



## Análisis de frecuencia regional de las precipitaciones máximas diarias en la región hidrográfica del Titicaca

Efrain Lujano Laura<sup>a</sup>, Oscar G. Felipe Obando<sup>b</sup>

<sup>A,B</sup>INGENIERO AGRÍCOLA, SERVICIO NACIONAL DE METEOROLOGÍA E HIDROLOGÍA (SENAMHI)  
[elujano28@gmail.com](mailto:elujano28@gmail.com), [ofelipe@senamhi.gob.pe](mailto:ofelipe@senamhi.gob.pe); Dirección postal: Av. Cahuide N° 224 2do Piso, Puno, Perú.

### INFORMACIÓN DEL ARTICULO

Art. Recibido 14/02/15  
Art. Aceptado 22/04/15  
Publicado: 30/04/15

#### PALABRAS CLAVE:

- \* Análisis regional
- \* índice de avenida
- \* momentos-L
- \* conglomerados
- \* regiones homogéneas

### ARTICLE INFO

Article Received 14/03/15  
Article Accepted 22/04/15  
Published:30/04/2015

#### KEY WORDS:

- \* Regional analysis
- \* index flood
- \* L-moments
- \* clusters
- \* homogeneous regions

### RESUMEN

La presente investigación se realizó en la región hidrográfica del Titicaca (lado Peruano), ubicado entre los Departamentos de Puno y Tacna, teniendo como objetivo desarrollar modelos regionales de precipitaciones máximas diarias, aplicado al problema de la estimación de precipitaciones en zonas sin información, cuyos resultados serán empleados en proyectos que requieren este tipo de información, generalmente para el diseño de obras hidráulicas. La información meteorológica, corresponde al Servicio Nacional de Meteorología e Hidrología con un periodo de registro de 1965 al 2013 y la metodología a utilizar consistió en identificar regiones homogéneas cuyos métodos usados se refieren a técnicas multivariadas jerárquicas y momentos-L, dividiendo el área de estudio en tres regiones homogéneas sobre la base de los estadísticos de medida central y variación de la precipitación total mensual. Para el Análisis probabilístico se utilizaron las distribuciones probabilísticas; generalizada de valores extremos, Log-Pearson III, logística generalizada, normal generalizada y Pareto generalizada, de donde las seleccionadas fueron generalizada de valores extremos (región norte), normal generalizada (región centro) y Pearson tipo 3 (región sur). Finalmente a partir de la distribución teórica seleccionada para cada región homogénea, se estimaron los índices de avenida y relativamente se establecieron 09 modelos regionales para diferentes periodos de retorno.

### REGIONAL FREQUENCY ANALYSIS DAILY MAXIMUM RAINFALL IN THE CATCHMENT TITICACA REGION

### ABSTRACT

This research was conducted in the catchment area of Titicaca (Peruvian side) located between the Departments of Puno and Tacna, aiming to develop regional models of daily maximum rainfall, applied to the problem of estimating rainfall in areas without information, whose Results will be used in projects that require such information, generally for the design of hydraulic works. Weather information is for the National Service of Meteorology and Hydrology with a registration period 1965 to 2013 and the methodology used was to identify homogeneous regions whose methods used refer to hierarchical and L-moments multivariate techniques, dividing the study area three homogeneous regions based on central statistical measurement and variation in total monthly precipitation. Probabilistic analysis for the probabilistic distributions were used; generalized extreme value, Log-Pearson III, generalized logistic, generalized normal and generalized Pareto, where selected were generalized extreme value (northern region), generalized average (central region) and Pearson type 3 (southern region). Finally, from the theoretical distribution selected for each homogeneous region, index flood were estimated and 09 regional models relatively different recurrence were established.

## INTRODUCCIÓN

Los eventos extremos de precipitación, ocasionan inundaciones en zonas con baja pendiente, deterioro de estructuras hidráulicas, avería de vías y erosión de suelos agrícolas que se producen por el exceso de precipitaciones en los meses de avenida, provocando daños económicos-sociales y todo ello se generan muchas veces por la mala estimación de estos eventos, extrapolación de datos a zonas sin información, mala estimación de la avenida de diseño y posteriormente un mal dimensionamiento hidráulico. Las breves muestras de máximos anuales generalmente no ofrecen bases sólidas para extrapolar un modelo ajustado a los períodos de retorno mucho más grandes que dos o tres veces el número de años de registro. Si se requieren estimaciones para períodos de retorno más grandes, otros métodos como el análisis de frecuencia regional, son necesarios (Oliveira, 2008). El Servicio nacional de meteorología e hidrología – Puno, administra actualmente 44 estaciones meteorológicas, a nivel de Departamento mostrando una escasez en la distribución espacial, debido a que cada estación local solo puede cubrir un área de 25 km<sup>2</sup> (OMM, 2011). En este contexto la ausencia de estaciones en zonas en las que no existe información meteorológica disponible para la estimación de eventos extremos de precipitación, la regionalización es un tema de gran importancia que toma mayor énfasis en la solución de estimar estos valores en zonas sin información, constituyendo un aporte esencial de prevención y control ocasionados por estos eventos.

Los recientes avances en el análisis de frecuencia regional incluyen el uso de momentos-L junto con el método de la avenida índice, como los informes de Hosking y Wallis (1997). Esta metodología se ha aplicado con éxito en las inundaciones de modelado en una serie de estudios de caso de Estados Unidos (Vogel *et al.*, 1993a), Australia (Pearson *et al.*, 1991; Vogel *et al.*, 1993b), Nueva Zelandia (Madsen *et al.*, 1997), el Sur de África (Mkhandi y Kachroo, 1997) y Sudáfrica (Mkhandi, 1995). El método de análisis regional de la avenida índice constituye una alternativa útil y toma en consideración las deficiencias de los métodos de análisis local, aumentando la longitud de los registros disponibles mediante la transferencia de información de las diferentes estaciones pluviométricas que componen una supuesta región homogénea (Rodríguez y Marreno de León, 2011). Cunnane (1988) y Grehys (1996), en su investigación presentan comparaciones detalladas de las diversas metodologías de estimación para el análisis regional de frecuencias (ARF).

Un análisis regional de frecuencia exhaustivo, requiere de la identificación de regiones homogéneas y para ello Gottschalk (1985) aplicó el análisis de conglomerados y componentes principales, encontrando que el análisis de conglomerados es un método adecuado para su uso. El método es ampliamente utilizado para la clasificación de los diferentes datos climáticos o hidrológicos (Jackson y Weinand, 1995; Nathan y McMahon, 1990; Masoodian, 1998 y Domroes *et al.*, 1998). Hall y Zhang (2004), también recomiendan el empleo del método conglomerados de Ward para identificar regiones homogéneas. El objetivo de la utilización de análisis de conglomerados con las variables hidrológicas son las observaciones de grupos o

variables en grupos sobre la base de la gran similitud de características hidrológicas, como las propiedades geográficas, físicas, estadística o estocástica (Badreldin y Feng, 2012). Hosking y Wallis (1997) sugieren considerar para la formación de regiones homogéneas, características del sitio tales como la localización, la altitud, la precipitación, la estación en la que se producen los eventos máximos, etc.

Por otro lado Grehys (1996) utiliza la técnica de momentos-L en todas las etapas de análisis regional incluyendo la delimitación de región homogénea, identificación y prueba de distribución de frecuencias regionales y la estimación de cuantiles. La experiencia también muestra que, en comparación con momentos convencionales, momentos-L están menos sujetos a sesgo en la estimación (Vogel y Fennessey, 1993; Hosking y Wallis, 1997). Parida y Moalafhi (2008) menciona que al tener una región homogénea los parámetros de una función regional podrían ser transferidas a sitios específicos sin información con confianza. Aunque las conclusiones obtenidas a través de este estudio se basa en los datos disponibles de más de 40 años, que sería bastante fiables debido a la eficacia de los métodos utilizados. No obstante Hosking y Wallis (1997) demostró que el método de momentos-L puede producir resultados muy fiables con tamaños de muestra pequeñas e incluso con valores atípicos. Murray *et al.* (2001), indica que el uso de métodos gráficos con diagramas de relación momentos-L en el proceso de selección de distribución de datos regionales puede ser engañosa. La selección de distribución de datos homogéneos es mejor basado en el promedio de la muestra y no en una línea de mejor ajuste a través de los puntos de datos. Estos resultados destacan la importancia de utilizar las pruebas de heterogeneidad junto con diagramas de relación momentos-L. El objetivo de la presente investigación es desarrollar modelos regionales de precipitaciones máximas diarias en la región hidrográfica del Títicaca.

## MATERIALES Y MÉTODOS

### Área de estudio

La zona de estudio, forma parte de la región hidrográfica del Títicaca y sistema hídrico TDPS, se caracteriza como un sistema de cuenca endorreica, ubicada entre los departamentos de Puno y Tacna. La superficie total es de 49,171.54 Km<sup>2</sup>, el régimen hidrológico es pluvial, presentándose las mayores precipitaciones y descargas en los meses de diciembre a marzo y las temperaturas más bajas concentradas generalmente en los meses de junio y julio. Los límites hidrográficos según la metodología de delimitación y codificación de unidades hidrográficas de Otto Pfafstetter, definen por el norte y este con la región hidrográfica del Amazonas y por el oeste y sur con la región hidrográfica del Pacífico. La figura 1, muestra el ámbito de estudio, ubicado entre las coordenadas UTM (WGS84, zona 19 sur); Este: 595,683 – 284,135, Norte: 7,992633 – 8,447692 y una altitud comprendida entre 3,810 a 5,750 msnm.

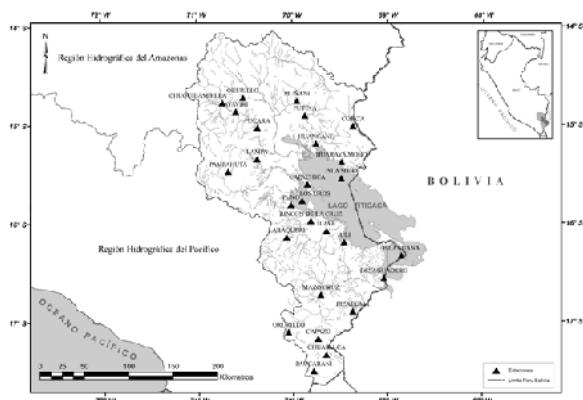


Figura 1. Ubicación del área de estudio

**Datos utilizados**

Información correspondiente a variables de precipitación total mensual y máxima diaria de 27 estaciones ubicadas dentro de la zona de estudio y administradas por el Servicio nacional de meteorología e hidrología. Las características del lugar de ubicación, se muestran en el tablal y la ubicación espacial en la figura 1.

Tabla 1. Características de ubicación de las estaciones meteorológicas

Nº	Estación	Latitud [°]	Longitud [°]	Altitud [msnm]	Nº	Estación	Latitud [°]	Longitud [°]	Altitud [msnm]
1	Ayaviri	-14.87	-70.59	3921	15	Laraqueñi	-16.15	-70.07	3952
2	Capachica	-15.62	-69.84	3827	16	Los Uros	-15.78	-69.9	3825
3	Capaza	-17.19	-69.74	4470	17	Mazocruz	-16.74	-69.71	3963
4	Chuapalca	-17.35	-69.65	4273	18	Muñani	-14.77	-69.95	3923
5	Chuquibambilla	-14.79	-70.73	3899	19	Orurillo	-14.73	-70.52	4048
6	Cojata	-15.03	-69.36	4325	20	Pampahuta	-15.48	-70.68	4292
7	Desaguadero	-16.57	-69.04	3813	21	Puacarani	-17.52	-69.78	4572
8	Huancané	-15.2	-69.75	3860	22	Pizacoma	-16.91	-69.37	3913
9	Huayna Moho	-15.39	-69.48	3818	23	Pucara	-15.04	-70.37	3869
10	Ilave	-16.09	-69.65	3837	24	Puno	-15.82	-70.02	3831
11	Isla Soto	-15.55	-69.49	3820	25	Putina	-14.91	-69.87	3853
12	Isla Suana	-16.34	-68.86	3809	26	Rincón de la Cruz	-15.99	-69.81	3875
13	Juli	-16.2	-69.46	3814	27	Vilacota	-17.12	-70.05	4420
14	Lampa	-15.36	-70.37	3929					

**Tratamiento de datos**

Las muestras de precipitaciones máximas diarias fueron sometidas a la prueba de independencia, homogeneidad e identificación de valores atípicos a través de pruebas no paramétricas como Kendall, Mann-Kendall y Grubbs-Beck respectivamente, cuya aplicación se realizó con el software SEAF (Sistema Especialista de Análisis de Frecuencia), desarrollado por la Universidad Federal de Minas Gerais – Brasil. El criterio principal para seleccionar estas estaciones fue tener disponible como mínimo al menos 15 años de registro histórico, puesto que datos demasiado cortos pueden producir errores de estimación más grande del parámetro y cuantiles, lo cual eventualmente puede inducir a una elección equivocada (Oliveira, 2008).

**Identificación de regiones homogéneas**

**Método de conglomerado**

La similaridad de regiones homogéneas se identificó en primera instancia utilizando el método de Ward. La realización de este análisis, se efectuó a través del software Minitab 16, utilizando los datos de precipitación promedio multimensual y el coeficiente de variación (Cv), eligiendo la máxima homogeneidad. La explicación detallada del algoritmo de Ward se puede encontrar en Rao y Srinivas (2006).

**Método de momentos - L**

Los momentos-L surgen de combinaciones lineales de los momentos ponderados probabilísticamente, introducidos por Greenwood *et al.* (1979) y constituyen un sistema alternativo al método tradicional de los momentos convencionales para describir las funciones de distribución. Los momentos ponderados probabilísticamente, de la variable aleatoria X con una función de distribución F(X), quedan definidos por la expresión:

$$M_{p,k,s} = E(X^p \{F(X)\}^k \{1 - F(X)\}^s) \quad (1)$$

donde E indica el valor de la Esperanza, siendo p, k y s números reales.

Son particularmente útiles los casos en que p = 1 con k = 0, y p = 1 con s = 0, denotados como  $\alpha_r = M_{1,0,r}$  y  $\beta_r = M_{1,r,0}$  respectivamente, siendo r = 0, 1, 2,...

Los primeros cuatro momentos-L, se expresan como combinaciones lineales de probabilidad ponderada. (Hosking y Wallis, 1997):

$$\lambda_1 = \alpha_0 = \beta_0 \quad (2)$$

$$\lambda_2 = \alpha_0 - 2\alpha_1 = 2\beta_1 - \beta_0 \quad (3)$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 + \beta_0 \quad (4)$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad (5)$$

El momento lineal de primer orden  $\lambda_1$  es el parámetro de localización o media de la distribución; el de segundo orden  $\lambda_2$  mide la escala de la distribución, indicando el grado de dispersión de los datos; el momento de tercer orden  $\lambda_3$  hace referencia a su asimetría; y el de cuarto orden  $\lambda_4$  indica la curtosis (apuntamiento) de la distribución.

No obstante, para poder comparar diferentes estaciones resulta imprescindible definir los momentos lineales de forma adimensional. Por ello, se calculan los ratios o cocientes de momentos lineales, dividiendo aquellos de mayor orden por la medida de escala  $\lambda_2$  (Hosking y Wallis, 1997):

$$\tau_r = \frac{\lambda_r}{\lambda_2}, r = 3,4,\dots \quad (6)$$

dando así una idea de la forma de la distribución con independencia de su escala de medida. Adquieren particular importancia los L-coeficientes de variación, asimetría y curtosis:

$$L - CV = \tau = \frac{\lambda_2}{\lambda_1} \quad (7)$$

$$L - asimetria = \tau_3 = \frac{\lambda_3}{\lambda_2} \quad (8)$$

$$L \text{ curtosis} = \tau_4 = \frac{\lambda_4}{\lambda_2} \quad (9)$$

Se puede afirmar que conceptualmente, los momentos-L son semejantes a los momentos convencionales. Sin embargo, Hosking y Wallis (1997) constatan que son varias las ventajas de los primeros desde el punto de vista teórico: capacidad para caracterizar un mayor rango de funciones de distribución, mayor robustez en presencia de valores atípicos y menor sensibilidad a la asimetría, entre otras.

**Medidas de discordancia**

Según lo explicado por Hosking y Wallis, (1997) si hay N sitios de un grupo de estaciones, el objetivo es identificar aquellos sitios que son extremadamente discordante con el grupo en su conjunto. La discordancia se mide en términos de los momentos-L de los datos de los sitios. Numéricamente la medida de discordancia  $D_i$  se define a partir de un estadístico.

Para un conjunto de datos con los sitios de N, la medida discordancia  $D_i$  para el sitio  $i$  se define por:

$$D_i = \frac{1}{3(N-1)} (u_i - \bar{u})^T S^{-1} (u_i - \bar{u}) \quad (10)$$

Un sitio es declarado discordante si  $D_i \geq 3.0$ , las estaciones individuales que superen este umbral deben ser sometidas a un minucioso análisis de calidad en busca de las posibles causas que diferencien a éstas del resto del grupo.

**Prueba de heterogeneidad**

Hosking y Wallis (1997), indica que se utiliza para estimar el grado de coherencia en un grupo de estaciones y evaluar cuándo éstas pueden ser tratadas en su conjunto como una región homogénea. En particular, la medida de heterogeneidad H, compara la variabilidad de los L momentos muestrales del grupo de estaciones que conforman la región con la esperada, obtenida a partir del empleo de técnicas de simulación. La región se declara heterogénea si el valor de H es suficientemente elevado. Hosking y Wallis (1997) sugieren considerarla como «aceptablemente homogénea» si  $H < 1$ , «posiblemente heterogénea» si  $1 \leq H < 2$ , y «definitivamente heterogénea» si  $H \geq 2$ . Valores negativos indican una dispersión de los L-CV muestrales menor que la esperada en caso de región homogénea con distribuciones locales independientes, siendo

la existencia de correlación positiva entre estaciones su causa más probable.

La prueba de heterogeneidad empleada está basada en el L-coeficiente de variación (L-Cv). Ésta usa la varianza ponderada del L-Cv de cada una de las estaciones de acuerdo con el estadístico:

$$V = \frac{I}{N} \cdot \frac{\sum_{i=1}^N n_i (t^{(i)} - \bar{t})^2}{\sum_{i=1}^N n_i} \quad (11)$$

La medida de heterogeneidad H se define cómo:

$$H = \frac{(V - \mu_v)}{\sigma_v} \quad (12)$$

donde  $\mu_v$  es la media y  $\sigma_v$  la desviación estándar.

**Selección de la distribución regional de frecuencia**

Para el ajuste de funciones de distribución en base a momentos-L, se utilizaron la distribución Logística generalizada (GLO), Generalizada de valores extremos (GEV), normal generalizada (GNO), Pearson tipo III (PE3) y Pareto Generalizado (GPA).

**Medida de la bondad de ajuste**

La bondad de ajuste juzgará en qué medida los momentos L-Cs y L-Ck de la distribución seleccionada se ajustan bien al promedio regional de L-Cs y L-Ck de los datos observados. La medida de la bondad de ajuste para una distribución de tres parámetros propuesta por Hosking y Wallis (1993) se expresa como:

$$Z^{DIS} = \frac{(t_4^R - \tau_4^{DIS})}{\sigma_4} \quad (13)$$

Se considera que el ajuste de una determinada distribución es adecuado si el estadístico  $Z^{DIST}$  es suficientemente cercano a cero, siendo un valor razonable para este criterio el grado de significación del 90%, que corresponde a  $|Z^{DIST}| \leq 1.64$ .

Donde  $t_4^R$  es, L-curtosis promedio regional,  $\tau_4^{DIS}$  L-curtosis de la distribución adaptada y  $\sigma_4$  desviación estándar promedio regional.

**Índice de avenida**

Propuesto por Darymple (1960) y establecido como un procedimiento formal del U.S. Geological Survey. A partir de la distribución teórica seleccionada para cada región homogénea se obtienen las precipitaciones máximas asociados a periodos de retorno de 2, 5, 10, 25, 50, 100, 500, 1000, 5000

y 1000 años cada uno de los cuales se dividen entre la precipitación máxima media (OMM, 2011) considerado como índice de avenida.

## RESULTADOS Y DISCUSIONES

### Prueba de independencia, homogeneidad y de valores atípicos

Todas las muestras han pasado las pruebas de independencia de Kendall y homogeneidad de Mann-Kendall, resultando datos independientes (DI) y homogéneos (DH), con un nivel de confianza de 95% y nivel de significancia de 5% (ver tabla 2). En cuanto a la presencia de valores atípicos, la prueba de

Grubbs y Beck han identificado la existencia de valores atípicos (HVA) en 06 estaciones, sin embargo, no fueron retirados de la serie de valores máximos porque no se encontraron razones objetivas para descartarlos, esto está de acuerdo con las recomendaciones de Cunnane (1989) en el tratamiento de los valores atípicos de series máximos anuales. La tabla 2 muestra los resultados de las pruebas de significancia de Grubbs y Beck, lo cual se contrasta los valores máximos y mínimos de la serie con los umbrales altos (Ua) y bajos (Ub), no existiendo valores atípicos (NHVA) en la mayoría de estaciones debido a que los valores de la serie se encuentran dentro de los umbrales establecidos.

Tabla 2. Prueba de significancia de Kendall, Mann-Kendall, Grubbs y Beck

N°	Muestra	Kendall				Man – Kendall		Grubbs y Beck (Ua)			Grubbs y Beck (Ub)			
		r	Z	Z	y  U	Alfa	Z <  Z	U	U >  U	[Ua]	Max	Max < Ua	Ub	Min
1	Cojata	-0.12	1.193	1.96	0.05	DI	0.776	DH	61.3	61.1	NHVA	14.7	19.2	NHVA
2	Putina	0.112	1.154	1.96	0.05	DI	0.411	DH	64.2	52.2	NHVA	16.1	16.7	NHVA
3	Huancané	0.086	0.854	1.96	0.05	DI	1.759	DH	80.3	68.7	NHVA	17.6	20.6	NHVA
4	Muñani	-0.115	1.118	1.96	0.05	DI	0.303	DH	50.7	45.4	NHVA	15	17.2	NHVA
5	Huaraya Moho	-0.068	0.676	1.96	0.05	DI	1.138	DH	84.1	74.3	NHVA	21.1	25	NHVA
6	Pucara	-0.043	0.418	1.96	0.05	DI	0.716	DH	73.6	76.4	NHVA	15.8	20.4	NHVA
7	Chuquibambilla	-0.065	0.596	1.96	0.05	DI	0.272	DH	65	59.2	NHVA	15.2	15.6	NHVA
8	Orurillo	0.253	1.665	1.96	0.05	DI	0.645	DH	58	55.1	NHVA	16.9	19.8	NHVA
9	Ayaviri	0.135	1.198	1.96	0.05	DI	1.41	DH	63	52.2	NHVA	15.1	16	NHVA
10	Lampa	-0.083	0.805	1.96	0.05	DI	0.119	DH	67.8	64	NHVA	17.5	19	NHVA
11	Pampahuta	-0.005	0.041	1.96	0.05	DI	1.526	DH	58.3	49.9	NHVA	20.8	22.9	NHVA
12	Laraqueri	0.078	0.787	1.96	0.05	DI	0.52	DH	65.5	54.8	NHVA	16.4	21.6	NHVA
13	Capachica	0.007	0.056	1.96	0.05	DI	0.824	DH	72.1	63.2	NHVA	16.7	22.3	NHVA
14	Juli	-0.032	0.281	1.96	0.05	DI	0.52	DH	86.7	100.8	HVA	19.3	25.4	NHVA
15	Los Uros	-0.092	0.831	1.96	0.05	DI	0.033	DH	63.2	57.1	NHVA	16.6	18.2	NHVA
16	Puno	0.185	1.871	1.96	0.05	DI	0.77	DH	76.7	78.2	HVA	17.4	23.6	NHVA
17	Rincón de la Cruz	-0.069	0.494	1.96	0.05	DI	0.882	DH	55.2	51	NHVA	19	23.2	NHVA
18	Desaguadero	0.053	0.511	1.96	0.05	DI	1.944	DH	83.6	71.8	NHVA	20.8	25	NHVA
19	Ilave	-0.133	0.994	1.96	0.05	DI	0.642	DH	52.5	46	NHVA	19	21	NHVA
20	Isla Suana	-0.103	0.661	1.96	0.05	DI	0.496	DH	71.5	75.6	HVA	21.9	26.5	NHVA
21	Isla Soto	-0.142	0.925	1.96	0.05	DI	0.943	DH	85.3	104.5	HVA	24.5	32.1	NHVA
22	Capazo	0.07	0.665	1.96	0.05	DI	1.316	DH	50.9	48.4	NHVA	12.9	16	NHVA
23	Paucarani	-0.059	0.529	1.96	0.05	DI	0.423	DH	54.9	38.3	NHVA	9.1	8.5	HVA
24	Mazocruz	-0.067	0.651	1.96	0.05	DI	0.018	DH	55.1	47.4	NHVA	12.9	13.4	NHVA
25	Chuapalca	0.167	1.489	1.96	0.05	DI	0.793	DH	50	54	HVA	13.8	16.3	NHVA
26	Pizacoma	0.265	1.442	1.96	0.05	DI	0	DH	64.2	61.2	NHVA	19.5	23.2	NHVA
27	Vilacota	0.123	1.009	1.96	0.05	DI	0.284	DH	59.1	41.2	NHVA	10.6	12	NHVA

**Identificación de regiones homogéneas**

La consistencia de muestras descritas anteriormente, dio como resultado la existencia de 27 muestras independientes y homogéneas, los cuales fueron utilizados para la identificación de regiones homogéneas, para lo cual se utilizaron los estadísticos de tendencia central (media) y medidas de variabilidad (coeficiente de variación) respectivamente, estimados a partir de los datos de precipitación total mensual con periodo de registro de 1956 al 2013. El coeficiente de variación y la media, han permitido ver la variabilidad y las características meteorológicas propias de una región, siendo variables muy necesarias para identificar regiones homogéneas.

**Método de conglomerado**

La figura 2, muestra el resultado de la formación de 03 regiones con un comportamiento similar denominado región norte y conformado por las estaciones Ayaviri, Lampa, Pucara, Chuquibambilla, Orurillo, Cojata, Putina, Huancané, Muñani y Huaraya Moho, la región centro conformado por Capachica, Juli, los Uros, Puno, Rincon de la Cruz, Laraqueri, Pampahuta, Desaguadero, Ilave, Isla Suana e Isla Soto y la region sur, conformados por Capazo, Paucarani, Mazocruz, Chuapalca, Pizacoma y Vilacota. La idea de la identificación de regiones homogéneas mediante este método son basados en las recomendaciones de Gottschalk (1985), Jackson y Weinand (1995), Nathan y McMahon (1990), Hosking y Wallis (1997), Hall y Zhang (2004) dado que es un método jerárquico que no requiere del conocimiento anticipado de un grupo homogéneo.

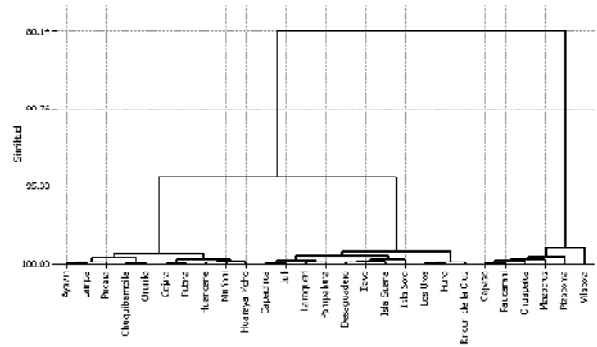


Figura 2. Dendrograma de Ward, distancia del coeficiente de correlación

**Método de momentos - L**

Las regiones preliminares identificadas por el algoritmo de Ward, son controlados por pruebas basadas en dispersión L-Cv, L-asimetría, L-curtosis, pruebas de discordancia y heterogeneidad. La tabla 3, muestran las estadísticas de momentos-L (Media, L-coeficiente de variación, L-asimetría y L-curtosis), mientras que en la figura 3 se puede apreciar los diagramas de dispersión de las relaciones L-CV con L-curtosis y L-asimetría con L-curtosis, deduciéndose que ninguna estación se encuentra alejada de manera significativa.

Tabla 3. Estadísticas L-coeficiente de variación, asimetría y curtosis

Región	Nº	Muestra	Tamaño de la muestra	Media [l <sub>1</sub> ]	L-CV [l <sub>2</sub> ]	L-asimetría [l <sub>3</sub> ]	L-curtosis [l <sub>4</sub> ]
Norte	1	Ayaviri	40	31.9	0.1504	0.0943	0.0797
	2	Chuquibambilla	43	32.6	0.1534	0.2844	0.1806
	3	Cojata	50	31.8	0.1653	0.2913	0.2083
	4	Huancané	49	39.0	0.1526	0.0867	0.1343
	5	Huaraya Moho	49	43.4	0.1431	0.1107	0.0940
	6	Lampa	47	35.5	0.1387	0.1232	0.1918
	7	Muñani	47	28.2	0.1259	0.0959	0.1553
	8	Orurillo	24	32.2	0.1463	0.1801	0.1300
	9	Pucara	49	35.5	0.1604	0.2192	0.2240
	10	Putina	52	33.1	0.1346	0.0552	0.1584
Centro	1	Capachica	42	36.0	0.1594	0.2944	0.1475
	2	Desaguadero	47	43.0	0.1443	0.1086	0.0914
	3	Isla Soto	24	47.4	0.1466	0.3296	0.2582
	4	Isla Suana	24	40.8	0.1396	0.2850	0.2398
	5	Ilave	30	32.2	0.1143	0.0683	0.0692
	6	Juli	42	42.6	0.1631	0.3522	0.2477
	7	Laraqueri	51	33.8	0.1461	0.1743	0.0697
	8	Los Uros	42	33.4	0.1407	0.1054	0.1096
	9	Pampahuta	45	35.5	0.1070	0.0897	0.1414
	10	Puno	50	38.0	0.1555	0.2694	0.2378
	11	Rincón de la Cruz	29	33.1	0.1228	0.1960	0.1281
Sur	1	Capazo	46	26.4	0.1472	0.2620	0.1561
	2	Chuapalca	40	27.0	0.1351	0.1307	0.1907
	3	Mazocruz	48	27.6	0.1472	0.0999	0.1349
	4	Paucarani	41	23.6	0.1758	0.0345	0.0744
	5	Vilacota	35	26.2	0.1774	0.0165	0.0524
	6	Pizacoma	18	36.5	0.1491	0.1396	0.0957

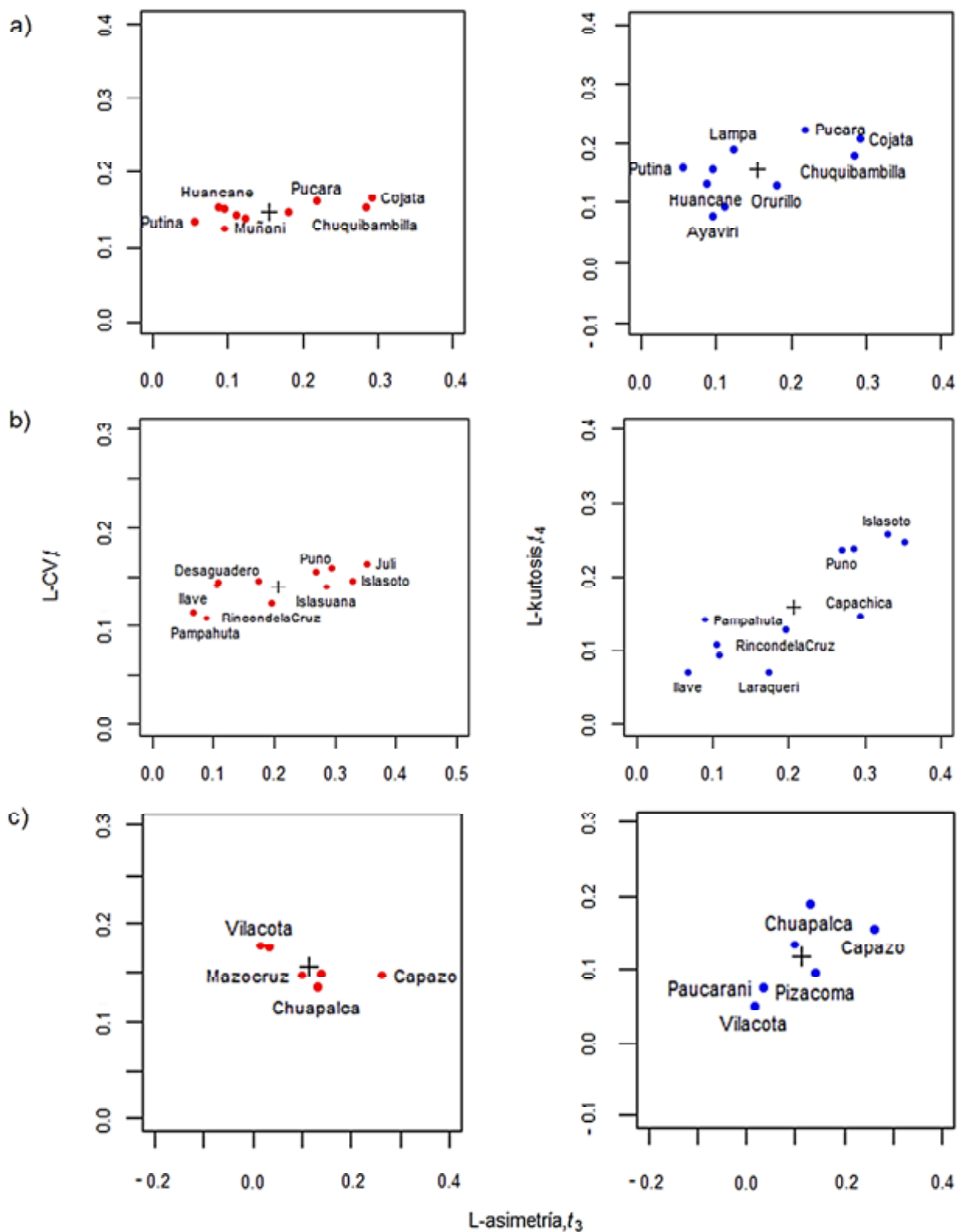


Figura 3. Relaciones de momentos-L región, a) norte, b) centro y c) sur

**Medidas de discordancia**

De la tabla 4, se deduce que para las regiones norte, centro y sur, según la cantidad de estaciones se tienen una discordancia crítica (Dc) igual a 2.49, 2.76 y 1.65 respectivamente, siendo para cada estación sus valores de discordancia inferiores a los umbrales críticos. Los resultados obtenidos son adecuados según lo descrito por Hosking y Wallis, 1997.

Tabla 4. Medidas de discordancia

N°	Norte: Dc=2.49		Centro: Dc=2.63		Sur: Dc=1.65	
	Muestra	Di	Muestra	Di	Muestra	Di
1	Ayaviri	1.26	Capachica	1.18	Capazo	1.63
2	Chuquibambilla	1.27	Desaguadero	1.05	Chuapalca	1.24
3	Cojata	1.11	Isla Soto	0.76	Mazocruz	0.29
4	Huancané	1.07	Isla Suana	0.58	Paucarani	0.75
5	Huaraya Moho	0.69	Ilave	0.88	Vilacota	0.69
6	Lampa	0.66	Juli	0.79	Pizacoma	1.41
7	Muñani	1.42	Laraqueri	1.13	-	-
8	Orurillo	0.47	Los Uros	1.03	-	-
9	Pucara	1.22	Pampahuta	1.48	-	-
10	Putina	0.84	Puno	1.00	-	-
11	-	-	Rincón de la Cruz	1.12	-	-

**Prueba de heterogeneidad**

Se establecen tres regiones homogéneas, con H-valores de -1.10 (norte), 0.24 (centro) y -0.32 (sur), los cuales satisfacen la condición de 'aceptablemente homogénea' ( $H < 1$ ), es decir, las estaciones de cada grupo tienen distribuciones de frecuencia idénticas, aparte de un factor de escala específica. La tabla 5, muestra la prueba de heterogeneidad, cuyos resultados son aceptables a lo descrito por Hosking y Wallis, 1997. La figura 5 muestra la espacialización del grupo de regiones homogéneas en el ámbito de estudio.

Tabla 5 Medidas de heterogeneidad ( $H_1$ ,  $H_2$  y  $H_3$ ) para las tres regiones

N°	Región	H1	H2	H3
1	Norte	-1.10	0.96	0.57
2	Centro	0.14	1.85	1.82
3	Sur	0.10	0.91	0.78

Los valores de  $H_1$  son inferiores al umbral de homogeneidad ( $H < 1$ )

**Elección de la distribución de frecuencia**

Se encontró una buena correspondencia entre los resultados de la prueba de bondad de ajuste y el diagrama de momentos-L, por lo tanto, tenemos para la región norte las distribuciones de GLO, GEV, GNO y PE3 como distribuciones regionales válidas y se rechaza la distribución GPA. En la región centro y sur, son aceptadas las distribuciones GEV, GNO y PE3, rechazando las distribuciones GLO y GPA (ver tabla 6 y figura 4). Sin embargo, las distribuciones que pueden describir de manera eficiente los datos observados de cada región homogénea, estaría dada por el menor valor de  $|Z^{DIST}|$ , entre ellos GEV, GNO y PE3, para las regiones norte, centro y sur respectivamente. La tabla 6 muestra los valores del estadístico Z, los valores sombreados en color gris son las recomendadas para cada región, mientras que la figura 4, el valor medio entre L-curtosis y L-asimetría estaría representado por el símbolo de un triángulo y sería la media regional.

Tabla 6 prueba de bondad de ajuste ( $|Z^{DIST}| \leq 1.64$ )

N°	Región	GLO	GEV	GNO	PE3	GPA
1	Norte	1.49	-0.85	-1.01	-1.61	-5.93
2	Centro	2.3	0.47	0.03	-0.88	-3.77
3	Sur	2.43	0.42	0.49	0.21	-3.73



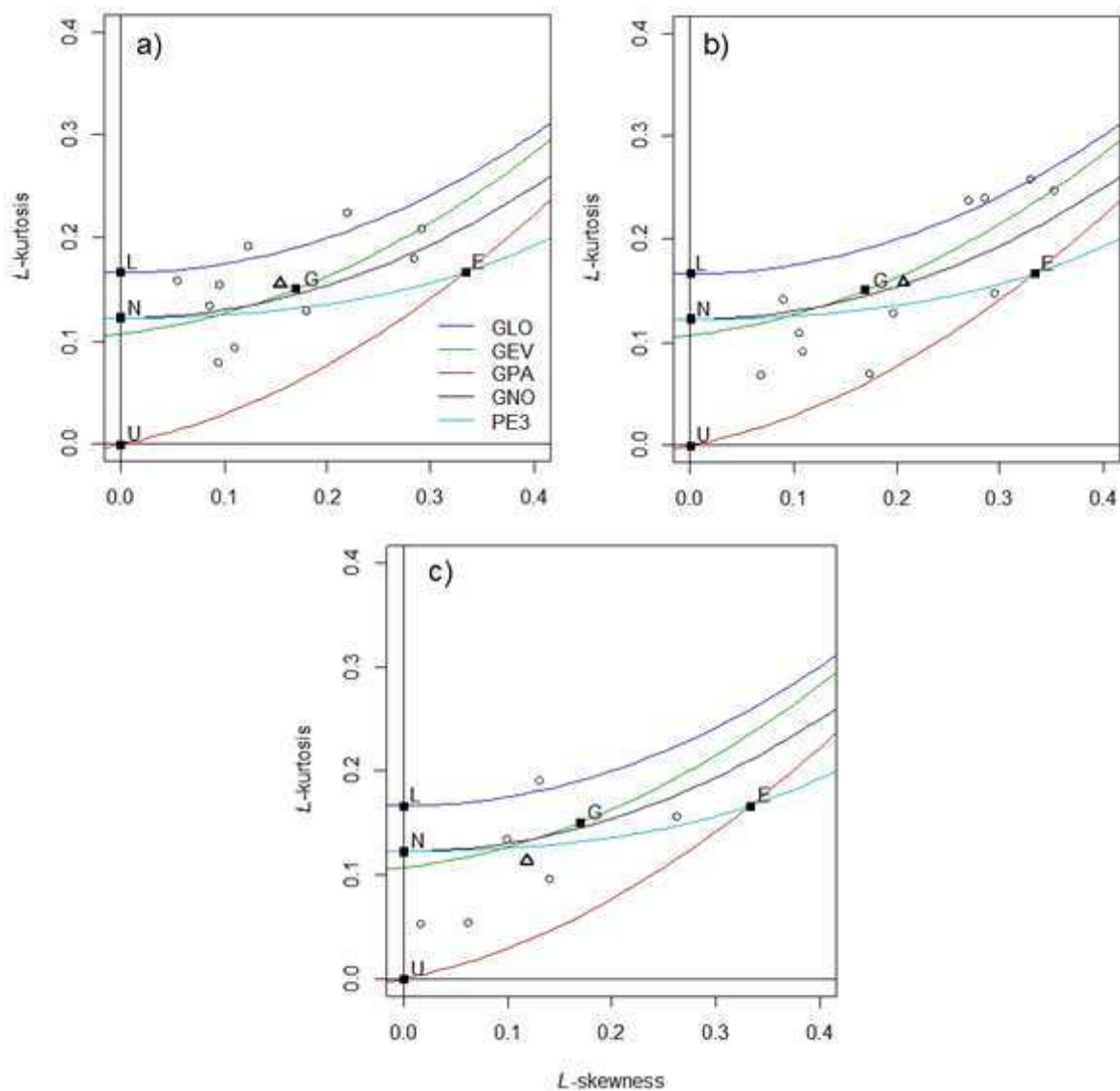


Figura 4. Diagrama de L-momentos de las distribuciones, a) región norte, b) región centro y c) región sur

#### Estimación de cuantiles regionales e índice de avenida

Obtenido la función de distribución para cada región homogénea, se estimaron los valores de precipitación para distintos períodos de retorno y dividiendo estos valores extremos por la precipitación máxima media anual ( $\bar{p}$ ) se obtuvieron los cuantiles regionales.

El índice de avenida espacial fue estimado mediante una regresión multivariada entre la precipitación máxima media ( $\bar{p}$ ), latitud, longitud y altitud para toda la región hidrográfica del Titicaca lado Peruano (figura 5), lo cual se multiplica con los cuantiles adimensionales mostrado en la tabla 7, cuyo gráfico contra los periodos de retorno representa la proporción más probable de precipitación máxima en la región homogénea.

Tabla 7. Cuantiles regionales y límites de confianza para diferentes periodos de retorno

Probabilidad	Tr. [años]	Región norte			Región centro			Región sur		
		Cuantil regional	Límite		Cuantil regional	Límite		Cuantil regional	Límite	
			Inf.0.05	Sup.0.95		Inf.0.05	Sup.0.95		Inf.0.05	Sup.0.95
0.5	2	0.96	0.92	0.99	0.95	0.90	0.99	0.97	0.92	1.00
0.8	5	1.20	1.17	1.23	1.18	1.15	1.21	1.22	1.18	1.28
0.9	10	1.35	1.30	1.40	1.34	1.25	1.41	1.37	1.32	1.43
0.95	20	1.50	1.40	1.62	1.50	1.33	1.66	1.51	1.43	1.59
0.98	50	1.68	1.51	1.96	1.70	1.43	2.02	1.67	1.56	1.83
0.99	100	1.82	1.57	2.27	1.86	1.49	2.34	1.79	1.66	2.01
0.995	200	1.95	1.62	2.61	2.02	1.55	2.69	1.90	1.74	2.19
0.998	500	2.12	1.69	3.14	2.24	1.63	3.22	2.04	1.84	2.43
0.999	1000	2.25	1.72	3.60	2.41	1.68	3.67	2.15	1.91	2.62

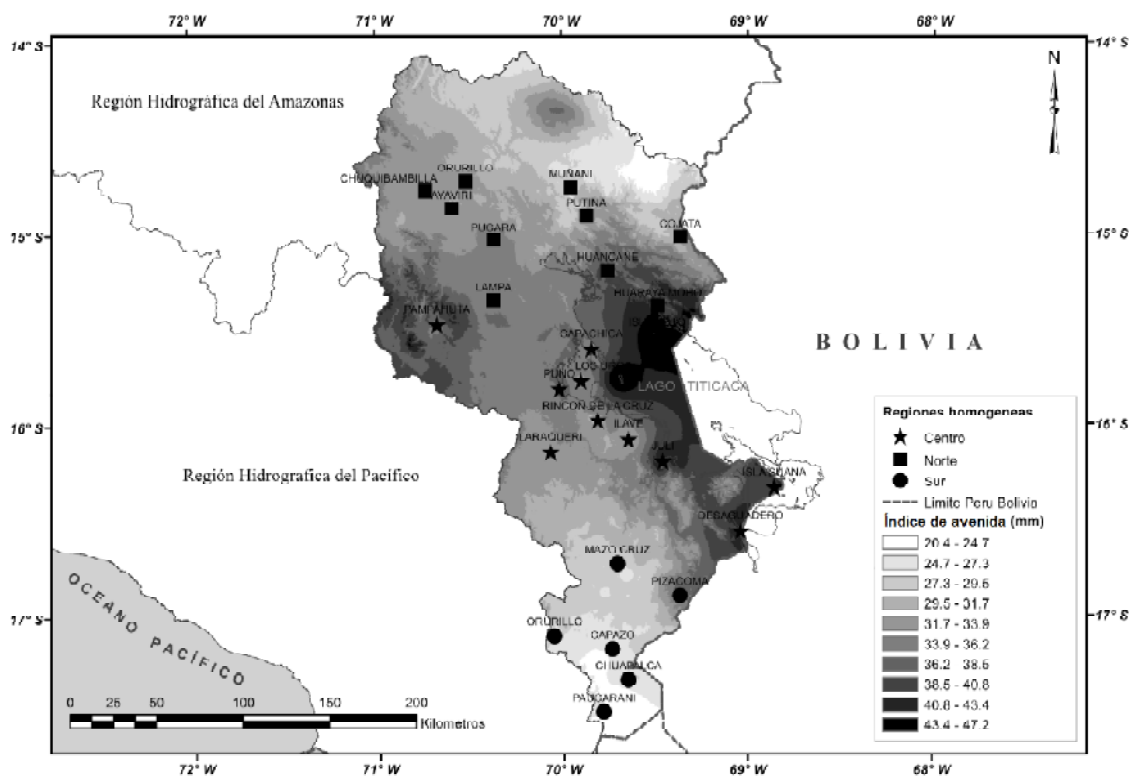


Figura 5. Distribución espacial de regiones homogéneas e índice de avenida

**Modelos regionales**

Se establecieron 09 modelos regionales basados en el índice de avenida () asociados a diferentes periodos de retorno, los cuales serán de utilidad para estimar precipitaciones máximas en sitios sin información. La tabla 8 muestra los modelos regionales establecidos para cada región homogénea.

Tabla 8. Modelos regionales basados en el índice de avenida

Probabilidad	Tr. [años]	Modelos regionales		
		Región norte	Región centro	Región sur
0.5	2	$P_{Tr2años} = 0.96^* \bar{P}$	$P_{Tr2años} = 0.95^* \bar{P}$	$P_{Tr2años} = 0.97^* \bar{P}$
0.8	5	$P_{Tr5años} = 1.20^* \bar{P}$	$P_{Tr5años} = 1.18^* \bar{P}$	$P_{Tr5años} = 1.22^* \bar{P}$
0.9	10	$P_{Tr10años} = 1.35^* \bar{P}$	$P_{Tr10años} = 1.34^* \bar{P}$	$P_{Tr10años} = 1.37^* \bar{P}$
0.95	20	$P_{Tr20años} = 1.50^* \bar{P}$	$P_{Tr20años} = 1.50^* \bar{P}$	$P_{Tr20años} = 1.51^* \bar{P}$
0.98	50	$P_{Tr50años} = 1.68^* \bar{P}$	$P_{Tr50años} = 1.70^* \bar{P}$	$P_{Tr50años} = 1.67^* \bar{P}$
0.99	100	$P_{Tr100años} = 1.82^* \bar{P}$	$P_{Tr100años} = 1.86^* \bar{P}$	$P_{Tr100años} = 1.79^* \bar{P}$
0.995	200	$P_{Tr200años} = 1.95^* \bar{P}$	$P_{Tr200años} = 2.02^* \bar{P}$	$P_{Tr200años} = 1.90^* \bar{P}$
0.998	500	$P_{Tr500años} = 2.12^* \bar{P}$	$P_{Tr500años} = 2.24^* \bar{P}$	$P_{Tr500años} = 2.04^* \bar{P}$
0.999	1000	$P_{Tr1000años} = 2.25^* \bar{P}$	$P_{Tr1000años} = 2.41^* \bar{P}$	$P_{Tr1000años} = 2.15^* \bar{P}$

### CONCLUSIONES

Mediante los resultados de aplicación del análisis regional, se establecen 09 modelos regionales basados en el índice de avenida para periodos de retorno de 2, 5, 10, 20, 50, 100, 200, 500 y 1000 años, cuyos resultados serán de utilidad para estimar precipitaciones máximas diarias en puntos o zonas donde no se dispone de información dentro de la región hidrográfica del Titicaca. Así mismo una recomendación práctica para entidades públicas y privadas que estén involucradas con el diseño de proyectos hidrológicos, la precipitación máxima para un determinado periodo de retorno en cualquier punto de interés se estima, una vez obtenido el valor promedio de la precipitación máxima (índice de avenida), multiplicando éste por el correspondiente cuantil regional adimensional, cuyo producto vendría a ser la intensidad de precipitación máxima en 24 horas, que sería el insumo de entrada a modelos hidrológicos lluvia escorrentía, obtención de curvas intensidad duración frecuencia y estimación de la avenida máxima de diseño para el dimensionamiento de estructuras hidráulicas, recordando que los valores obtenidos son sólo aproximaciones de un determinado patrón de razonamiento que refleja algunas de nuestras convicciones y preferencias particulares ya que las verdaderas distribuciones que describen el comportamiento probabilístico del fenómeno en cuestión no son conocidos.

### REFERENCIAS BIBLIOGRAFICAS

Álvarez, M., Puertas, J., Soto, B. y Díaz, F. (1999). Análisis regional de las precipitaciones máximas en Galicia mediante el método del índice de avenida. *Ingeniería del Agua*, 6(4), 283-290.

Badreldin, G.H. y Feng, P. (2012). Regional Rainfall Frequency Analysis for the Luanhe Basin by Using L-moments and Cluster Techniques, 1, 126-135.

Cunnane, C. (1988). Methods and merits of regional flood frequency analysis. *Journal of Hydrology*, 100, 269-290.

Cunnane, C. (1989). *Statistical distributions for flood frequency analysis*. Geneva, Switzerland: Secretariat of the World Meteorological Organization.

Dalrymple, T. (1960). Flood frequency analysis. Water Supply Paper 1543-A, USGS. Washington. United States.

Domroes, M., Kaviani, M. y Schaefer, D. (1998) An Analysis of Regional and Intra-annual Precipitation Variability over Iran using Multivariate Statistical Methods. *Theoretical and Applied Meteorology*, 61, 151-159.

Escalante, C.A. y Reyes, L. (2002). Técnicas Estadísticas en Hidrología. Facultad de Ingeniería, Universidad Nacional Autónoma de México, 298p.

Gottschalk, L. (1985). Hydrological regionalization of Sweden. *Hydrology Science Journal*. 30, 65-83.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C. y Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15(5), 1049-1054.

GREHYS (1996). Inter-comparison of regional flood frequency procedures for Canadian rivers. *Journal of Hydrology*, 186, 85-103.

Gupta, H. V., Sorooshian, S. y Yapo, P. O. (1999). Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*, 4(2), 135-143.

Hosking, J.R.M. y Wallis, J.R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research*, 29(2), 271-281.

Hosking, J.R.M. y Wallis, J.R. (1997). Regional Frequency Analysis: An Approach based on L-moments. Cambridge University Press, Cambridge, UK.

Hosking, J.R.M. (2009). L-moments. R package, version 1.5. URL: <http://CRAN.R-project.org/package=lmom>.

Jackson, I.J. y Weinand, H. (1995). Classification of tropical rainfall stations: a comparison of clustering techniques. *Journal Climatology*, 15(9), 985-994.

Madsen, H., Pearson, C.P. y Rosbjerg, D. (1997). Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events 2. Regional modeling. *Water Resources Research*, 33(4), 759-769.

Masoodian, S.A. (1998) An analysis of Tempo-Spatial variation of precipitation in Iran. Ph.D.thesis in climatology, University of Isfahan, Iran.

McCuen, R.H., Knight, Z. y Cutter, A.G (2006). Evaluation of the Nash Sutcliffe Efficiency Index. *Journal Hydrology Engineering*, 11, 597-602.

- Mkhandi, S. (1995). Choosing a distribution for flood frequency analysis. 7th South African National Hydrology Symposium.
- Mkhandi, S. y Kachroo, S. (1997). Regional flood frequency analysis for Southern Africa. Southern African FRIEND. Technical Documents in Hydrology.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., y Veith, T. L. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulation. *Transactions of the ASABE*, 50(3), 885-900.
- Murray, C.P., Wang, Q.J., Vogel, R.M. y McMahon, T.A. (2001). The utility of L-moment ratio diagrams for selecting a regional probability distribution. *Hydrological Sciences Journal des Sciences Hydrologiques*, 46(1), 147-155.
- Nash, J.E. y Sutcliffe, J.V. (1970). River flow forecasting through conceptual models, part I: A discussion of principles. *Journal of Hydrology*, 10 (3), 282–290.
- Nathan, R.J. y McMahon, T.A. (1990) Identification of homogeneous regions for the purpose of regionalization. *Journal of Hydrology*, 121, 217-238.
- Oliveira, C. M. y Naghettini, M. (2008). SEAF – A prototype of an expert system for at-site frequency analysis of hydrological annual máxima. *Avances en Recursos Hidráulicos*, 18, 37-56.
- Organización Meteorológica Mundial (2011). Guía de prácticas hidrológicas, gestión de recursos hídricos y aplicación de prácticas hidrológicas. Volumen II.
- Parida, B.P. y Moalaffi, D.B. (2008). Regional rainfall frequency analysis for Botswana using L-Moments and radial basis function network. *Physics and Chemistry of the Earth*, 33, 614–620.
- Pearson, C.P., Mckerchar, A.I. y Woods, R.A. (1991). Regional flood frequency analysis of western Australian data using L-moments. *International Hydrology and Water Resources Symposium*, 631-632.
- Rodríguez, Y.L. y Marrero de Leon, N.V. (2011). Análisis regional de series de lluvias máximas. *Ingeniería Hidráulica y Ambiental*, XXXII(2), 34-45.
- Rao, A.R. y Srinivas, V.V. (2006). Regionalization of watersheds by hybrid cluster analysis. *Journal of Hydrology*, 318, 37–56.
- Vogel, R.M. y Fennessey, N.M. (1993). L-moment diagrams should replace product moment diagrams. *Water Resources*, 29(6), 1745–1752.
- Vogel, R.M., Thomas, W.O. y McMahon, T.A. (1993a). Flood-flow frequency model selection in southwestern United States. *Journal of Water Resources Planning and Management*. 119(3), 353-366.
- Vogel, R.M., McMahon, T.A. y Chiew, F.H.S. (1993b). Flood-flow frequency model selection in Australia. *Journal of Hydrology*, 146, 421-449.
- Zhang, J. y Hall, M.J. (2004). Regional flood frequency analysis for the Gan-Ming River basin in China. *Journal of Hydrology*, 296, 98–117.