

Regionalización de la cuenca del río Jequetepeque mediante el método de agrupación y el vector regional

Regionalization Jequetepeque River Basin the method of grouping and the Regional Vector

Diana María Guerrero Pérez

MSc en Recursos Hídricos

Instructora Ambiental

dmguerrerop@misena.edu.co

Servicio Nacional de Aprendizaje

SENA Regional Nariño

Centro Internacional de Producción

Limpia - LOPE



Resumen

En este trabajo se analizó la precipitación de 14 estaciones pluviométricas durante el periodo 1966 – 2006 en la cuenca del Río Jequetepeque (Perú), permitiendo la regionalización de la precipitación y la caracterización del régimen de lluvias utilizando el método de agrupación combinado con el análisis de componentes principales; utilizando el método del Vector Regional lograron determinarse patrones de lluvia en zonas de la cuenca donde no existe información. En la zona sur de la cuenca, se observó una precipitación menor (452–512mm) en comparación con la del norte (859 – 930mm), habiendo una clara influencia de la cordillera en ésta región. Esta metodología puede ser aplicable en Colombia dado a la escasez de estaciones hidrológicas con datos faltantes de información.

Palabras claves:

Cuenca río Jequetepeque, precipitación anual, Regional Vector, Clúster.

Abstract

In this work the precipitation of 14 rain gauge was analyzed from 1966 to 2006, in the Jequetepeque River Basin (Peru), allowing the regionalization of the precipitation and characterization of rainfall patterns, using the method of combining group analyzing the main components. It was achieved to determine rainfall patterns in the river basin areas where there is no information using the Regional Vector Method. In the southern of the river basin, lower precipitation (452–512mm) compared to the north (859 – 930mm) was observed, it has been a significant impact of the mountain range in this region. This methodology may be applied in Colombia given the lack of hydrological stations and missing data information.

Key words:

Jequetepeque River Basin, annual precipitation, Regional Vector, Cluster.

Introducción

Uno de los rasgos climáticos más importantes que caracterizan un lugar, es el régimen pluviométrico. Para conocer la periodicidad con que ciertas cantidades de lluvia pueden registrarse en cualquier punto de una determinada región, suele recurrirse al análisis de las cantidades obtenidas a partir de las series de datos disponibles de las estaciones meteorológicas o pluviométricas ubicadas en las distintas zonas y realizar ajustes mediante funciones de distribución apropiadas (Ávila, 2012).

Es así que éste análisis permite generar zonas homogéneas para determinar la variabilidad en el tiempo o de correlación con variables como la altura sobre el nivel del mar u otras meteorológicas. La clasificación de las estaciones pluviométricas en grupos homogéneos según el ciclo anual, conduce a la

obtención de una regionalización hidrológica de la zona, ésta puede ser utilizada para la transferencia de información de sitios con registros a otros, en los que se requieren mediciones pero no existen (Bárdossy & Kumar, 2011). Las metodologías de regionalización involucran dos etapas principales, la identificación de grupos de cuenca hidrológicamente homogéneas y la aplicación de un método regional para cada región homogénea (Ouarda et al, 2008).

Así también, se conoce que la cuenca del río Jequetepeque hace parte de la vertiente del Pacífico del Perú, comprendiendo un área de 4.372 km² y se encuentra bajo la influencia del clima del Pacífico y del Atlántico (López & Girón, 2007). El objetivo general de éste trabajo era la aplicación y evaluación de la técnica de regionalización por el método del clúster y la clasificación de las estaciones

pluviométricas de la cuenca a través de pruebas de homogeneidad y comparando los estadísticos de la muestra: media, varianza, entre otros (Ganancias, 2010), igualmente se utilizó la técnica del vector regional para la ubicación de zonas con semejante comportamiento pluviométrico y generando índices de vectores anuales.

Es necesario aclarar que los datos utilizados proceden de estaciones pluviométricas de la Autoridad Nacional del Agua de Perú (ANA - Perú), de los años 1966 hasta 2006 para completar los datos por el método de la correlación múltiple para 14 estaciones.

1. Área de Estudio

La cuenca del río Jequetepeque está ubicada en la costa Norte del Perú, entre los paralelos 7° 6' y 7° 30' de latitud sur y los meridianos 78° 30' y 79° 40'

longitud oeste del meridiano de Greenwich (Figura. 1). Tal ubicación corresponde a la vertiente occidental de la Cordillera de los Andes y tiene un área total de 698, 200 hectáreas; distribuida entre los departamentos de La Libertad (provincias de Pacasmayo y Chepén) y Cajamarca (provincias de Cajamarca, Contumazá, San Pablo y San Miguel), abarcando un total de seis provincias y 30 distritos (López y Girón, 2007). Los niveles altitudinales varían entre 0 y 4188 msnm, con una accidentada topografía y con rangos de precipitación de 100 a 1800 mm anuales.

Las características de pluviosidad en la cuenca muestran información entre 50mm de precipitación al nivel del reservorio Gallito Ciego y de más de 1000 mm en la zona alta de la cuenca, que marcan la enorme diferencia de humedad ambiental presente en la cuenca. Además de las

datos anuales para 14 estaciones pluviométricas durante el período 1966 - 2006. En la Tabla 1 se presenta la información respectiva para cada estación.

Teniendo en cuenta que los datos obtenidos son de diferentes períodos comprendidos entre 1966 y 2006, se realizó el Análisis Exploratorio de Datos (AED) para la completación y corrección de saltos, y tendencias de los mismos con el fin de obtener información confiable al momento de correr

Los datos de precipitación de la cuenca del río Jequetepeque, fueron obtenidos de la página de la Autoridad Nacional del Agua,



Figura 1. Mapa de ubicación de la cuenca del río Jequetepeque a escala 1:750.000 y ubicación de estaciones pluviométricas de la cuenca media y alta.

tanto el clúster como el vector regional. La información se procesó en Microsoft Excel, para los estadísticos y diagramas se utilizó Infostat ® versión estudiantil.

El AED proporciona métodos sencillos para organizar y preparar los datos, detectar fallos en el diseño y recogida de datos, tratamiento y evaluación de datos ausentes como la

Tabla 1. Información general de estaciones pluviométricas de la Cuenca del Río Jequetepeque.

Estación	ID	LATITUD (m)	LONGITUD (m)	ALTURA (msnm)	LONGITUD (AÑOS)	RANGO DE AÑOS
Chillete	Chillete	737413	201740	850	41	1966 - 2006
Lives	Lives	717223	9216579	2000	41	1966 - 2006
Contumaza	Contumaza	741024	9186972	2452	41	1966 - 2006
Porcon	Porcon	772509	9221848	3000	41	1966 - 2006
Llapa	Llapa	741217	9227535	2798	35	1966 - 1986; 1993 - 2006
Yumagual	Yumagual	774433	9199318	3450	11	1966 - 1976
San Juan	San Juan	776502	9193345	2353	32	1966; 1968; 1971 - 1988; 1990; 1996 - 2006
Hacienda Llagaden	Hacienda Llagaden	755846	9196060	2000	27	1966 - 1968; 1971 - 1988; 1990 - 1995
Monte Grande	Monte Grande	703914	9200965	489	12	1995 - 2006
Quebrada Honda	Quebrada Honda	759686	9236666	3550	32	1966 - 1997
San Miguel	San Miguel	737533	9227552	2590	11	1996 - 2006
San Pablo	San Pablo	739324	9216481	2290	11	1996 - 2006
Tembladera	Tembladera	706099	9198191	450	15	1966 - 1980
Magdalena	Magdalena	759209	9197563	1300	32	1966 - 1980; 1988 - 1993; 1996 - 2006

identificación de casos atípicos (Salvador y Gallardo, 2003). Para este caso, se utilizaron diagramas de caja, boxplot e histogramas utilizando datos anuales, se completó la serie utilizando la correlación múltiple y corrigiendo tendencias y saltos obtenidos del software TREND con los test de Mann-Kendal y Cusum respectivamente, verificados con T de student; teniendo en cuenta que los saltos dividen el grupo de datos en dos submuestras, es necesario corregir la serie utilizando las siguientes fórmulas (Villón, 2011).

$$x'(t) = \frac{X_t - \bar{X}_1}{S_1(X)} * S_2(X) + \bar{X}_2 \quad (1)$$

$$x'(t) = \frac{X_t - \bar{X}_2}{S_2(X)} * S_1(X) + \bar{X}_1 \quad (2)$$

Donde:

$x'(t)$ = valor corregido de saltos.

x_t = valor a ser corregido.

La ecuación (1) se utiliza cuando se deben corregir los valores de la submuestra de tamaño n_1 , y la ecuación (2) si se deben corregir los datos de la submuestra de tamaño n_2 (Villón, 2011).

En el análisis de homogeneidad, utilizando el método del clúster jerárquico, se pretendió encontrar un conjunto de grupos con variables asignadas a los diferentes datos de precipitación, usando la distancia euclidiana y el método de Ward; para la determinación de los grupos homogéneos se tuvo en cuenta (Álvarez, Hotait & Rivera, 2011):

- Partir de un conjunto de N individuos de los que se dispone de una información cifrada por un conjunto de n variables (una matriz de datos de N individuos y n variables).
- Criterio de similaridad para determinar: una matriz de similaridades que permita relacionar la semejanza de los individuos entre sí (Matriz de N individuos * N individuos).
- Seleccionar el algoritmo de clasificación para determinar

la estructura de agrupación de los individuos. d) Especificar la estructura mediante diagramas arbóreos, dendogramas u otros gráficos.

Teniendo en cuenta que al integrar diferentes individuos en un clúster se tiene el riesgo de perder información, es necesario utilizar el análisis de componentes principales, siendo ésta una técnica estadística que permite sintetizar la información, reducir la dimensión de las variables y perder la menor cantidad de información posible (Terradéz, 2003) para este proceso se utilizó Infostat®, igualmente se utilizó el método de Ward que permite unir grupos con menos incremento en la suma de cuadrados generando grupos más homogéneos, para ello se calcula la media de todas las variables en cada conglomerado, luego se calcula la distancia entre cada caso y la media del conglomerado, sumando después las distancias entre todos los casos. Finalmente se

agruparon los conglomerados que generan menos aumentos en la suma de las distancias dentro de cada conglomerado, generando grupos homogéneos y con tamaños similares (de la Fuente, 2011).

En el análisis de homogeneidad utilizando el método del vector regional, con el software Hydraccess, permitió generar índices de vectores a nivel anual durante el período de estudio (1966 - 2006) que representan zonas de semejante comportamiento pluviométrico, para este estudio se utilizó el método de Brunet Moret, debido a que filtra los índices de las estaciones que se alejan demasiado del promedio, dado que éste se ve influenciado por los valores extremos (Espinoza, 2005).

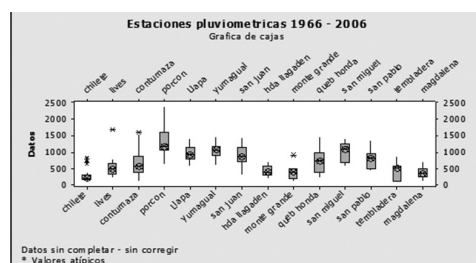
3. Resultados

El procesamiento de los datos se inició con el AED.

En la Figura 2 se presenta un diagrama de caja o box-plot para determinar la distribución de la precipitación de las 14 estaciones pluviométricas de la cuenca del río Jequetepeque en el período de estudio (1966 y 2006), indica que los registros de precipitación en la mayoría de las estaciones tienden a concentrarse en la parte media de la distribución, a excepción de la estación Porcon, igualmente se presentan los valores atípicos en las estaciones Chilete, Lives, Contumazá y Monte Grande. Teniendo en cuenta los datos obtenidos en la Tabla 1, se procedió a completar los datos faltantes por el método de correlación múltiple, utilizando tres y cuatro estaciones base (Chiles, Lives, Contumazá y Porcon). La correlación obtenida para todas las estaciones y el R^2 se presentan en la Tabla 2.

Para los datos obtenidos, se evaluaron las tendencias y saltos utilizando el test de Mann-Kendal y verificados con el test

de T de student; las estaciones que presentaron variaciones, se corrigieron utilizando la metodología empleada por Villón (2011). Teniendo en cuenta que menos del 50% de las estaciones presentan saltos y tendencias, se procedió a realizar la corrección utilizando el método descrito por Villón (2011).



Los asteriscos representan valores atípicos considerando el límite superior en cada estación.

A partir de los resultados obtenidos de la completación y corrección, se generaron los grupos homogéneos tanto para el método del clúster como para el vector regional. En el caso del clúster, inicialmente

Estación	Correlación		
	2 estaciones	3 estaciones	4 estaciones
Chillete	-	-	-
Lives	-	-	-
Contumaza	-	-	-
Porcon	-	-	-
Llapa	-	0.75 - 0.57	-
Yumagual	-	0.69 - 0.43	-
San Juan	-	0.79 - 0.63	-
Hacienda	-	0.71 - 0.60	-
Llagaden	-	0.81 - 0.67	-
Monte Grande	-	-	-
Quebrada	-	-	0.73 - 0.53
Honda	-	-	0.77 - 0.59
San Miguel	-	0.93 - 0.85	-
San Pablo	0.77 - 0.59	-	-
Tembladera	-	-	0.99 - 0.99
Magdalena	-	-	-

Figura 2. Diagramas de caja de la precipitación anual en las 14 estaciones pluviométricas ubicadas en la cuenca del río Jequetepeque comprendido entre los años 1966 - 2006.

se introdujeron las variables en software Infostat® para generar el ACP y el diagrama de aglomeración. El ACP permitió generar componentes para cada uno de los datos introducidos, presentando un resumen de los componentes 1 y 2 para cada variable así: X: 0.74 y -0.59; Y: 0.56 y 0.79; Altura: 0.95 y -0.03; Precipitación 0,88 y 0.03.

Los valores obtenidos a partir del ACP, se introdujeron en el clúster para generar los grupos homogéneos utilizando el método de Ward y la media euclidiana, minimizando la pérdida de información, los grupos generados por éste método determinan las semejanzas que existen entre las variables utilizadas.

Como se observa en la Figura 3, se generaron tres grupos homogéneos determinados de acuerdo con el análisis de los resultados obtenidos en las tablas informativas de la matriz de distancias y la matriz de distancia ultramétrica.

De acuerdo con el dendograma generado, se determinaron las zonas homogéneas en el mapa de ubicación de las estaciones. Se observó que los tres grupos generados están relacionados con la ubicación geográfica de cada estación; en la Figura 4 se presentan los grupos generados en el método del clúster.

Una vez obtenidas las zonas homogéneas por el método del clúster, se procedió a generar con las mismas variables, el vector regional, utilizando Hydracces; este método permitió generar tres grupos representados por vectores con desviación estándar de desvío (DED) de 0.19, 0.23 y 0.31 respectivamente; es importante resaltar que la correlación de las estaciones respecto al vector regional, debe ser mayor a 0.7, determinando que 8 de las 14 estaciones analizadas, no tienen una buena correlación con el vector generado en cada grupo (Tabla 4 y Figura 5).

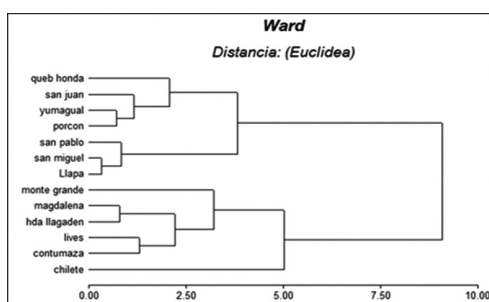


Figura 3. Dendrograma de las estaciones analizadas mediante el método de Ward. Se utilizaron como variables la altitud, precipitación total anual y las coordenadas UTM de cada estación.

ZONA 1			ZONA 2			ZONA 3		
ESTACIÓN	D.E.D	CORREL/ VECTOR	ESTACIÓN	D.E.D	CORREL/ VECTOR	ESTACIÓN	D.E.D	CORREL/ VECTOR
Porcon	0.163	0.276	Contumaza	0.484	0.315	Chilete	0.506	0.631
Llapa	0.159	0.667	Yumagual	0.321	0.299	Lives	0.348	0.734
San Juan	0.159	0.736	Monte Grande	0.473	0.647	Hda Llagaden	0.297	0.477
Qda Honda	0.222	0.569	San Pablo	0.264	0.854	Tembladera	0.424	0.383
San Miguel	0.217	0.685				Magdalena	0.45	0.302

Tabla 4. Correlación de las estaciones con respecto al vector regional generado. Las estaciones con valores menores a 0,70 en la correlación con el vector, no son representados por éste.



Figura 4. Zonas homogéneas mediante el método del clúster jerarquizado con método de conglomeración de Ward y como medida la distancia euclídea. En este caso se muestran tres grupos claramente diferenciados.

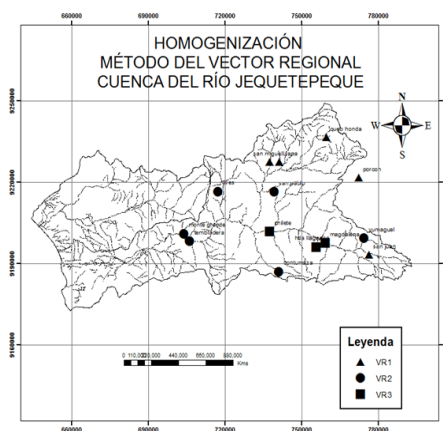


Figura 5. Zonas Homogéneas generadas a partir del método del vector regional. En éste caso, se presentan 3 grupos generados con DED de 0.19, 0.23 y 0.31 respectivamente. VR1: Vector 1, VR2: Vector 2, VR3: Vector 3.

Con base en la homogenización obtenida por los dos métodos, se determinó la producción de información a partir de la interpolación de datos, obteniendo por el método del clúster, una tasa de precipitación para cada zona de: 452 – 512 mm, 859 – 930 mm y 930 – 1014 mm respectivamente. Para el caso de vector regional, se generaron tasas de precipitación así: 860 – 928 mm, 753 – 807 mm y 414 – 489 mm para cada zona respectivamente.

Para el método de aglomeración, en la zona sur de la cuenca (Figura 4), se presenta una precipitación menor (452 – 512 mm) en comparación con la precipitación presentada en la zona norte de la misma (859 – 930 mm), habiendo una clara influencia de la cordillera en esta región.

El vector regional mostró que en la zona 3 existe una influencia de la ubicación geográfica y la altura sobre el nivel del

mar, presentando una menor precipitación (413 – 486 mm) en comparación con la zona 1, que al verse influenciada por la cordillera, presenta una mayor precipitación (859 – 928 mm).

4. Conclusiones

Mediante los dos métodos de homogenización se logró determinar tres regiones homogéneas con comportamiento pluviométrico similar, dos de las cuales no difieren en su distribución general.

El vector regional permitió generar tres grupos homogéneos con DED de 0,19; 0,22 y 0,31, lo que hace necesario una corrección de los dos últimos grupos que determinarían una correcta correlación entre las variables introducidas.

La homogenización por el método del clúster determinó tres grupos que se distribuyeron de acuerdo

con la relación de altura y precipitación.

La homogenización por el método del clúster es de fácil aplicación e interpretación, si se determinan los componentes principales y la similitud entre variables, el método del vector regional requiere un mayor análisis para determinar las zonas homogéneas.

La estación Chilete presentó errores al momento de generar las zonas homogéneas en los dos métodos; para el caso del clúster, se corrigió utilizando el análisis de componentes principales, pero en el vector regional produjo complicaciones en todos los vectores generados, contaminando el resultado del grupo 3.

Las tasas de precipitación generadas están determinadas por la precipitación media obtenida en los 41 años de

análisis: en el método del clúster se generaron tasas de zona 1: 452 - 512 mm, zona 2: 859 - 930 mm y zona 3: 930 - 1014 mm. Para el caso de vector regional, se generaron tasas de precipitación así: zona 1: 859 - 928 mm, zona 2: 412 - 648 mm y zona 3: 827 - 927 mm.

5. Recomendaciones

Teniendo en cuenta las características de la variación de la precipitación debido a las regiones altitudinales, se recomienda la utilización del método de vector regional debido a que utiliza diferentes variables (precipitación, altura, ubicación geográfica, entre otras), que son analizadas en conjunto para determinar zonas con características similares, además de contribuir a la completación de datos de las series hidrológicas, característica muy común en la región Andina colombiana.

6. Agradecimientos

Se agradece al SENA a través de la estrategia SENNOVA de la Regional Nariño, por el acompañamiento en este proceso.

Bibliografía

Ávila Parra, K. (2012). Análisis del comportamiento de eventos extremos de precipitación en la zona centro y sur de Chile continental. Tesis MSc, Universitat de Barcelona, Barcelona, España.

Bárdossy, A. & Kumar, S. (2011). Regionalization of hydrological model parameters using data depth. *Hydrology research*, 42(5): 356 - 371.

Coaguilla, et al. (2013). Análisis comparativo de métodos de interpolación espacial utilizando datos de evapotranspiración. Anais XVI Simpósio de Sensoriamento Remotop - SBR, Foz, do Iguaçu, PR, Brasil, 13 a 18 de abril de 2013, INPE. Pág 4080 - 4085

Espinoza J.C. 2010. El Método del Vector Regional de Índices Pluviométricos. Manual escrito con motivo del curso taller: Utilización del método del Vector Regional con HYDRACCESS. SENAMHI - Perú.

De la Fuente F. S. (2011). Análisis de conglomerados. Universidad Autónoma de Madrid. 57 p.

Salvador F, M y Gargallo, P. (2003): "Análisis Exploratorio de Datos", [en línea] <http://ciberconta.unizar.es/leccion../aed/ead.pdf>, estadística [9 de noviembre de 2013].

Ganancias Martínez, F. (2010). Evaluación de metodologías de regionalización hidrológica: Aplicación a los caudales máximos de cuencas representativas de la región Sur - Oeste de la Provincia de Córdoba. Tesis MSc, Universidad Nacional de Córdoba, España.

López, F & Girón, E. (2007). Pago por servicios ambientales para

generar una nueva dinámica de desarrollo rural en los andes. Informe parte 1: Análisis biofísico (Modelo SWAT). Condesan - Perú.

Ouarda, T. et al (2008). Intercomparison of regional flood frequency estimation methods at ungauged sites for a Mexican case study. *Journal of Hydrology* 340, 40 - 50.

Oviedo Motta, J. et al (2005). Comportamiento agroclimático y pronóstico de lluvias para las cuencas de los ríos Chancay, Lambayeque, Jequetepeque, Rimac, Pampas y Apurímac. Proyecto: Pronóstico de sequías a nivel de cuencas para programas de prevención. 4to Informe Técnico SENAMHI. 33p.

Ríos, I. et al. (2000) Estudio de evaluación ambiental territorial y de planteamientos para la reducción de la contaminación de origen minero en la cuenca del río Jequetepeque. Universidad Nacional de Ingeniería, Lima (Perú).

Secretaría de Agricultura, Ganadería y Desarrollo Rural, Pesca y Alimentación (SAGARPA), (2012). Hidrología aplicada a las pequeñas obras hidráulicas México, 20p.

Terradéz, M. (2003). Análisis de componentes principales. Universitat Oberta de Catalunya. Proyecto E-Math. 11p.

Toro, g. y Melos, C (2009). Aplicación de métodos de interpolación geoestadísticos para la predicción de niveles digitales de una imagen satelital con líneas perdidas y efecto sal y pimienta. *Tecnura* 12 (24): 58-69

Villón V., M. (2011). Hidrología estadística. Instituto Tecnológico de Costa Rica, Escuela de Ingeniería Agrícola. Lima (Perú).