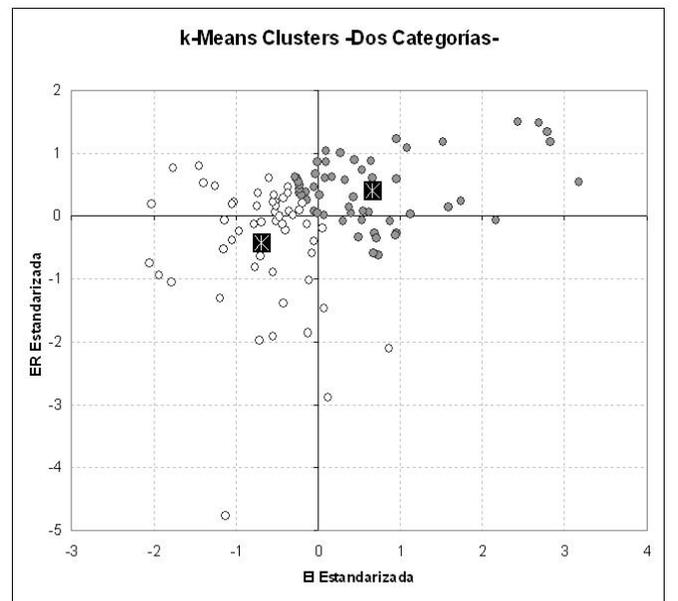
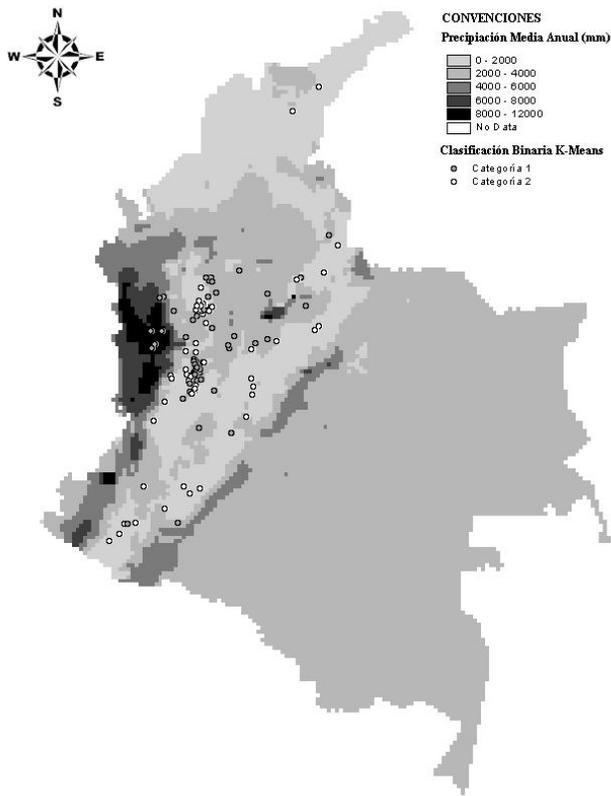
En las Figuras 4 y 6 puede observarse los resultados de las clasificaciones con el agrupamiento simple y el agrupamiento de los k-Centroides (“*k-Means clustering*”) considerando dos categorías. Puede considerarse que los dos grupos resultantes representan dos zonas con diferentes características de disponibilidad de agua. En los diagramas los círculos grises representan las zonas en las cuales hay exceso relativo de disponibilidad de agua con respecto a los círculos de color blanco, que pueden considerarse como representativos para zonas en las cuales hay escasez de agua. Con respecto al mapa de precipitación anual, puede observarse que las zonas correspondientes a escasez obtenidas mediante la clasificación por entropía corresponden a aquellas más áridas dentro del territorio colombiano, como por ejemplo la Guajira, mientras que las zonas correspondientes a disponibilidad corresponden a aquella en las cuales la precipitación es elevada, por ejemplo la Región Pacífico Norte en el departamento de Chocó.



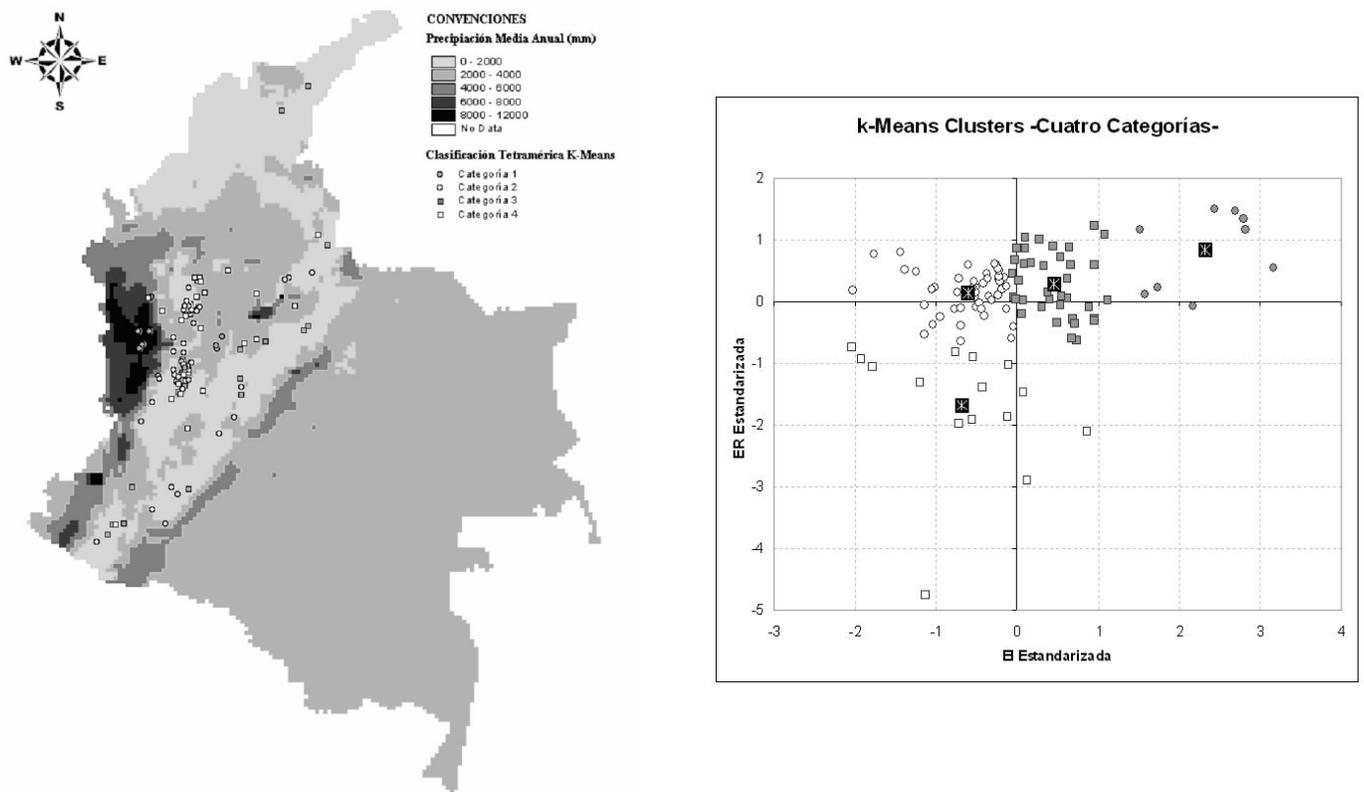
**Figura 4.** Clasificación simple binaria. Círculos grises corresponden a categoría 1, círculos blancos corresponden a categoría 2.



**Figura 5.** Clasificación simple tetramérica. Círculos grises corresponden a categoría 1, círculos blancos corresponden a categoría 2, cuadros grises corresponden a categoría 3, cuadros blancos corresponden a categoría 4.



**Figura 6.** k-Means Clusters –Dos Categorías-. Círculos grises corresponden a categoría 1, círculos blancos corresponden a la categoría 2. Los cuadros negros corresponden los centroides de las categorías para la última iteración



**Figura 7.** k-Means Clusters –Cuatro Categorías-. Círculos grises corresponden a categoría 1, círculos blancos corresponden a categoría 2, cuadrados grises corresponden a categoría 3, cuadrados blancos corresponden a categoría 4. Los cuadros negros corresponden los centroides de las categorías para la última iteración

### 3.2.2. Clasificación en cuatro categorías

El método de clasificación tetramérica simple divide el dominio de las entropías en cuatro regiones, cada una de las cuales corresponde a cada cuadrante del sistema cartesiano  $EI-ER$ , cada cuadrante se considera como característico para la disponibilidad potencial de recurso hídrico. Los resultados obtenidos se muestran en la Figura 5.

El cuadrante superior ( $EI > 0$  y  $ER > 0$ , en la Figura 5) representa las zonas en las cuales se presenta gran disponibilidad de agua anualmente y tal disponibilidad se distribuye bien a lo largo del año, es decir, la precipitación anual se distribuye a lo largo de todos los meses y no se concentra en una sola época. Como puede observarse en la Figura 5, dentro de esta categoría se encuentran las localizaciones de estaciones ubicada en el norte de la región Pacífica colombiana (círculos de color gris).

El cuadrante inferior ( $EI < 0$  y  $ER < 0$ , en la Figura 5) representa las zonas en las cuales la oferta hídrica por

precipitación es baja y se encuentra concentrada dentro de ciertas épocas del año. El algoritmo de evaluación de disponibilidad encontró que las zonas colombianas que presentan las características anteriormente comentadas se encuentran principalmente en el norte del país, y corresponden a zonas como por ejemplo la Guajira o en la zona norte de la cordillera oriental (cuadros de color gris en la Figura 5).

El primer cuadrante intermedio ( $EI < 0$  y  $ER > 0$ , en la Figura 5) se refiere a las zonas en las cuales se presenta baja disponibilidad de agua, pero ésta se distribuye a lo largo de todo el año. Para éste tipo de zonas se necesitan embalses de regulación multianual para suplir la demanda cuando ésta supere la oferta natural hídrica. En este punto cabe aclarar que es necesario analizar la conveniencia económica de la construcción y la operación de dichos embalses, lo cual es un problema abierto. En el análisis realizado para el territorio colombiano se encontró que dichas zonas corresponden a las marcadas mediante puntos blancos en la Figura 5, como por ejemplo algunos sectores localizados en la cordillera central.

El segundo cuadrante intermedio ( $EI > 0$  y  $ER < 0$ , en la Figura 5) representa las zonas en las cuales hay gran disponibilidad de agua, pero ésta no se distribuye uniformemente lo largo de todo el año, sino que se concentra en algunas épocas del año, para tales zonas se necesitan embalses que regulen el agua de tal manera que se supla la demanda en los meses en los cuales se presenta baja precipitación (cuadrados blancos en la Figura 5).

En la Figura 7 se muestra el esquema resultante para la clasificación por centroides con base a cuatro categorías. Puede observarse que surgen dos grupos, los cuales se denominan como “disponibilidad alta extrema” y “disponibilidad baja extrema”, éstas categorías corresponden a los círculos y cuadrados grises respectivamente en la Figura 7, y dentro ellas se agrupan las observaciones de disponibilidad potencial de recursos hídricos para las cuales se estimaron valores elevados de  $EI$  y  $ER$  y valores bajos  $EI$  y  $ER$ , respectivamente. Adicionalmente se forman dos categorías intermedias de disponibilidad potencial de recursos hídricos, correspondientes a los puntos y cuadrados en blanco en la Figura 7. Dentro de éstas clases se encuentran diversas combinaciones de entropías, como por ejemplo  $EI > 0$  y  $ER < 0$ ,  $EI > 0$  y  $ER > 0$ , para la categoría dos (círculos blancos), y  $IE > 0$  y  $ER > 0$ ,  $EI < 0$  y  $ER < 0$ , para la categoría cuatro (cuadrados blancos). Para tales grupos puede presentarse abundancia o escasez de acuerdo con alguno de los dos tipos de entropía considerados, pero no en extremo para ambas al mismo tiempo. La interpretación con respecto a la disponibilidad para estos grupos puede decirse que es semejante a los cuadrantes intermedios en la clasificación simple.

### 3.3. Consideraciones acerca del tamaño de clase en el histograma de frecuencia de la precipitación mensual

La escogencia del tamaño de las clases en la estimación del histograma de frecuencias para la precipitación mensual tiene efectos en el cálculo de la entropía de intensidad no estandarizada ( $IE$ ). Puede verse que la media de  $EI$  aumenta a medida que aumenta el número de clases, como se observa en la Tabla 1. Por lo anterior se decidió realizar varias clasificaciones variando el valor del tamaño de clase para el histograma. Se realizaron histogramas de 10 clases (200mm por clase), 20 clases (100mm por clase) y 40 clases (50mm por clase), y se elaboraron esquemas parecidos a los mostrados en las Figuras 4 a 7. A pesar de que los valores obtenidos para  $EI$  y para el diagrama de dispersión cambian, las clasificaciones que se obtienen son muy parecidas y consistentes, con respecto a la hipótesis del método del cálculo y clasificación de la entropía comentado en el aparte correspondiente a la metodología aplicada, según la cual la precipitación media anual aumenta a medida que aumenta la entropía de intensidad  $EI$ .

### 3.4. La clasificación binaria y el número de Budyko

La relación  $B=P/PET$  se conoce como el número de Budyko y resume la relación entre la disponibilidad hídrica ( $P$ ) y la disponibilidad energética ( $PET$ ), ésta relación ha sido usada para caracterizar la disponibilidad hídrica de las plantas, la aridez y la productividad neta de los ecosistemas (Arias y Poveda, 2006).

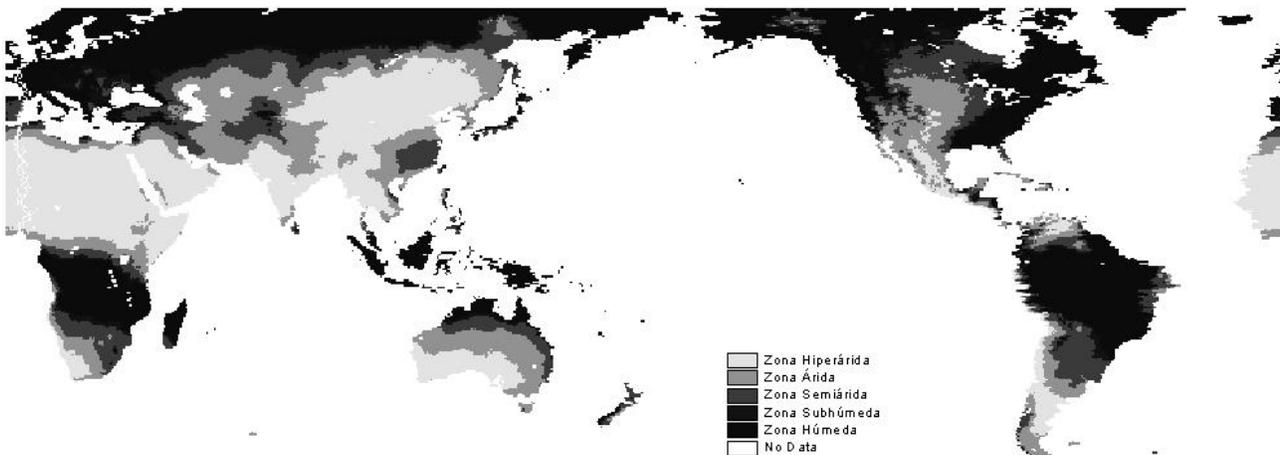
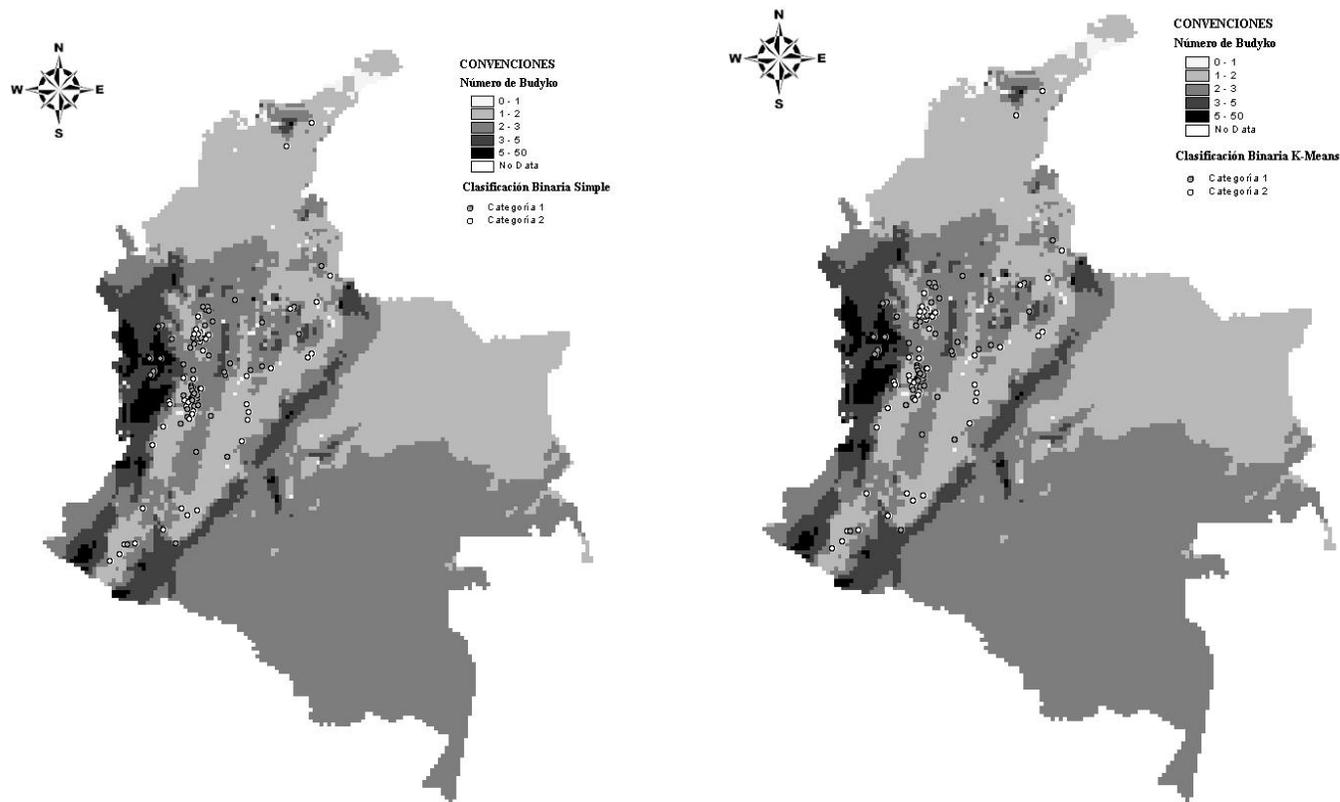


Figura 8. Clasificación para la disponibilidad de agua en el mundo (Tomada de UNESCO, 1997)

UNESCO ha propuesto una clasificación de la disponibilidad de agua alrededor del mundo con base en el número de Budyko (Ver Figura 8), según la cual se delimitan 5 zonas, denominadas: (i) zona húmeda ( $P/PET \geq 0.65$ ), (ii) zona seca subhúmeda ( $0.5 \leq P/PET < 0.65$ ), (iii) Zona semiárida ( $0.2 \leq P/PET < 0.5$ ), (iv) Zona árida ( $0.05 \leq P/PET < 0.2$ ), (v) Zona hiperárida ( $P/PET < 0.05$ ) (UNESCO, 1984). Maruyama et al. (2005),

realizaron un análisis de disponibilidad potencial de recursos hídricos utilizando criterios de entropía a escala mundial y encontraron que para la clasificación binaria simple, las estaciones clasificadas con alta disponibilidad se ubicaron dentro de las zonas húmedas propuestas por UNESCO ( $B > 0.65$ ), mientras que la clasificación binaria obtenida mediante el algoritmo “*k-Means*” delineaba las fronteras de las zonas húmeda y subhúmeda.



**Figura 9.** Mapa de número de Budyko al cual se le han superpuesto las clasificaciones de disponibilidad potencial de recursos hídricos binarias simple (mapa de la izquierda) y *k-Means* (mapa de la derecha) para las estaciones analizadas

En la Figura 9 se muestra un mapa del número de Budyko ( $B=P/PET$ ) obtenido para Colombia mediante álgebra de mapas. Para la estimación de la precipitación media a largo plazo se utilizó el mapa interpolado mediante kriging con deriva externa del ENA, mientras que para la evaporación potencial se utilizó un mapa obtenido mediante la ecuación de Cenicafé. Ambos mapas se extractaron de la base de datos de HidroSIG. (ver Poveda et al., 2006 a y b).

Sobre el mapa del número de Budyko para Colombia se ha superpuesto la clasificación binaria de la disponibilidad potencial de recursos hídricos mediante los métodos de clasificación simple y “*k-Means*”. De la Figura 9 puede verse que la mayoría de las estaciones que presentan deficiencias en disponibilidad potencial de

recursos hídricos, (círculos blancos), se ubican dentro de las zonas en las cuales se ha estimado que el número de Budyko es menor que uno (zonas de gris claro), mientras que las estaciones para las cuales se estima que no presentan deficiencias en disponibilidad potencial de recursos hídricos (círculos grises) se encuentran en su gran mayoría, en zonas para las cuales el número de Budyko es mayor que uno.

#### 4. CONCLUSIONES

Los cálculos de las entropías de intensidad y distribución, junto con la implementación computacional de algoritmos de clasificación son útiles para obtener zonas características con respecto a la disponibilidad potencial de recursos hidráulicos. Con base en éste cálculo, junto

con la proyección de la demanda del recurso se pueden desarrollar estrategias de manejo y regulación del recurso hídrico en un país con tanta variabilidad climática como Colombia. Adicionalmente, un análisis como el usado aquí puede ayudar a los climatólogos, meteorólogos e hidrólogos a clasificar de zonas climáticas representativas, ya sea a partir de registros puntuales o a partir de series de mapas de precipitación o variables relacionadas con la actividad de las plantas, como por ejemplo el NDVI.

Los resultados obtenidos mediante la clasificación binaria para los algoritmos de clasificación considerados en el análisis, son una muestra cuantitativa de que la clasificación de la disponibilidad potencial de recursos hidráulicos para una zona da indicios acerca de sus condiciones climáticas medias de largo plazo. Aún más, el cálculo y la clasificación de la disponibilidad del agua con base en la entropía, surge como un indicador inicial acerca de las características más generales del balance de la disponibilidad hídrica y energética si se relaciona con el número de Budyko.

Por otro lado, el cálculo de la disponibilidad potencial de recursos hidráulicos mediante el uso de la entropía informacional de Shannon depende fuertemente de la información contenida dentro del registro de las estaciones de climatológicas tenidas en cuenta para ser utilizadas en el análisis, por lo cual es importante estudiar la calidad de la información disponible. Resulta obvio el hecho de que considerar estaciones con registros cortos de precipitación o con información de poca calidad puede sesgar los resultados de la estandarización de la entropía y, por lo tanto, la clasificación de la disponibilidad del recurso podría modificarse sustancialmente. Mas aún, disponer de un número considerable de estaciones que se encuentren bien distribuidas dentro del dominio espacial que se está estudiando, contribuye a obtener una mejor caracterización espacial de la disponibilidad del recurso hídrico.

El procedimiento de clasificación de la disponibilidad potencial del recurso hídrico mediante los algoritmos que se tratan en éste estudio no deja de ser subjetivo, tanto por la escogencia del número de clases, como por la escogencia de la distribución de las mismas. Como perspectiva de trabajo futuro podría considerarse el uso de la clasificación no supervisada mediante algunos de los algoritmos existentes como por ejemplo “*Fuzzy k-Means*”, los cuales podrían utilizar lógica difusa para caracterizar la pertenencia parcial a las categorías obtenidas mediante la clasificación.

## 5. BIBLIOGRAFIA

- Arias, P. y Poveda G., 2006, Acerca de la Productividad Primaria Neta en Colombia. Revista Academia Colombiana de Ciencias. En imprenta.
- Goovaerts, P., 1999, Geostatistics For Natural Resources Evaluation. Oxford University Press. Applied Geostatistics Series. New York.
- Kapur, J. y Kesavan H., 1992, Entropy Optimization Principles With Applications. Academic Press Inc. San Diego.
- Kawashi, T, Maruyama, T. y Singh, V., 2001, Rainfall Entropy For Delineation Of Water Resources In Japan. Journal of Hydrology 246. pp. 36-44.
- Macqueen, J., 1967, Some Methods For Classification And Analysis Of Multivariate Observations. En: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1: pp. 281-297. En internet: [http://www.elet.polimi.it/upload/matteucc/Clustering/tutorial\\_html/kmeans.html](http://www.elet.polimi.it/upload/matteucc/Clustering/tutorial_html/kmeans.html)
- Maruyama, T. Kawashi, T y Singh, V., 2005, Entropy-Based Assessment And Clustering Of Potential Water Resources Availability. Journal of Hydrology, 309, pp. 104-113.
- Poveda, G, Vélez, J.I., Mesa, O.J., Cuartas, L.A., Barco, O.J., Mantilla, R.I., Mejia, J.F., Hoyos, C.D., Ramirez, J.M., Botero, B.A., Mejia, M.I., Ceballos, L.I., Zuluaga, M.D., Arias, P.A., Giraldo, J.D. y Quevedo, D. I., 2006a, Linking long-term water balances and statistical scaling to estimate river flows along the drainage network of Colombia. Journal of Hydrologic Engineering, ASCE, En imprenta.
- Poveda, G, Mesa, O.J., Velez, J.I., Mantilla, R., Ramirez, J. M., Hernández, O.O., Borja, A. F. y Urzola, J. A., 2006b, HidroSIG: An interactive digital Atlas of Colombia's hydro-climatology, Journal of Hydroinformatics. En imprenta,
- Viera, S. y Moreno, H., 2002, Caracterización Del Ciclo Diurno De Precipitación En Los Andes Tropicales De Colombia : Región Sur. Trabajo dirigido de grado. Universidad Nacional de Colombia Sede Medellín. Facultad de Minas.
- UNESCO, 1984, Climate, Drought and Desertification. Nature Resources 20 (1), pp. 2-8.
- UNESCO, 1997, Learning to Combat Desertification. UNESCO Education Kit on Desertification, En Internet: [http://www.wateryear2003.org/en/ev.php-URL\\_ID=5137&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://www.wateryear2003.org/en/ev.php-URL_ID=5137&URL_DO=DO_TOPIC&URL_SECTION=201.html).