

REGIONALIZATION OF THE IDF CURVE FOR THE PROVINCES OF LOS RÍOS, GUAYAS AND EL ORO

Velasco Ramos Marcos¹, Garaicoa Velásquez Joshua², Rodríguez Santos Josué, MSc³

velascojesus.ramos@gmail.com¹ <https://orcid.org/0000-0003-2812-5086>,

jgaraicoa25@gmail.com² <https://orcid.org/0000-0001-6200-9156>,

josue.rodriguez@ug.edu.ec³

Universidad de Guayaquil - Ecuador.

Abstract-

The small amount of information in the country regarding precipitation, information which usually presents missing data, problem that generates the present research that aims to present the analysis of the parameters k , m and n belonging to the Bernard equation for the development of Intensity, Duration and Frequency (IDF) curves, using the regionalization method with historical data from different meteorological stations located in Ecuador, specifically in the provinces of Los Rios, Guayas, El Oro and neighboring areas.

Regionalization allows us to obtain data from projections based on information obtained through interpolation methods, the best known being ordinary kriging (KO) and inverse distance weighting (IDW), with which a comparison will be made to choose the one that has the highest degree of affinity with the provinces mentioned above and their meteorological data.

The regionalization allows the development of the present model, which is reliable in areas where the available data is poor and the meteorological stations do not cover efficiently.

The results provided by the KO method had better adjustment in the study provinces, presenting error percentages for Los Rios, Guayas 1, Guayas 2, Guayas 3 and El Oro of 0,92%, 7,64%, 14,21%, 19,83% and 4,47% for parameter k and 13,72%, 19,14%, 1,74%, 8,40% and 6,76% for parameter m respectively, being parameter n constant, it was omitted in the analysis.

From field data regionalization, parameters that have high error percentages, needs especial considerations to validate the data, this due to the unreliable source of the information.

KEYWORDS: Regionalization IDF curves, Los Rios, Guayas, El Oro, Spatial analysis.

Digital Object Identifier: (only for full papers, inserted by LACCEI).

ISSN, ISBN: (to be inserted by LACCEI).

DO NOT REMOVE

REGIONALIZACIÓN DE LA CURVA IDF PARA LAS PROVINCIAS DE LOS RÍOS, GUAYAS Y EL ORO.

Velasco Ramos Marcos¹, Garaicoa Velásquez Joshua², Rodríguez Santos Josué, MSc³
velascojesus.ramos@gmail.com¹ <https://orcid.org/0000-0003-2812-5086>,
jgaraicoa25@gmail.com² <https://orcid.org/0000-0001-6200-9156>,
josue.rodriguez@ug.edu.ec³
Universidad de Guayaquil - Ecuador.

Resumen-

La poca información existente en el país respecto a parámetros de intensidad de precipitación, información que por lo general presenta datos faltantes, problema que genera la siguiente investigación, misma que tiene como objetivo presentar el análisis de los parámetros k , m y n pertenecientes a la ecuación de Bernard para el desarrollo de curvas Intensidad, duración y frecuencia (IDF), mediante el método de la regionalización con la data histórica de 208 estaciones meteorológicas ubicadas en el Ecuador, específicamente en las provincias de Los Ríos, Guayas, El Oro y zonas colindantes.

La regionalización nos permite obtener datos a partir de proyecciones basadas en información obtenida, esto mediante métodos de interpolación, siendo los más conocidos Kriging ordinario (KO) y la ponderación de distancia inversa (IDW) con los cuales se realizará una comparación para elegir el que tenga mayor grado de afinidad con las provincias antes mencionadas y su data meteorológica.

La regionalización permite desarrollar el presente modelo, el mismo que es confiable en zonas en las que la data disponible es escasa y las estaciones meteorológicas no abarcan de manera eficiente.

Los resultados proporcionados por el método KO tuvieron mejor ajuste en las provincias de estudio, presentando porcentajes de error para Los Ríos, Guayas 1, Guayas 2, Guayas 3 y El Oro de 0,92%, 7,64%, 14,21%, 19,83% y 4,47% para el parámetro k y 13,72%, 19,14%, 1,74%, 8,40% y 6,76% para el parámetro m respectivamente, siendo el parámetro n constante se lo omitió en el análisis.

De regionalizar la data de campo, los parámetros que presenten porcentajes de error altos deben tener una consideración especial para su validación, esto debido a la fuente de información la cual es poco confiable.

Palabras claves: Regionalización curvas IDF, Los Ríos, Guayas, El Oro, Análisis espacial.

Abstract-

The small amount of information in the country regarding precipitation, information which usually presents missing data, problem that generates the present research that aims to present the analysis of the parameters k , m and n belonging to the Bernard equation for the development of Intensity, Duration and Frequency (IDF) curves, using the regionalization method with historical data from different meteorological stations located in Ecuador, specifically in the provinces of Los Rios, Guayas, El Oro and neighboring areas.

Regionalization allows us to obtain data from projections based on information obtained through interpolation methods, the best known being ordinary kriging (KO) and inverse distance weighting (IDW), with which a comparison will be made to choose the one that has the

highest degree of affinity with the provinces mentioned above and their meteorological data.

The regionalization allows the development of the present model, which is reliable in areas where the available data is poor and the meteorological stations do not cover efficiently.

The results provided by the KO method had better adjustment in the study provinces, presenting error percentages for Los Rios, Guayas 1, Guayas 2, Guayas 3 and El Oro of 0,92%, 7,64%, 14,21%, 19,83% and 4,47% for parameter k and 13,72%, 19,14%, 1,74%, 8,40% and 6,76% for parameter m respectively, being parameter n constant, it was omitted in the analysis.

From field data regionalization, parameters that have high error percentages, needs especial considerations to validate the data, this due to the unreliable source of the information.

KEYWORDS: Regionalization IDF curves, Los Rios, Guayas, El Oro, Spatial analysis.

I. INTRODUCCIÓN

Según la hidrología, la precipitación es la principal fuente de agua en la superficie terrestre, y los estudios relacionados con el uso y la gestión del agua suelen comenzar con sus mediciones [1], los análisis relacionados con las mismas deben pasar por un proceso minucioso, necesario para darle fiabilidad a los resultados, debido a todas las incongruencias que afectan a los datos en su recolección.

La recolección de data es uno de los factores más importantes en el desarrollo de un análisis hidrológico, debido al impacto que tiene sobre la regionalización al ser la base para la interpolación espacial el cual es un método que permite extrapolar los datos puntuales relativos a cada estación meteorológica a un espacio de información deficiente más extenso [2], en el Ecuador el ente regulador de la mayor red meteorológica es el Instituto Nacional de Meteorología e Hidrología (INAMHI) el cual se encarga de la recepción, recolección y análisis de la data histórica meteorológica.

Recolección de data histórica

Comprende la etapa inicial de un análisis hidrológico, se basa en la toma de datos mediante estaciones meteorológicas para su posterior recopilación y tratamiento.

La situación actual de la data meteorológica en el país presenta inconvenientes en su recolección, como:

- Años sin registro en la mayoría de las estaciones.
- Data privatizada de difícil acceso.
- Retrasos en el procesamiento de data obtenida.
- Data actualizada no disponible.

Por lo antes mencionado se tiene que recurrir a métodos estadísticos que permitan complementar la data faltante y tener un mayor registro para el análisis.

Complementación de datos faltantes

En la proyección de datos hidrológicos, cuando las estaciones presentan una falta de información de carácter diaria, mensual o anual, existe una diversidad de metodologías a seguir, desde métodos tradicionales hasta avances tecnológicos como el uso de programas estadísticos, satélites y modelamientos. Los tradicionales son los más usados por su viabilidad, estos se basan en modelos matemáticos donde se establece una relación entre estaciones objetivo (data incompleta) y estaciones base (data completa) [3].

La presente investigación para la complementación de datos se propuso dos métodos el de regresión lineal y el U.S National Weather Service.

Regresión lineal. - Consiste en la suposición de que dos variables se relacionan de forma lineal, se recomienda que esta sea acompañada de una prueba de homogeneidad (t de student) para determinar diferencias significativas entre la estación base y la estación objetivo, una vez completado los datos, expresada con la siguiente formula [4]:

$$y = a + bx \quad (1)$$

Dónde:

x = variable independiente (conocida)

y = variable dependiente (predecible)

a = intercepto, punto donde la línea de regresión cruza el eje y , es decir valor de y cuando $x=0$

b = pendiente de la línea o coeficiente de regresión, es decir es la cantidad de cambio de y asociada a un cambio unitario de x

U.S National Weather Service. - Esta metodología empírica, ha sido verificada matemáticamente, consiste en ponderar los valores observados. Siendo igual al recíproco del cuadrado de la distancia de las estaciones vecinas. Esto se puede aplicar para datos faltante diarios, mensuales o anuales, puede definirse con la siguiente ecuación [5]:

$$P_x = \frac{\sum \left(P_i * \frac{1}{(D_i)^2} \right)}{\sum \frac{1}{(D_i)^2}} \quad (2)$$

Dónde:

P_x = valor de precipitación

P_i = dato observado para la fecha faltante, de la estación auxiliar circundante (mínimo 2) en mm

D_i = distancia entre estación circundante y la estación incompleta en km

Prueba de datos atípicos

Cuando se realiza el proceso de complementación de datos, estos suelen presentar anomalías en la data de precipitación calculada. Los datos dispersos corresponden a los valores que no siguen la misma tendencia apartándose de la media de los valores nativos. La selección de este procedimiento de detección de valores atípicos se basó en el ensayo de varios procedimientos en un logaritmo simulado log-Pearson Tipo III, en los datos de inundaciones observadas y la comparación de los resultados [6].

La metodología del Water Resources Council aconseja utilizar los coeficientes de sesgo en la tabla I.

Tabla I. Coeficientes de Sesgo para distintos casos: Ven T. Chow – Hidrología Aplicada.
Fuente: Propia

COEFICIENTE DE SESGO	TIPO DE PRUEBA
+0,4	Datos dudosos alto
-0,4	Datos dudosos bajo
-0,4 y +0,4	Datos dudosos alto y bajos

Las siguientes fórmulas presentadas en la tabla II se utilizan para estimar los umbrales altos y bajos basados en los coeficientes de sesgo.

Tabla II. Fórmulas para estimar los umbrales para la prueba de datos dudosos en unidades logarítmicas: Ven T. Chow – Hidrología Aplicada.
Fuente: Propia

ECUACIÓN	TIPO DE UMBRAL
$y_H = \bar{y} + k_n S_y$	Alto
$y_L = \bar{y} - k_n S_y$	Bajo

\bar{y} .- Media de la muestra
 k_n .- Valor en función de la cantidad de datos a un nivel de significancia del 10 % (ver Tabla III)
 S_y .- Desviación estándar

El parámetro K_n presentado en la tabla II se lo estima a partir del número de datos de la muestra analizada, rigiéndose por los valores mostrados en la tabla III.

Tabla III. Valores K_n para diferentes tamaños de muestra(n) en la prueba de datos dudosos con un nivel de significancia del 10%.
Fuente: Flood Flow Frequency [6]

n	K_n	n	K_n	n	K_n	n	K_n
10	2,036	43	2,710	76	2,922	118	3,073
11	2,088	44	2,719	77	2,927	119	3,075
12	2,134	45	2,727	78	2,931	120	3,078
13	2,175	46	2,736	79	2,935	121	3,081
14	2,213	47	2,744	80	2,940	122	3,083
15	2,247	48	2,753	81	2,945	123	3,086
16	2,279	49	2,760	82	2,949	124	3,089
17	2,309	50	2,768	83	2,953	125	3,092

18	2,335	51	2,775	84	2,957	126	3,095
19	2,361	52	2,783	85	2,961	127	3,097
20	2,385	53	2,790	86	2,966	128	3,100
21	2,408	54	2,798	87	2,970	129	3,102
22	2,429	55	2,804	88	2,973	130	3,104
23	2,448	56	2,811	89	2,977	131	3,107
24	2,467	57	2,818	90	2,981	132	3,109
25	2,486	58	2,824	91	2,984	133	3,112
26	2,502	59	2,831	92	2,989	134	3,114
27	2,519	60	2,837	93	2,993	135	3,116
28	2,534	61	2,842	94	2,996	136	3,119
29	2,549	62	2,849	95	3,000	137	3,122
30	2,563	63	2,854	96	3,003	138	3,124
31	2,577	64	2,860	97	3,006	139	3,126
32	2,591	65	2,866	98	3,011	140	3,129
33	2,604	66	2,871	99	3,014	141	3,131
34	2,616	67	2,877	100	3,017	142	3,133
35	2,628	68	2,883	101	3,021	143	3,135
36	2,639	69	2,888	102	3,024	144	3,135
37	2,650	70	2,893	103	3,027	145	3,140
38	2,661	71	2,897	104	3,030	146	3,142
39	2,671	72	2,903	114	3,061	147	3,144
40	2,682	73	2,908	115	3,064	148	3,146
41	2,692	74	2,912	116	3,067	149	3,148
42	2,700	75	2,917	117	3,070		

Análisis de frecuencia

El objetivo del análisis de frecuencia radica en la estimación de intensidades precipitaciones o caudales, según sea necesario, estos expresados para distintos periodos de retorno.

Para que el estudio de frecuencia se ajuste a un modelo matemático de distribución, debe cumplir los siguientes supuestos [7]:

- La muestra estudiada presenta eventos aleatorios.
- Los procesos naturales son estables a lo largo del tiempo.
- La muestra nos proporciona los parámetros para la población.

Las distribuciones usadas en esta investigación son: Gamma de dos parámetros, Log Pearson III y Gumbel tipo I, aplicable a las lluvias máximas horarias o diarias por lo general se ajustan a las distribuciones antes presentadas [5].

Distribución Gamma de dos parámetros

La distribución gamma es muy útil para describir variables hidrológicamente asimétricas sin necesidad de la transformación logarítmica ya que tiene una forma que varía sutilmente, similar a la función densidad de probabilidad típica [8]. Esta puede ser expresada con la siguiente ecuación [4]:

$$F(x) = \int_0^x \frac{x^{\gamma-1} e^{-\frac{x}{\beta}}}{\beta^{\gamma} \Gamma(\gamma)} dx, \quad x \geq 0 \quad (3)$$

Dónde:

γ = parámetro de forma.

β = parámetro de escala.

Γ = función gamma completa, definida por:

$$\Gamma(\gamma) = \int_0^{\infty} x^{\gamma-1} e^{-x} dx, \quad \text{confluir si } \gamma > 0$$

Distribución log Pearson tipo III

Utilizada para el análisis de las curvas IDF, esta se genera a partir de la transformación logarítmica de la distribución Gamma [9], viene dada por la siguiente ecuación [4]:

$$F(x) = \int_0^x \frac{(\ln x - x_0)^{\gamma-1} - \frac{\ln x - x_0}{\beta}}{x^{\beta} \Gamma(\gamma)} dx \quad (4)$$

En la cual:

x_0 = parámetro de posición

β = parámetro de escala

γ = parámetro de forma

Distribución Gumbel tipo I

Es un caso especial de la formula generalizada, está dada por la siguiente relación [4]:

$$F(x) = e^{-e^{-\frac{(x-\mu)}{\alpha}}} \quad (5)$$

Dónde:

α = es el parámetro de escala

μ = es el parámetro de posición, llamado también valor central o moda

Es imprescindible corroborar el mejor ajuste de acuerdo a la naturaleza de la data y los datos obtenidos mediante las distribuciones teóricas. Por lo tanto, se debe aplicar pruebas que den su validación.

Pruebas de bondad de ajuste

El objetivo de estas pruebas es medir la incertidumbre que resulta de realizar una hipótesis estadística empírica sobre una población. Se trata de determinar si la frecuencia de la serie estudiada se ajusta a una función de probabilidad teórica específica que se eligió previamente basándose en la data obtenida [4].

Smirnov-Kolmogorov

Consisten en comparar las diferencias existentes, entre la probabilidad empírica de los datos de la muestra y la probabilidad teórica, tomando el valor máximo del valor absoluto entre ellas, expresada con la siguiente formula [4]:

$$\Delta = \max |F(x) - P(x)| \quad (6)$$

Dónde:

Δ = estadístico de Smirnov-Kolmogorov

$F(x)$ = probabilidad de la distribución teórica

$P(x)$ = probabilidad empírica

La tabla IV presenta los valores críticos Δ_0 con distintos niveles de significación, en base a las siguientes criterios establecen el ajuste del metodo:

$\Delta < \Delta_0 \rightarrow$ buen ajuste al nivel de significacion elegido

$\Delta > \Delta_0 \rightarrow$ mal ajuste al nivel de significación elegido, siendo necesaria elegir otra distribución

Tabla IV. Valores críticos de del valor estadístico Δ_0 del estadístico Δ , para valores de N y distintos niveles de significancia.
Fuente: Máximo Villón Béjar [4]

Tamaño de muestra	Nivel de significación				
	0,2	0,15	0,1	0,05	0,01
1	0,900	0,925	0,950	0,975	0,995
2	0,684	0,726	0,776	0,842	0,929
3	0,565	0,597	0,642	0,708	0,828
4	0,494	0,525	0,564	0,624	0,733
5	0,446	0,474	0,510	0,565	0,669
6	0,410	0,436	0,470	0,521	0,618
7	0,381	0,405	0,438	0,486	0,577
8	0,358	0,381	0,411	0,457	0,543
9	0,339	0,360	0,388	0,432	0,514
10	0,322	0,342	0,368	0,410	0,490
11	0,307	0,326	0,352	0,391	0,468
12	0,295	0,313	0,336	0,375	0,450
13	0,284	0,302	0,325	0,361	0,433
14	0,274	0,292	0,314	0,349	0,418
15	0,266	0,283	0,304	0,338	0,404
16	0,258	0,274	0,295	0,328	0,392
17	0,250	0,266	0,286	0,318	0,381
18	0,244	0,259	0,278	0,309	0,371
19	0,237	0,252	0,272	0,301	0,363
20	0,231	0,246	0,264	0,294	0,356
25	0,210	0,220	0,240	0,270	0,320
30	0,190	0,200	0,220	0,240	0,290
35	0,180	0,190	0,210	0,230	0,270
N>35	$\frac{1,07}{\sqrt{N}}$	$\frac{1,14}{\sqrt{N}}$	$\frac{1,22}{\sqrt{N}}$	$\frac{1,36}{\sqrt{N}}$	$\frac{1,63}{\sqrt{N}}$

Periodo de retorno. - Es una estimación en años en la que el valor de una tormenta de diseño puede ser igualado o superado. Al estimar un tiempo de retorno, se examina la relación entre la probabilidad prevista del suceso, la durabilidad estructural y el riesgo de fallo admisible [10].

Relación a la lluvia de 24 horas

El uso de las relaciones con la precipitación máxima en 24 horas permite conocer la magnitud de esta expresada en horas. D.F. Campos A. adoptó estos valores después de comparar 8 criterios de estimación [5], presentados en la tabla V.

Tabla V. Valores concluidos para relaciones a lluvia de duración 24 horas.
Fuente: Campos Aranda [5]

Duración en horas									
1	2	3	4	5	6	8	12	18	24
0.30	0.39	0.46	0.52	0.57	0.61	0.68	0.80	0.91	1.00

Curva Intensidad, Duración y Frecuencia (IDF)

Las curvas IDF es una metodología para estimación y proyección de datos de precipitación, existen diferentes modelos matemáticos para estas curvas que en su mayoría están en función del periodo de retorno y distribuciones estadísticas [10].

Modelo matemático. - Charles W. Sherman y Merrill Bernard fueron los primeros en parametrizar esta triple relación a principios del siglo XX. El modelo de Bernard correlaciona

simultáneamente la intensidad, la duración y el periodo de retorno de las precipitaciones mediante una familia de curvas; la ecuación podría expresarse de la siguiente manera [5] :

$$I = \frac{K T^m}{D^n} \quad (7)$$

Dónde:

I= intensidad en mm/hr
D= duración de la lluvia en min
T= periodo de retorno en años
K, m, n.- parámetros de ajuste

Prueba de coeficiente de correlación múltiple

Esta prueba estadística busca conocer la correlación entre tres o más variables, en la que dos o más de ellas se asocian y la tercera queda inalterada. Relaciones más cercanas a 1 señalan una buena correlación, se calcula con la siguiente ecuación:

$$R = \left(1 - \frac{S^2_{Y|X_1+X_2}}{S^2_Y} \right)^{\frac{1}{2}} \quad (8)$$

Dónde:

$$S^2_{Y|X_1+X_2} = \frac{1}{n-3} \sum (y_i - \hat{y}_i)^2 \quad (9)$$

$$S^2_Y = \frac{1}{n-1} \sum (y_i - \bar{y}_i)^2 \quad (10)$$

\hat{y}_i = valor estimado de y para X_{1i} y X_{2i}

Regionalización

Es una metodología que utiliza diferentes tipos de interpolación espacial utilizadas para proyectar información en lugares que haya escasos o inexistencia de datos, en base a información del sitio y zonas colindantes.

Las interpolaciones más comunes para data observada en la hidrología son IDW y KO [11].

Método de Ponderación Inverso a la Distancia

El método IDW consiste en el cálculo de parámetros en zonas sin data conocida, mediante una combinación lineal entre los valores en puntos de la muestra ponderada y el inverso de su distancia, debido a esto los puntos más cercanos tendrán mayor influencia sobre el punto de estudio, este método es determinístico, exacto, local y gradual. El método puede ser definido con la siguiente formula [11]:

$$\hat{z} = \frac{\sum_{i=1}^n \frac{1}{(d_i)^p} z(x_i)}{\sum_{i=1}^n \frac{1}{(d_i)^p}} \quad (11)$$

Dónde:

\hat{z} = valor a ser estimado
 $z(x_i)$ = diferentes muestras de entrada usadas para la interpolación
 d_i = distancia desde el punto muestreado al punto a interpolar
 p = valor ponderado de la distancia

Método de Interpolación Kriging Ordinario

Este método utiliza la posición geográfica, la cual depende de la auto correlación espacial de la variable para realizar las interpolaciones, Kriging se basa en funciones matemáticas para incorporar más peso en las posiciones más cercanas a los puntos de muestreo y menores pesos en posiciones más alejadas, y así crear nuevos puntos interpolados basados en estas combinaciones lineales de datos. Puede definirse con la siguiente ecuación [11]:

$$\hat{z}_{(x_0)} - \mu = \sum_{i=1}^n \lambda_i [z(x_i) - \mu(x_0)] \quad (12)$$

Dónde:

\hat{z} = valor a ser estimado

μ = media estacionaria conocida

λ_i = peso de kriging

n = numero de puntos de muestra

$\mu(x_0)$ = media de las muestras

II. METODOLOGÍA

Análisis de data de precipitación

Se considera el siguiente procedimiento:

- Para la presente investigación se utilizó la precipitación máxima de 24 horas anual disponible en los anuarios hidrológicos del INAMHI, ver anexo 1 se presenta el formato de la información ponderada por esta institución.
- Llevar a cabo la complementación de datos faltantes para la estación que lo amerita, se analizaran las condiciones de aplicación para cada método, siendo estos la regresión lineal ecuación (1) y U.S National Weather Service ecuación (2).
- Posterior a la complementación de los datos y corroborando tener un mínimo 10 años, se aplica la prueba de datos atípicos donde la tabla I nos presenta los valores de sesgo y su correspondiente categoría, se usará la tabla II donde se encuentran las ecuaciones para determinar los rango máximos y mínimos de precipitación en conjunto con la tabla III.
- Luego de filtrar la data, se empleará el software libre Hydrognomon para el análisis de distribuciones teóricas, se estudiaron las distribuciones Gumbel tipo I ecuación (5), Gamma incompleta ecuación (3) y Log Pearson III ecuación (4).
- Con el Hydrognomon se aplicará la prueba de bondad de ajuste de Smirnov-Kolmogorov ecuación (6), los valores críticos a comparar para los distintos niveles de significación se encuentran en la tabla IV, se observan los porcentajes de ajuste con respecto a las distribuciones teóricas mencionadas.
- La prueba de bondad de ajuste como resultado indica la distribución más predominante presente en las estaciones, con la cual se estimó la precipitación para los distintos

tiempos de retorno elegidos, además se aplicó el factor de ajuste de 1,13 para tener una estimación más acercada a la realidad, recomendada por la Organización Meteorológica Mundial (OMM).

- Con la data tratada y aplicado los distintos tiempos de retorno, se estimaron los parámetros del modelo matemático de Bernard para elaborar las curvas IDF en base al proceso mencionado por Campos Aranda con los coeficientes para lluvias máximas presentes en la tabla V [5].
- Para corroborar el tipo de relación que tienen las variables entre ellas se les aplico la prueba del coeficiente de correlación ecuación (8) (9) (10).

Regionalización mediante el software libre QGIS

Luego de comprobar la afinidad de las variables se procede a realizar el siguiente procedimiento:

- Una vez creado el proyecto se realiza su configuración, siendo uno de los primeros pasos a considerar el sistema de referencia de coordenadas, en nuestro caso el Universal Transversal de Mercator (UTM), se recalca que se debe tener cuidado al elegir la zona UTM para el Ecuador ya que presenta 4 zonas de coordenadas diferentes, siendo estas 17S, 18S, 17N y 18N, ver anexo 2 correspondiente a la cuadrícula por zonas.
- Se ingresa el archivo shp de la división por provincia, luego se selecciona y se exporta como un solo archivo para tener delimitada el área de estudio en el análisis con QGIS.
- Se convierten las coordenadas de las estaciones a formato UTM para su ingreso a QGIS, se retiraron estaciones de estudio por provincia para poder realizar un análisis comparativo.
- Con la información ingresada se realizan las interpolaciones IDW ecuación (11) y KO ecuación (12) para cada parámetro correspondiente al modelo matemático.
- Se deben realizar varias pruebas con las estaciones de estudio hasta encontrar la mejor distribución con la que se obtendrán los porcentajes más bajos de error en la prueba de validación (error relativo).

Visualización de datos en página web

Siendo esta una continuación del trabajo Regionalización de las Curvas de Intensidad-Duración-Frecuencia (IDF) en las Provincias de Santa Elena y Manabí [12], el siguiente link <https://arcg.is/9rCCy0> presenta la información referente a la regionalización de los parámetros k, m y n, para las provincias analizadas hasta la fecha.

III. RESULTADOS

Recolección de información

- En la recopilación inicial se obtuvieron 203 estaciones dentro de las zonas de estudio y colindantes.
- En base a criterios estadístico se concluyó que la data historia mínima debe ser 10 años, además en este proceso se seleccionaron estaciones de apoyo que no cumplen con el mínimo de 10 pero serán de utilidad para la posterior complementación de datos.

Tabla VI. Cantidad de estaciones por criterio de años.
Fuente: Propia

Criterio	#Estaciones	Observación
≥ 10 años	75	Aceptada
< 10 años	9	Apoyo

Complementación de datos

- Durante la complementación de datos se analizó la posibilidad de utilizar las estaciones de apoyo cuyas condiciones iniciales permitieran completar el mínimo de 10 años mediante esta metodología, obteniendo un total de 76 estaciones.

Prueba de datos atípicos

- Luego de esta prueba se descartaron 3 estaciones por tener menos de 10 años al final del análisis, quedando con 73 estaciones restantes.

Análisis de distribuciones teóricas y prueba de bondad

- Mediante hydrognomon se analizaron las distintas distribuciones teóricas obteniendo los siguientes resultados en base a la prueba de bondad.

Tabla VII. Cuento de estaciones basado en distribuciones con mejor ajuste.
Fuente: Propia

Distribuciones teóricas	Cantidad de estaciones
Log Pearson III	32
Gumbel tipo I	9
Gamma de dos parámetros	32

- Las distribuciones teóricas más predominantes son Gamma y Log Pearson III, con esto presente se optó a la búsqueda de otro parámetro de análisis, el Dr. Víctor Manuel Ponce [13] partiendo de estudios del Flood Studies Report menciona que los métodos de tres parámetros como Log Pearson III suelen ser generalmente más confiables, sin embargo cuando la extensión de la data histórica es corta los métodos de dos parámetros como Gamma han demostrados ser más confiables, por lo tanto en base a lo antes mencionado se decidió utilizar la distribución gamma incompleta para el desarrollo de las curvas IDF.

- Para mantener un margen de aceptación teniendo como referencia a la prueba de bondad se tomaron las estaciones mayores al 70% de ajuste, quedando con un total de 66 estaciones.

Regionalización

Se presentan las Distribuciones de las estaciones al inicio y final del análisis con las interpolaciones IDW y KO, además de los modelos resultantes.

Distribución inicial

En la figura 1 se presentan las estaciones resultantes del filtrado de las pruebas estadísticas y las estaciones prueba utilizadas en el análisis.

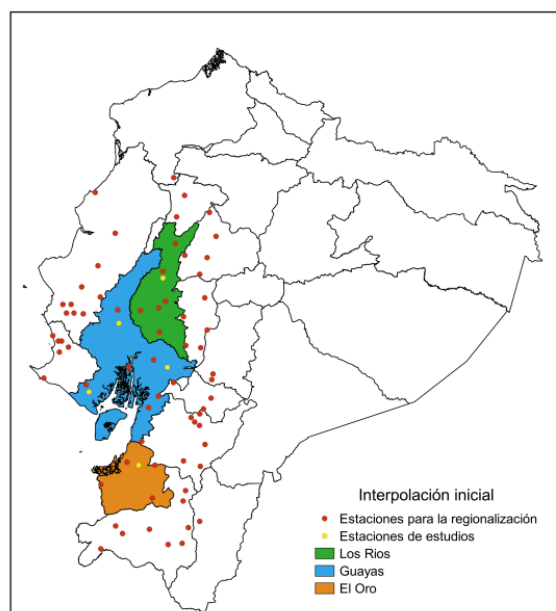


Fig. 1. Distribución inicial.
Fuente: Propia

Distribución final

En la figura 2 se muestra la mejor distribución luego de realizar aproximadamente 9 distribuciones, donde los mejores ajustes se obtuvieron excluyendo a 3 estaciones del análisis, quedando un total de 58 estaciones.

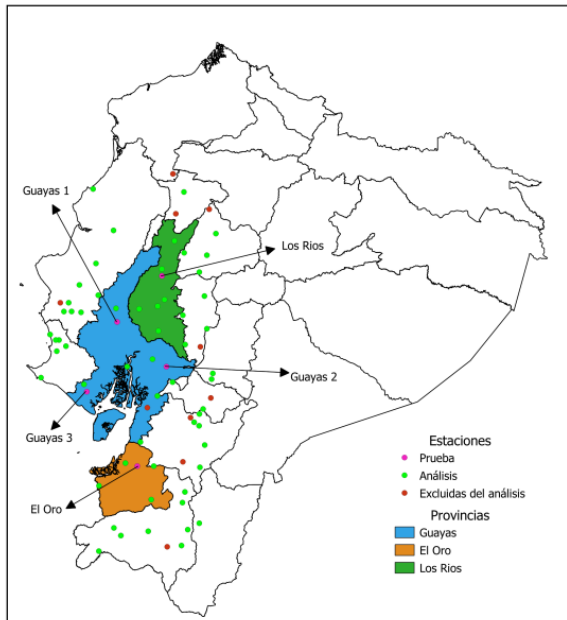


Fig. 2. Distribución final con las estaciones de estudio.
Fuente: Propia

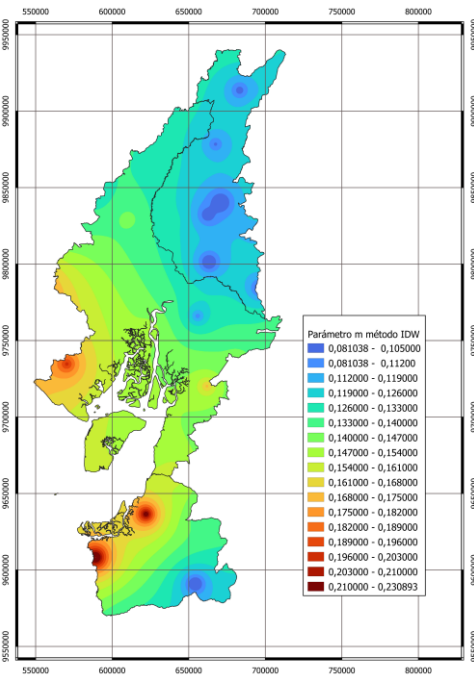


Fig. 4. Regionalización del parámetro m por el método IDW.
Fuente: Propia

Las figuras 3, 4, 5, 6 exponen los mapas por intervalos de color por cada parámetro y método realizado.

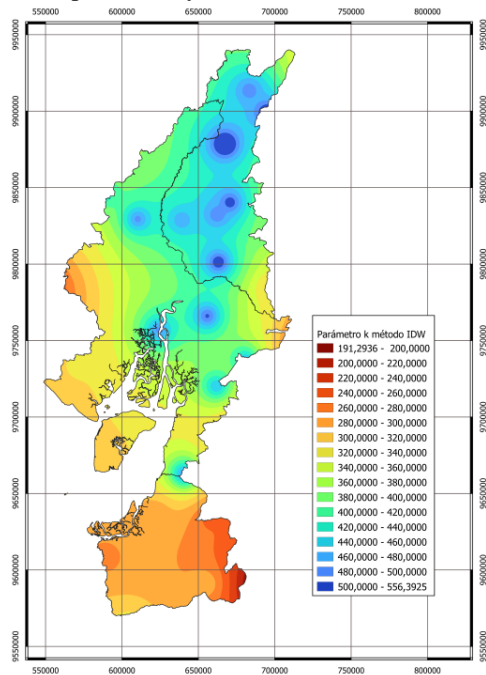


Fig. 3. Regionalización del parámetro k por el método IDW.
Fuente: Propia

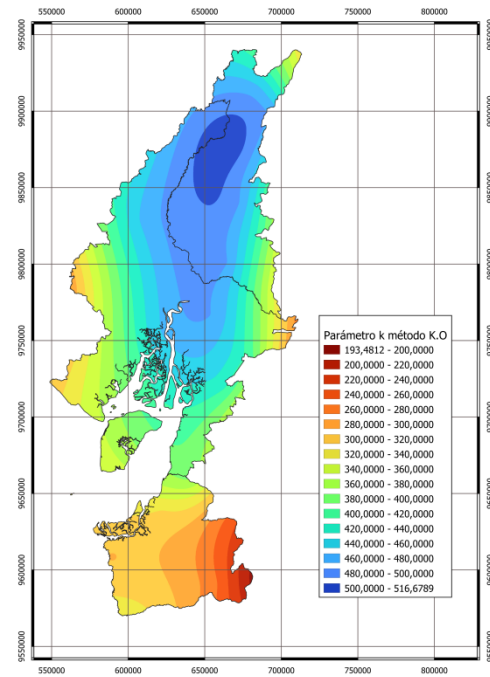


Fig. 5. Regionalización del parámetro k por el método KO.
Fuente: Propia

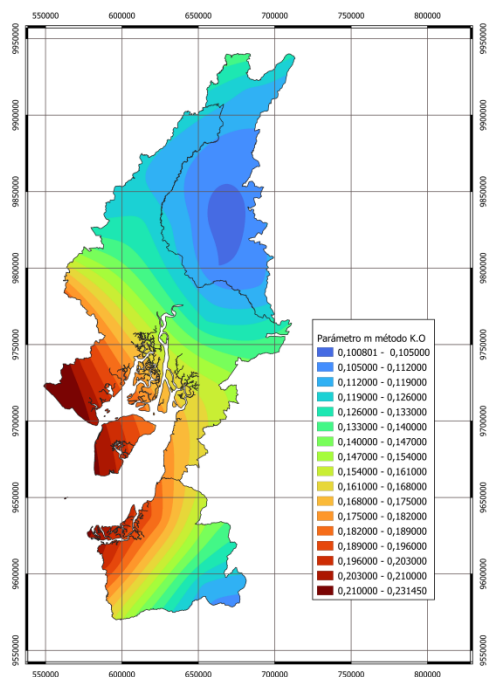


Fig. 6. Regionalización del parámetro m por el método KO.
Fuente: Propia

- La tabla VIII expone el error en porcentaje tras la prueba de validación para las estaciones prueba ver figura 2.

Tabla VIII. Errores porcentuales para los coeficientes k y m considerando los métodos IDW y KO.
Fuente: Propia

Provincia	k (IDW)	k (K.O)	m (IDW)	m (K.O)
Los Ríos	4,67%	0,92%	6,29%	13,72%
Guayas 1	6,15%	7,64%	25,96%	19,14%
Guayas 2	23,16%	14,21%	4,87%	1,74%
Guayas 3	11,84%	19,83%	23,23%	8,41%
El Oro	1,15%	4,47%	0,27%	6,77%

IV DISCUSIÓN

- La poca accesibilidad a la información publicada, misma que comprende el periodo 1990-2013 y el poco cuidado de la información hidrológica presente en los anuarios con muchas series de años y datos faltantes, datos incompletos que agravan los análisis hidrológicos, teniendo que recurrir a resultados poco confiables, además de análisis rigurosos para la poca serie de data disponible.
- En el análisis inicial de la recolección de datos se consideró 203 estaciones, de las cuales el 63.05% (128) no cumplen con un mínimo de 10 años de registro, tener más información respecto a estas estaciones mejoraría drásticamente los resultados de la regionalización.

- Para registros meteorológicos con gran cantidad de datos faltantes, la metodología establecida por U.S National Weather Service ha demostrado ser más eficaz, para la complementación de la data hidrológica, esto debido a la flexibilidad del método, la información solicitada para cada dato individual suele ser de carácter puntual.
- Se realizó interpolaciones, con las distribuciones de mejor ajuste en base a la prueba de bondad, para cada estación meteorológica. Como resultado se presentaron porcentajes de error de mayor magnitud, esto debido a que al estimar las precipitaciones para cada periodo de retorno se observó un gran desfase al usar varios métodos, sobre todo con el método Log Pearson III. Se concluyó que es recomendable aplicar una sola distribución para todas las estaciones con el método Gamma Incompleta.
- Para la generación de las curvas IDF se determinaron estaciones de prueba (color anaranjado), a partir de las cuales se generaron las estaciones de análisis (color verde).
- Del paso anterior se determinó: para la provincia de Los Ríos la estación de análisis fue M470-INAMHI; la provincia del Guayas se la caracterizo en tres zonas, Guayas 1 correspondiente a la zona centro-norte, con la estación de análisis M476-INAMHI, Guayas 3 colindante con la zona costera y su estación de análisis fue M777-INAMHI, mientras que para Guayas 2 correspondiente a la zona este de la provincia con la estación de análisis M218-INAMHI. Finalmente, para la provincia de El Oro se determinó la estación de análisis M040-INAMHI.
- Comparando los parámetros k y m, obtenidos de manera tradicional contra los regionalizados aplicados a la ecuación (7). Ecuación que hace referencia al modelo matemático establecido por Merrill Bernard, se observa que a mayor tiempo de retorno mayor la variación en la intensidad, debido a la poca data histórica utilizada en el análisis.
- Como se muestra en la figura 2 y tabla VIII se concluye que el error en la regionalización es inversamente proporcional al número de estaciones de prueba consideradas, mismas que generaron las estaciones de análisis y área de estudio, presentando la provincia del Guayas con errores mayores debido a la poca cantidad de estaciones y gran superficie en comparación con las provincias de Los Ríos y El Oro.
- La figura 3 y 5 presentan gráficamente los resultados por intervalos del método IDW y KO para al parámetro k, en ambas se observan que los valores más

altos se encuentran en la zona norte del área de estudio, ocupada en la provincia de Los Ríos mayoritariamente sin embargo en el método KO se ve una distribución más homogénea contrario al IDW en el que denota picos de información en áreas reducidas.

- La figura 4 y 6 presentan gráficamente los resultados por intervalos del método IDW y KO para el parámetro m, en la zona norte correspondiente mayormente a la provincia de Los Ríos, se establecen los valores menores para ambos métodos, mientras que en las zonas de costa reflejan los valores más altos de la serie, cabe mencionar que si hablamos de datos homogéneamente distribuidos se destaca el método KO debido que el método IDW presenta concentraciones de valores provocando puntos destacables por todo el área.
- Como se muestra en la tabla VIII para el parámetro k el menor error lo obtiene el KO en la provincia de Los Ríos siendo este de 0,92%, IDW presenta el mayor error en Guayas 2 con un 23,16%. En la misma tabla se observa que para el parámetro m, el menor error se encuentra en El Oro con un 0,27% corresponde al IDW mientras que el mayor error encontrado está en Guayas 1, teniendo un 25,96% mediante el método IDW.

REFERENCIAS

- [1] F. J. Aparicio Mijares, Fundamentos de Hidrología de Superficie, México : Limusa, 1992.
- [2] G. Fernández y D. Milla, «Validez de los datos de precipitación media mensual (1970-2000) generados por el modelo WorldClim V2.0 para Venezuela,» *Terra Nueva Etapa*, vol. XXXIV, n° 56, 2018.
- [3] R. A. Rospigliosi Cerrato, Métodos de complementación de datos hidrológicos faltantes, Moquegua, 2018.
- [4] M. Villón Béjar, Hidrología Estadística, Tecnológica de Costa Rica, 2006.
- [5] D. F. Campos Aranda, Procesos del Ciclo Hidrológico, San Luis Potosí: Universitaria Potosina, 1998.
- [6] Water Resources Council (Us). Hydrology Committee, «Flood Flow Frequency,» *Guidelines For Determining*, vol. 17B, 1982.
- [7] N. M. Cardenas Perez, Análisis de frecuencia de eventos extremos hidrológicos provenientes de mas de una población., Bogotá, 2006.
- [8] E. V. Lobo Reyes , Artist, *Estudio Hidrológico e hidraulico para la Identificación de la Vulnerabilidad a Desbordes de la Cuenca Baja del Rio Coata-Puno*. [Art]. Universidad Nacional Agraria de Molina, 2017.
- [9] D. Koutsoyiannis, D. Kozonis y A. Manetas, «A mathematical framework for studying rainfall intensity-duration-frequency relationships,» *Journal of Hydrology*, vol. 206, n° 1-2, pp. 119-135, 1998.
- [10] Miniterio de Transportes y Comunicaciones, Manual de hidrología, Hidráulica y drenaje, 2008.
- [11] M. F. Rocha Salamanca y D. M. Gómez Martínez, «Análisis Comparativo de Métodos de Interpolación Espacial Utilizando Datos de Sísmica de Reflexión Post-Stack,» de *XVI Simposio Internacional SELPER 2014*, Medellín, 2014.
- [12] N. Nuñez y P. Corapi, «Regionalización de las Curvas de Intensidad-Duración-Frecuencia (IDF) en las Provincias de Santa Elena y Manabí,» *Revista Tecnológica ESPOL*, vol. 34, 2022.
- [13] V. M. Ponce, Engineering Hydrology, Principles and Practices, 2014.
- [14] Instituto Nacional de Meteorología e Hidrología, «Anuario Meteorológico,» 2006.
- [15] A. Morton, «DMPA Distribution Mapping Software,» [En línea]. Available: <https://www.dmap.co.uk/index.htm>. [Último acceso: 15 enero 2023].

ANEXOS

Anexo 1: Presentación tipo de datos hidrológicos en los anuarios del INAMHI

M175		EL PROGRESO-GUAYAS					INAMHI								
MES	HELIOFANIA (Horas)	TEMPERATURA DEL AIRE A LA SOMBRA (°C)					HUMEDAD RELATIVA (%)			PUNTO DE ROCÍO (°C)	TENSION DE VAPOR (hPa)	PRECIPITACION(mm)		Número de días con precipitación	
		ABSOLUTAS		M E D I A S			Máxima	Mínima	Media			Suma	Máxima		
		Máxima	Mínima	da	Máxima	Mínima				Mensual	Máxima	Mínima	da	Media	Mensual
ENERO															
FEBRERO															
MARZO															
ABRIL															
MAYO					31,4	21,2	25,9		77	21,3	25,5	0,0	0,0	1 0	
JUNIO					28,7	20,0	24,0		78	19,7	23,1	1,2	1,2	7 1	
JULIO			17,0	3	26,8	19,5	23,2		80	19,3	22,5	0,0	0,0	1 0	
AGOSTO			18,0	2	29,2	19,9	24,0		79	19,9	23,4	0,5	0,5	3 1	
SEPTIEMBRE		33,0	12	17,8	7	29,8	20,1	24,3		79	20,2	23,8	0,0	0,0	1 0
OCTUBRE					29,8	20,1	24,6		89	22,6	27,7	0,0	0,0	1 0	
NOVIEMBRE					29,9	20,7	24,8		88	22,5	27,6	0,0	0,0	1 0	
DICIEMBRE					31,2	21,7	26,2		83	23,0	28,6	10,2	9,0	21 2	
VALOR ANUAL															

Fig. 1. Ponderación de datos de la estación M175 en el año 2006.

Nota: Espacios en blanco no registraron los eventos

Fuente: Instituto Nacional de Meteorología e Hidrología [14]

Anexos 2

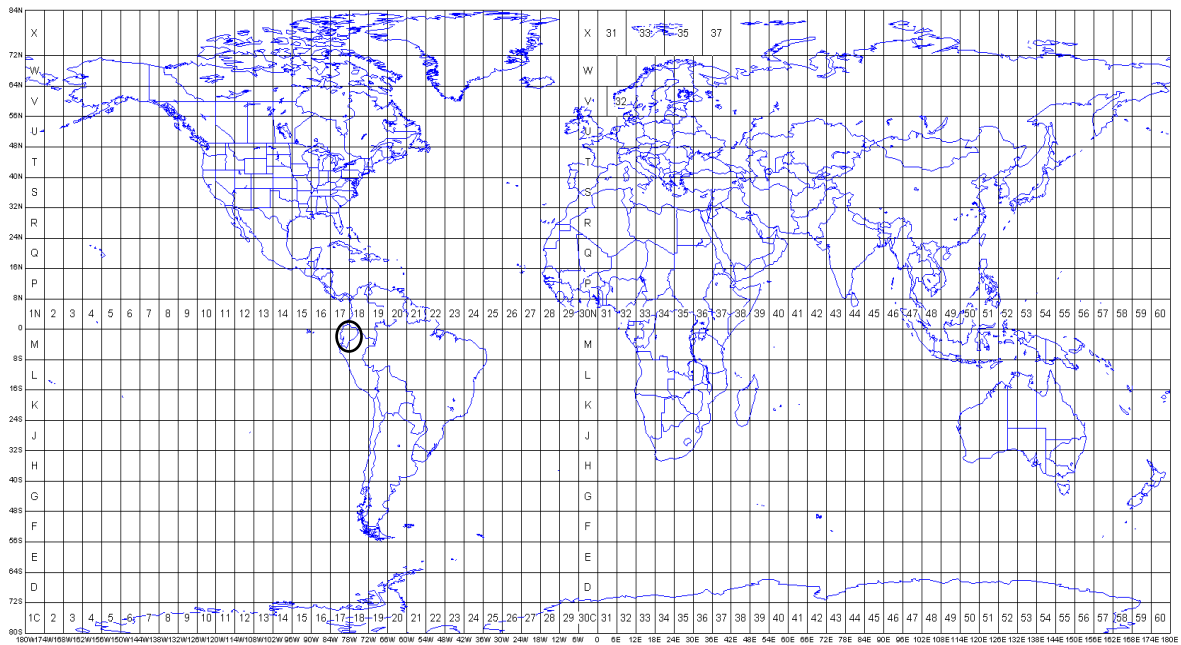


Fig. 2. Cuadrícula por zonas para el sistema Universal Transversal de Mercator (UTM) del mundo.

Fuente: DMAP [15]