

Determination of Intensity-Duration-Frequency (IDF) equation due to presence of ENSO and Niño Costero. Study case: Piura city

Marina Farías de Reyes, Magister en Ingeniería¹ y Milagros Ruiz Tesen, Bachiller en Ingeniería¹

¹University of Piura, Perú, marina.farias@udep.pe, mruiz.t92@gmail.com

Abstract— *The IDF curves consider the stationary climate, contrary to climate change. In Peru, the recurrent presence of El Niño, statistically shows changes in climate. We propose to update existing curves including recent events and appropriately cut the series to work with the most recent and reflect the current climate trend. As a case study, new IDF curves were elaborated for Piura, averaging Log-Normal 3p and Doble Gumbel models, and a general equation.*

Keywords- *IDF curves, Probability Distributions, Log-Normal 3, Double Gumbel, ENOS*

Digital Object Identifier (DOI):<http://dx.doi.org/10.18687/LACCEI2018.1.1.191>

ISBN: 978-0-9993443-1-6

ISSN: 2414-6390

Determinación de ecuación Intensidad-Duración-Frecuencia en presencia de ENSO y Niño Costero. Caso: ciudad de Piura

Marina Farías de Reyes, Magister en Ingeniería¹ y Milagros Ruiz Tesen, Bachiller en Ingeniería¹

¹University of Piura, Perú, marina.farias@udep.pe, mruiz.t92@gmail.com

Abstract— Las curvas IDF son una herramienta fundamental para modelar los caudales de diseño en lugares sin registros hidrométricos. Usualmente se determinan a partir de la modelación de tormentas históricas, asumiendo que el clima tiene la misma tendencia a lo largo del tiempo, lo que no es correcto ante las evidencias de cambio climático. En el Perú, la presencia recurrente de eventos El Niño, costero y global, genera que en los últimos años, a partir de El Niño de 1983, los registros de precipitaciones indiquen un cambio en el clima, corroborado estadísticamente. Se propone actualizar las curvas IDF existentes para incluir los últimos eventos y evaluar recortar apropiadamente las series históricas para trabajar con series más recientes que refleje la tendencia climática actual. A las intensidades definidas para la ciudad de Piura, como caso de estudio, se les ha ajustado modelos Log-Normal 3p y Doble Gumbel, cuyas predicciones se han promediado para obtener las curvas actualizadas y una ecuación general que relaciona las tres variables de diseño.

Keywords-- Curvas IDF, Distribuciones de probabilidad, Log-Normal 3, Doble Gumbel, ENSO.

I. INTRODUCCIÓN

Para diseñar las obras hidráulicas que tienen por finalidad controlar la escorrentía y proteger la infraestructura es necesario realizar un estudio hidrológico previo, que permita determinar el caudal que se requerirá evacuar. Por lo general, no se cuenta con registros de caudales para modelarlos probabilísticamente, como es el caso en el diseño de reservorios o algunos puentes, de modo que en su ausencia es necesario recurrir a modelación hidrológica de la cuenca de interés. Así, los sistemas de manejo de aguas pluviales tendrán en las curvas de Intensidad-Duración-Frecuencia (IDF) una herramienta de diseño estándar.

Cabe resaltar que, los cuantiles de precipitación extrema representados por las curvas IDF estarán sujetos a alteraciones a lo largo del tiempo [1]. Esto debido a los cambios en el ciclo hidrológico generados por el aumento de los gases de efecto invernadero. Una forma de reducir la vulnerabilidad es cuantificar los efectos potenciales del cambio climático y adaptarse a ellos [2]. En la referencia [1] se indica que en la actualidad, un enfoque común es adoptar un único período de referencia para determinar curvas IDF, y comparan curvas IDF desarrolladas tomando datos históricos de diferentes períodos de tiempo, así como escenarios de clima futuros proyectados, encontrando que hay una diferencia significativa en las intensidades de lluvia. Ante esas evidencias, en este trabajo se destaca la importancia de diseñar las obras con curvas IDF elaboradas con datos de precipitaciones actualizados y que reflejen la

tendencia actual del comportamiento climático, frente al enfoque clásico que considera el clima invariable, y busca contar con series históricas lo más largas posibles, y por tanto antiguas, para representar un universo asumido en régimen permanente. Se presenta la metodología de determinación de las curvas IDF para la ciudad de Piura, en la costa norte peruana, considerando la presencia recurrente de Fenómenos El Niño (FEN), ya sea de carácter global (ENSO) o costero, llegando a obtener una ecuación general para la intensidad en función de la duración de la lluvia y el período de retorno en la ciudad.

II. MARCO TEÓRICO

El desarrollo del presente estudio involucra varios pasos diferentes, como se muestra en la Fig. 1 y se describen con mayor detalle a lo largo de esta sección. El enfoque metodológico tuvo como objetivo la obtención de las curvas IDF para la Ciudad de Piura con información de precipitación actualizada; y posteriormente la determinación de la expresión matemática de las mismas.

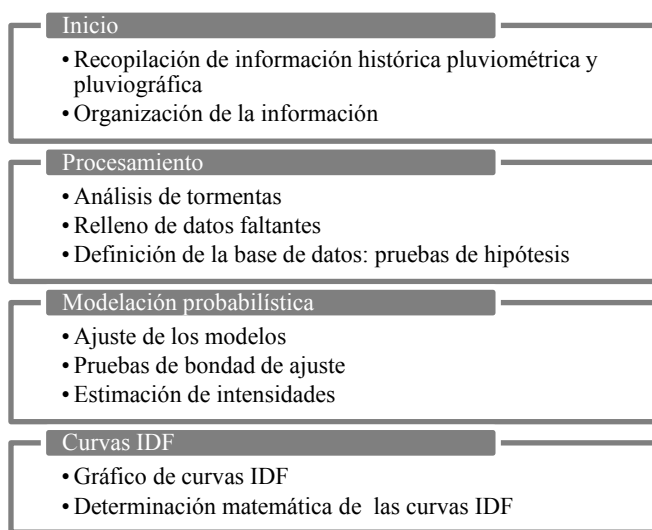


Fig. 1. Secuencia de trabajo empleada para el análisis

A. Pruebas de hipótesis

Estas pruebas de hipótesis se realizaron con la finalidad de evaluar si la media y la desviación estándar de las series A y B indican que el universo ha cambiado, es decir si hay evidencia estadística de un cambio climático en la serie.

A.1. Prueba T-Student para comparación de medias

En la referencia [3], se explica que esta prueba de hipótesis se utiliza para saber si las medias de dos muestras

son significativamente diferentes. Para tal fin, se deben distinguir dos casos: muestras grandes, por ejemplo, mayor de 50 valores, y muestras pequeñas.

Si los tamaños de ambas muestras son suficientemente grandes, se puede demostrar que el estadígrafo z definido por la expresión (1), tiene una distribución normal con promedio 0 y varianza 1; y la probabilidad de sobrepasar el valor del estadígrafo calculado, se obtiene de las tablas de la distribución normal [3].

$$z = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (1)$$

En el caso en el que el tamaño de alguna o de ambas muestras sea más reducido, lo cual es una situación común en hidrología, entonces el estadígrafo t definido por la ecuación (2), sigue una ley de distribución de Student [3]:

$$t = \frac{m_1 - m_2}{s \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (2)$$

Con una varianza que tiene en cuenta los tamaños efectivos de las muestras:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (3)$$

Si la hipótesis de que las muestras provienen de una misma distribución es verdadera, el estadígrafo t sigue una ley de distribución de Student con $n_1 + n_2 - 2$ grados de libertad. Por consiguiente, se debe obtener de la tabla de distribución de Student la probabilidad de sobrepasar el valor del estadígrafo calculado, bajo el supuesto de la hipótesis nula, es decir, que ambas muestras provienen de la misma población o universo [3].

A.2. Prueba F para comparación de dos varianzas

En esta prueba, se define el estadígrafo F como la razón entre la varianza mayor y la varianza menor. Si ambas varianzas provienen de la misma distribución, aproximadamente normal, el estadígrafo F sigue una distribución llamada ley de Fisher-Snedecor [3], con $g_1 = n_1 - 1$ y $g_2 = n_2 - 1$ grados de libertad.

$$F = s_1^2 / s_2^2 \quad (4)$$

Por consiguiente, se debe obtener de la tabla de distribución de Fisher la probabilidad de sobrepasar el valor del estadígrafo calculado, bajo el supuesto de la hipótesis nula, es decir, que ambas muestras provienen de la misma población o universo.

B. Modelación probabilística

Los datos sistematizados deben ser modelados probabilísticamente, con el fin de relacionar la magnitud y la frecuencia de los eventos. Cada modelo o función de distribución tiene sus propios parámetros. Es común el empleo de la media aritmética, relacionada con la magnitud de las variables. De igual manera, se halla la dispersión o variación de los datos de la muestra, utilizando la desviación estándar de cada muestra [4].

Existen diversos modelos probabilísticos que se pueden ajustar a las variables, a través de los cuales es posible predecir su comportamiento en función del período de retorno. Para realizar estimaciones con estos modelos se debe evaluar previamente su bondad, realizando una prueba de bondad ajuste [5].

Las distribuciones de probabilidad comúnmente utilizadas en Hidrología son entre otras:

- Distribución Normal, principalmente sus variantes Log-Normal de 2 ó 3 parámetros.
- Distribución Gamma, con sus variantes Exponencial, Gamma de 2 parámetros, Gamma de 3 parámetros o Pearson tipo III y Log-Pearson tipo III.
- Distribución Gumbel, con sus variantes Doble Gumbel y Log-Gumbel.

Para este análisis las distribuciones que se ajustaron de manera satisfactoria a las series de intensidades máximas diarias fueron la de Doble Gumbel y la Log-Normal de 3 parámetros.

B.1. Modelo probabilístico de Log-Normal de 3 parámetros

Las variables físicas de interés en hidrología (precipitación, caudal, evaporación y otras) son generalmente positivas, por lo cual es usual que presenten distribuciones de frecuencias asimétricas [3].

Algunos investigadores han propuesto aplicar una transformación logarítmica a la variable de interés y luego utilizar el modelo de distribución normal para la variable transformada. La distribución así obtenida se denomina Log-Normal de dos parámetros, que son la media y la desviación estándar de la variable transformada y , tal que [3]:

$$y = \ln(x) \quad (5)$$

Asimismo, hay una distribución Log-Normal de 3 parámetros, donde el tercer parámetro, denominado parámetro de posición, es el límite inferior x_o tal que la variable transformada es [5]:

$$y = \ln(x - x_o) \quad (6)$$

Función de densidad Log-Normal 3p

$$f(x) = \frac{1}{(x - x_o) \sigma_y \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left[\frac{\ln(x - x_o) - \mu_y}{\sigma_y} \right]^2 \right) \quad (7)$$

Para $x_o \leq x < \infty$, donde:

μ_y : Parámetro de escala

σ_y : Parámetro de forma

x_o : Parámetro de posición

Ajustar un modelo a una serie de datos consiste en estimar los parámetros de la distribución. Un método para realizarlo en una muestra de longitud N es el método de máxima verosimilitud:

$$\mu_y = \frac{1}{N} \sum_{i=1}^N \ln(x_i - x_0) \quad (8)$$

$$\sigma_y = \left[\frac{1}{N} \sum_{i=1}^N [\ln(x_i - x_0) - \mu_y]^2 \right]^{\frac{1}{2}} \quad (9)$$

El parámetro de posición se calcula mediante un proceso iterativo de modo que el coeficiente de asimetría tienda a cero, es decir que indique que el histograma de la variable es simétrico.

B.2. Modelo probabilístico de Doble Gumbel

Fisher y Tippet encontraron que la distribución del valor máximo (o mínimo) seleccionado de muestras de tamaño n se aproxima a una distribución límite cuando el tamaño de la muestra se aumenta. Cuando las distribuciones iniciales representan colas que pueden aproximarse por funciones de tipo exponencial, se encuentra que los valores extremos tienden a tener distribución tipo I o Gumbel [6].

Fariás, Dos Santos y Palacios explican que mientras la función de valores extremos tipo I o Gumbel está pensada para una población, la distribución de Doble Gumbel permite modelar el comportamiento de dos poblaciones consideradas mutuamente excluyentes que intercalan su ocurrencia bajo una misma variable climatológica [7]. Es el caso de las precipitaciones en la costa norte de Perú, donde se presentan años normales y años con presencia de El Niño, con magnitudes totalmente diferentes a las observadas el resto de años.

En la referencia [8], se plantea una función de distribución de probabilidad para dos poblaciones:

$$F(x) = (1-p)F(x; \theta_1) + p.F(x; \theta_2) \quad (10)$$

Donde:

p : Probabilidad de que la variable x pertenezca a la segunda población, que agrupa a elementos sucedidos en condiciones climatológicas ordinarias.

Función de acumulada de Gumbel

$$F(x) = \exp \left(- \exp \left(- \frac{(x - \alpha)}{\beta} \right) \right) \quad (11)$$

Donde:

x : variable hidrológica

α : Parámetro de posición de la función

β : Parámetro de escala de la función

Estos parámetros se pueden estimar conociendo que la media y la desviación estándar de la función son respectivamente:

$$\mu = \alpha + \gamma\beta \quad (12)$$

$$\sigma = \beta \frac{\pi}{\sqrt{6}} \quad (13)$$

Siendo γ la constante de Euler-Mascheroni ($\gamma \approx 0.577216$).

De modo que si se tuviera una muestra de tamaño infinito se podrían estimar los parámetros de la función, α y β , a través de los estadígrafos de la muestra, con las expresiones (14) y (15), [7]:

$$\beta = s_x \frac{\sqrt{6}}{\pi} \quad (14)$$

$$\alpha = \bar{x} + \gamma\beta \quad (15)$$

Donde \bar{x} y s_x son la media y la desviación estándar muestrales, respectivamente.

Sin embargo, en la práctica se trabaja con muestras de tamaño finito, por tanto, el valor de los parámetros debe modificarse. Gumbel obtuvo valores modificados minimizando la sumatoria de cuadrados de los errores perpendiculares a la recta de ajuste de valores extremos. Las ecuaciones que obtuvo están en función del tamaño de la muestra, n , y de los estadígrafos [3]:

$$\beta = \frac{s_x}{s_n} \quad (16)$$

$$\alpha = \bar{x} + y_n \cdot \beta \quad (17)$$

Donde:

y_n : Valor esperado de la variable reducida y_m

s_n : Desviación estándar de la variable reducida y_m

Estos estadígrafos pueden obtenerse de la variable reducida ordenando la muestra de manera decreciente y calculando para cada x_i la correspondiente variable reducida y_m con la expresión (18) de Weibull:

$$y_m = -\ln \left[-\ln \left(\frac{n+1-m}{n+1} \right) \right] \quad (18)$$

Donde:

n : Tamaño de la muestra

m : Número de orden

En la referencia [3] se recomienda trabajar con tablas que entregan el valor de y_n y s_n directamente en función del tamaño de la muestra n , ya que éstos no dependen de los valores muestrales. Se propone el uso de las expresiones para evitar el uso de las tablas y facilitar el cálculo (19) y (20).

$$y_n = 0.0252 \ln(n) + 0.4477 \quad (19)$$

$$s_n = 0.0993 \ln(n) + 0.7635 \quad (20)$$

Función de distribución Doble Gumbel

Con base en la ecuación (10), la probabilidad de excedencia de Gumbel de dos poblaciones se expresa mediante la expresión (21), [7]:

$$F(X) = (1-p)e^{-e^{-\frac{(x-\alpha_1)}{\beta_1}}} + p.e^{-e^{-\frac{(x-\alpha_2)}{\beta_2}}} \quad (21)$$

Esta función de distribución cuenta con cinco parámetros, dos de cada población, α_1 , β_1 , α_2 y β_2 , mientras el quinto es la proporción de mezcla, p .

C. Pruebas de bondad de ajuste

Las pruebas de bondad de ajuste consisten en comprobar gráfica y estadísticamente si la frecuencia empírica de la serie analizada, se ajusta a una determinada

función de probabilidades teórica seleccionada a priori, con los parámetros estimados con base en los valores muestrales [5].

Las pruebas estadísticas tienen por objeto medir la certidumbre que se obtiene al hacer una hipótesis estadística sobre una población, es decir calificar el hecho de suponer que una variable aleatoria, se distribuya según una cierta función de probabilidad [5].

De uso frecuente son la prueba χ^2 , la de Kolmogorov-Smirnov y la del error cuadrático estándar. La primera trabaja con valores agrupados, por lo que su aplicación no es muy práctica; la segunda no requiere agrupar valores, se concentra en determinar el punto en el que estas dos distribuciones muestran la mayor divergencia, pero no evalúa las diferencias en su conjunto; mientras que la tercera es una mezcla de ambas pruebas, trabajando con los datos independientes y sumando las divergencias parciales entre las probabilidades de excedencia, teórica y empírica [7].

En este estudio se aplicó el test de Kolmogorov-Smirnov como medida de bondad del ajuste.

El test de Kolmogorov-Smirnov es una prueba no paramétrica que permite probar si dos muestras provienen del mismo modelo probabilístico. Este test consiste en comparar las diferencias absolutas entre dos funciones de distribución acumulada, la empírica y la teórica [3].

Por lo tanto, es necesario definir tanto la frecuencia observada o empírica acumulada, como la frecuencia teórica acumulada. En el primer caso, para determinar la frecuencia observada acumulada de la serie, se ordenan los datos de menor a mayor y se trabaja con la probabilidad de excedencia asignada a cada valor de la muestra, calculada con la fórmula de Weibull (22), como se cita en la referencia [9]:

$$P_{emp}(x) = \frac{m}{n+1} \quad (22)$$

Donde:

$P_{emp}(x)$: frecuencia empírica u observada acumulada.

m : número de orden

n : número total de datos

En el segundo caso, la frecuencia teórica acumulada, $F(x)_i$, se determina a través de la función probabilística teórica.

Cuando han sido obtenidas ambas frecuencias para cada evento de la serie analizada, se selecciona el valor máximo el valor absoluto, de la diferencia entre el valor observado y el valor de la recta teórica del modelo [5], es decir:

$$\Delta_{max} = \max |P_{emp}(x)_i - F(x)_i| \quad (23)$$

Una vez obtenido el valor del estadígrafo Δ_{max} , se compara con el $\Delta_{crítico}$, proveniente de tablas de valores críticos de esta prueba de bondad de ajuste, con un tamaño de muestra n y asumiendo un 95% de confianza o 5% de significancia. Luego se decide si el ajuste es o no adecuado, considerando las siguientes afirmaciones:

- Si $\Delta_{max} > \Delta_{crítico}$ se rechaza H_0 (el ajuste no es adecuado).
- Si $\Delta_{max} < \Delta_{crítico}$ no se rechaza H_0 (el ajuste es adecuado).

D. Determinación matemática de las curvas IDF

Las curvas IDF no sólo se construyen de manera gráfica, sino que también pueden expresarse como ecuaciones con el fin de evitar la imprecisión de lectura de la intensidad de lluvia de diseño en una gráfica [10].

El planteamiento de las primeras relaciones matemáticas, que modelan a dichas curvas, data desde el inicio de los años treinta del siglo pasado, y fueron realizadas por Sherman, en 1931 y Bernard, en 1932 [11]. Estos modelos han conducido a numerosos investigadores del mundo a generar diversas expresiones matemáticas que simulan las curvas IDF [12], algunas de las cuales se presentan a continuación:

Modelo propuesto por Bernard en [11]:
$$i_t^T = \frac{kT^m}{t^n} \quad (24)$$

Modelo propuesto por Chow [10]:
$$i_t^T = \frac{kT^m}{t^n + b} \quad (25)$$

Modelo propuesto por Koutsoyiannis et al. en [13]:
$$i_t^T = k \left\{ \frac{m - \ln[-\ln(1 - 1/T)]}{(t+b)^n} \right\} \quad (26)$$

Modelo propuesto por ASCE en [14]:
$$i_t^T = \frac{A}{t^B + C} \quad (27)$$

Modelo propuesto por ASCE en [14]:
$$i_t^T = \frac{a}{(t+b)^c} \quad (28)$$

Donde T es período de retorno, t es la duración e i_t^T la intensidad máxima de precipitación y A, B, C, a, b, c, k, m y n son parámetros adimensionales.

Los estudios para relacionar las lluvias de duraciones menores a un día, expresadas en valores horarios, con las lluvias diarias registradas, incluyen generalmente el periodo de retorno y el tiempo como variables [15].

Chen en [16], usando series anuales, propuso la siguiente fórmula para expresar la intensidad de una lluvia:

$$i_t^T = \frac{a \cdot i_t^{10} \cdot \log \left\{ 10^{2-x} \cdot [\ln(T/(T-1))] \right\}^{-(x-1)}}{(t+b)^c} \quad (29)$$

La relación es válida para periodos de retorno mayores a un año y para duraciones entre 5 minutos y 24 horas. Los parámetros a, b y c se obtienen de un gráfico presentado por dicho autor en función de la razón entre la lluvia de una hora y de 24 horas para el mismo periodo de retorno (K). El exponente x es la razón entre las lluvias de 1 hora de 10 y 100 años de periodo de retorno.

Varas y Farías [15] analizaron distintos trabajos regionales encontrados acerca de las relaciones IDF, los cuales se unificaron de acuerdo a la ecuación de Chen para series anuales, obteniendo una síntesis de las ecuaciones para determinar los parámetros regionales a, b y c empleados en dicha expresión según la región de aplicación, sobre la base del parámetro k. Dicha combinación de parámetros entregó valores parecidos entre los diferentes trabajos en el rango de duración de la lluvia entre 1 y 24 horas.

III. RECOPIACIÓN DE INFORMACIÓN HISTÓRICA

Se contó con información histórica de lluvias de 45 años entre las estaciones meteorológicas Miraflores y Universidad de Piura (UDEP), por ser las más representativas de la zona. Ambas se ubican en la ciudad de Piura, en la provincia y departamento del mismo nombre, a poco más de 2 km de distancia. La Tabla I presenta los detalles de ubicación de las estaciones.

TABLA I
UBICACIÓN DE LAS ESTACIONES PLUVIOGRÁFICAS DE ESTUDIO

Estación	Longitud W	Latitud S	Distrito	Categoría
Miraflores	80°36'59"	05°10'31"	Castilla	PLU
Udep	80°38'18"	05°10'14"	Piura	PLU-PG

La estación Miraflores empezó su operación en el año 1972 contando con pluviógrafo y pluviómetro, a cargo del Proyecto Especial Chira Piura (PECHP), hasta el año 1992; y en 1993 pasó a cargo del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI), hasta la actualidad, en que cuenta con un pluviómetro convencional.

La información histórica de lluvias de la estación Miraflores fue brindada por el departamento de Hidrometeorología del PECHP, consistente en las precipitaciones diarias máximas anuales de los años 1971 al 2017, intensidades máximas para diversas duraciones y años, y bandas pluviográficas de los años 1984, 1987, 1988, 1989 y 1990.

La estación meteorológica UDEP se encuentra en el Campus de la Universidad de Piura, a cargo del Laboratorio de Física General y empezó su operación en febrero de 1991. Esta estación meteorológica cuenta con un pluviógrafo basculante Campbell Scientific; modelo TE525 con precisión de 0.25 mm.

El Instituto de Hidráulica, Hidrología e Ingeniería Sanitaria de la Universidad de Piura (IHHS) proporcionó la base de datos pluviográficos de la estación meteorológica UDEP, años 1995-1998 y 2005-2017, consistente en una base de datos de lluvia en intervalos de 30 minutos, salvo los datos de 2017 que están configurados cada 10 minutos.

IV. PROCESAMIENTO DE LA INFORMACIÓN

A. Análisis de tormentas

Obtenida toda la información histórica de lluvias posible, se procedió a realizar el análisis de tormentas respectivo con la finalidad de obtener las intensidades máximas de cada año hidrológico (01 septiembre-31 de agosto) y así actualizar la base de datos de intensidades máximas). Se describe a continuación el procesamiento de la información recogida en cada estación meteorológica.

La base de datos de lluvias cada media hora de la estación Udep se procesó en hoja de cálculo Excel, determinando las intensidades de precipitación para periodos de duración de 30, 60, 120, 180, 240, 360, 720 y 1440 minutos a partir todos los datos. Se obtuvieron las intensidades máximas anuales para cada duración.

De las bandas pluviográficas de las máximas tormentas de la estación Miraflores se realizó el correspondiente análisis, a fin de extraer la información numérica a hojas de cálculo, y de esta manera obtener las intensidades máximas

para cada duración, tal como se realizó con los registros históricos de la estación UDEP, y para cada año hidrológico.

Como resultado del análisis de las dos estaciones se actualizó la matriz de precipitaciones máximas diarias e intensidades, obteniendo una versión preliminar, que contiene información de los años 1972 a 2017; con excepción del año 1993, del cual no se encontraron registros en ninguna de las dos estaciones. Esta matriz está completa en un 73% y se requiere completarla para elaborar las curvas IDF, sobre todo porque de algunos años solo se dispone de registros de precipitaciones máximas diarias.

B. Relleno de datos faltantes de intensidades

El primer paso fue completar las intensidades en los años que presentaban menor cantidad de datos faltantes. Así, se inició con el año hidrológico 2016-2017 determinando la intensidad de 15 minutos de duración, i_{15} , interpolando las intensidades adyacentes, i_{10} e i_{30} del mismo año, y luego la i_{240} en diversos años faltantes, interpolando del mismo modo los valores anuales correspondientes a la i_{180} e i_{360} (Fig. 2).

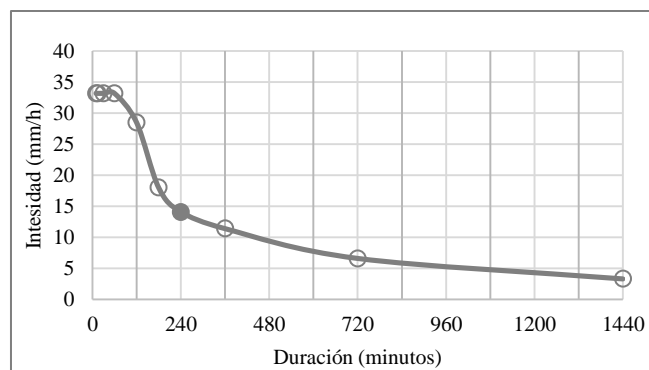


Fig. 2 Ejemplo de determinación de intensidad (i_{240}) por interpolación.

A continuación se completó la serie i_{60} , por presentar pocos datos faltantes y ser representativa del comportamiento pluviográfico, correlacionándola con la P_{max} día (Fig. 3).

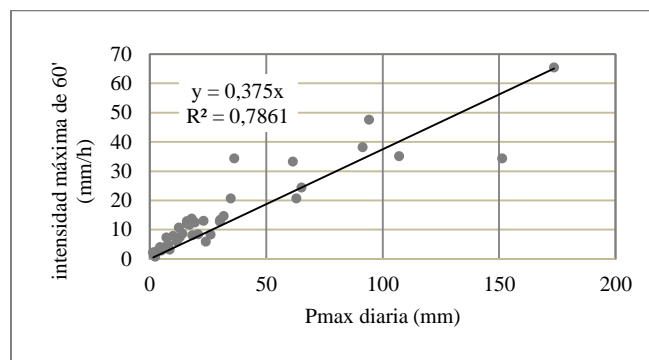


Fig. 3. Correlación de la serie i_{60} y P_{max} día.

Finalmente, se evaluó la correlación de las series de intensidades restantes respecto a las series completas, P_{max} día e i_{60} , efectuando el relleno con aquella que presentara mejor correlación (Fig. 4). Esto permitió completar todos los datos faltantes y obtener la matriz de intensidades máximas anuales para el período 1972-2017, exceptuando el año 1993 sin registros.

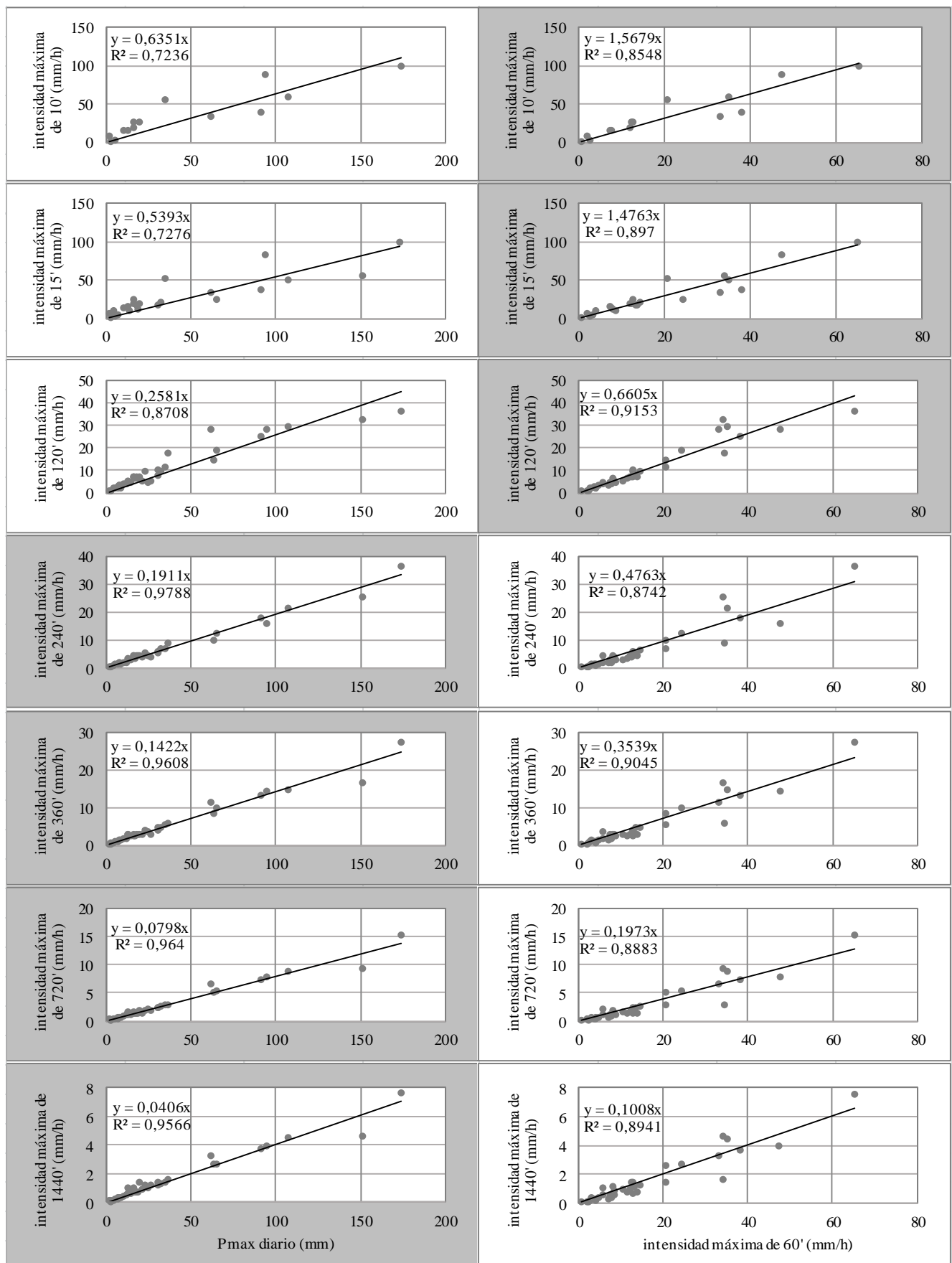


Fig. 4 Evaluación de la correlación de las diferentes series de intensidades máximas con la serie de P_{max} día (izquierda) y con i_{60} (derecha) para el relleno de datos faltantes. Los recuadros sombreados indican las mejores correlaciones que se han usado para el relleno.

C. Definición de la base de datos

La Región Piura está ubicada en la zona árida del norte peruano, presenta años muy secos y algunos años de considerable cantidad de lluvia, producto de la presencia de Fenómenos El Niño (FEN), ya sea de carácter global (ENSO) o costero. La Fig. 5 muestra la variabilidad de las precipitaciones diarias máximas anuales para el período 1972-2017, donde se aprecia que a partir de 1983 el régimen hídrico de la zona podría haber aumentado.

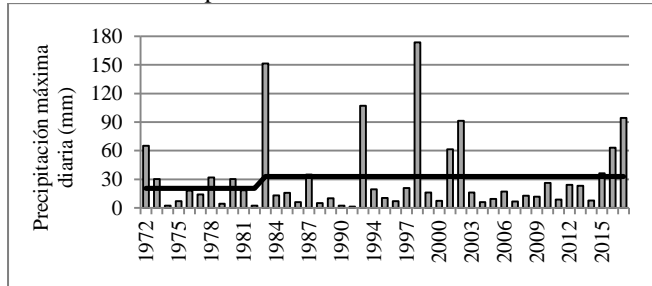


Fig. 5 Variabilidad de las precipitaciones máximas diarias en Piura y su media.

Existen evidencias de que los caudales máximos del río Piura (Fig. 6) presentan este mismo cambio en la serie aún más larga, entre 1926-2008 C. Su trabajo analiza estadísticamente esta hipótesis y propone una división de la serie hidrométrica, basada en que, en los últimos años, la intensidad de los FEN y sus respectivos caudales han incrementado, dejando prever que posiblemente la serie ha cambiado y por lo tanto las series 1926-1982 y 1983-2008 pueden no pertenecer al mismo universo.

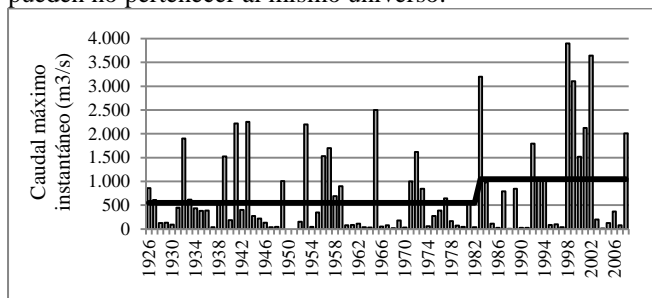


Fig. 6 Variabilidad de los caudales máximos anuales en el río Piura.

Basados en los hallazgos antes mencionados, se dividió la matriz de series de intensidades máximas anuales en dos grupos de series para analizar si todas las series pertenecen o no a un mismo universo; y decidir en base a ello qué series se propondría para la construcción final de las curvas IDF.

Al dividir la base de datos y analizar dichos grupos de series de datos, se observa que estas series tienen parámetros estadísticos diferentes. Sin embargo, y de acuerdo con las conclusiones de varios otros estudios a nivel mundial, se considera que estas diferencias no pueden ser directamente consideradas como la confirmación de cambios en las tendencias y parámetros principales de una cuenca; pero se puede analizar si existe diferencia en las medias y desviaciones estándar de las divisiones de las series mediante pruebas de hipótesis [17]

Para realizar el análisis se escogieron tres variables representativas de toda la matriz: Pmax día, i_{10} e i_{60} . Los datos de cada variable fueron divididos en dos series, A

(1972-1982) y B (1983-2017), de acuerdo a lo propuesto en [17].

C.1. Estadística descriptiva de las series

Para verificar y comprobar la calidad y confiabilidad de las series antes mencionadas, primero se evaluaron los principales estadígrafos para apreciar en qué medida los datos se agrupan o dispersan en torno a ese valor central (Tabla II y Tabla III). Los estadígrafos se aprecian muy diferentes entre las series. Las series B (1983-2017) superan notablemente los valores en diferentes parámetros de las series completa y A; por tal motivo se decidió continuar con las pruebas de hipótesis necesarias para comprobar si realmente difieren dichos parámetros de cada una de las series.

TABLA II
ESTADÍSTICA DESCRIPTIVA DE LA PMAX DIA

Parámetros	Serie total	Serie A	Serie B
Media	29.8	20.4	32.9
Mediana	16.0	18.1	16.0
Desviación estándar	38.4	18.7	42.7
Varianza	1474	348	1821
Curtosis	5.4	2.5	3.9
Coefficiente de asimetría	2.3	1.4	2.1
Rango	172.1	63.1	172.1
Mínimo	1.5	2.2	1.5
Máximo	173.6	65.3	173.6
Cuenta	45.0	11.0	34.0

TABLA III
ESTADÍSTICA DESCRIPTIVA DE LA i_{60}

Parámetros	Serie total	Serie A	Serie B
Media	13.2	9.6	14.4
Mediana	8.2	8.6	8.0
Desviación estándar	13.9	7.0	15.4
Varianza	193	48	237
Curtosis	3.8	0.5	2.6
Coefficiente de asimetría	1.9	0.7	1.7
Rango	64.6	23.5	64.6
Mínimo	0.7	0.8	0.7
Máximo	65.3	24.3	65.3
Cuenta	45.0	11.0	34.0

C.2. Aplicación de las pruebas de hipótesis a las series

Para la aplicación de la prueba de hipótesis de comparación de medias a las series elegidas se planteó como hipótesis nula: “El promedio de las series B es igual al promedio de las series A” y como hipótesis alternativa, que existe diferencia significativa entre los promedios.

Se prosiguió evaluando el tamaño de las muestras, menores que 50 en ambos casos, 11 y 34 años respectivamente. Por lo explicado, se aplicó la prueba de t-Student, obteniendo los resultados de la Tabla IV y la Tabla V para las variables Pmax día e i_{60} , respectivamente.

TABLA IV
PRUEBA T-STUDENT APLICADA A LAS SERIES DE PMAX DIA

Parámetros	Serie A	Serie B
Media	20.4	32.9
Varianza	347.9	1820.8
Observaciones	11.0	34.0
Varianza agrupada	1478.3	
Diferencia hipotética de las medias	0	
Grados de libertad	43.0	
Estadístico t	0.938	
P(T<=t) una cola	0.180	

Con los resultados de los estadígrafos t , se obtuvieron las probabilidades de excedencia de 0.18 y 0.16, respectivamente, lo que representa que, con esos niveles de significancia o mayores, 20% por ejemplo, se puede admitir que las medias de los dos grupos de series estudiadas pertenecen a universos diferentes.

TABLA V

PRUEBA T-STUDENT APLICADA A LAS SERIES DE LA VARIABLE I_{60}

Parámetros	Serie A	Serie B
Media	9.6	14.4
Varianza	48.4	237.0
Observaciones	11	34
Varianza agrupada	193.1	
Diferencia hipotética de las medias	0	
Grados de libertad	43	
Estadístico t	0.992	
$P(T \leq t)$ una cola	0.163	

Por otro lado, en C.1 se puede observar que los valores de las varianzas de las series B son notablemente mayores a los de las series A. Esta evidencia no es suficiente para rechazar la hipótesis de que las muestras tienen varianzas iguales, por lo que se aplicó la prueba de hipótesis F de comparación de dos varianzas, para lo cual se estableció como hipótesis nula que las muestras provienen de universos con varianzas iguales y como hipótesis alternativa que la varianza de la serie B es mayor a la varianza de las series A.

Los valores F calculados a partir de las muestras son mayores al F crítico, entonces podemos rechazar la hipótesis nula (Tabla VI y Tabla VII). Por lo tanto, se puede considerar aceptar la hipótesis de que las muestras no provienen de universos con la misma varianza, con un nivel de significancia del 5%.

TABLA VI

PRUEBA F APLICADA A LAS SERIES DE LA VARIABLE P_{MAX} DÍA

Parámetros	Serie A	Serie B
Media	32.91	20.40
Varianza	1820.8	347.9
Observaciones	34	11
Grados de libertad	33	10
F calculado	5.23	
$P(F \leq f)$ una cola	0.00427	
Valor crítico para F (una cola: 5%)	2.69	

TABLA VII

PRUEBA F APLICADA A LAS SERIES DE LA VARIABLE I_{10}

Parámetros	Serie C	Serie B
Media	14.42	9.64
Varianza	236.95	48.40
Observaciones	34	11
Grados de libertad	33	10
F calculado	4.90	
$P(F \leq f)$ una cola	0.00558	
Valor crítico para F (una cola: 5%)	2.69	

C.3. Base de datos final

Basados en el análisis de la estadística descriptiva, las pruebas de hipótesis y el trabajo realizado en [17], se concluye que las series A y B provienen de universos diferentes, esto es que se está ante un cambio en las características climatológicas de las lluvias en la ciudad de

Piura, que debe ser considerado para determinar las lluvias de diseño de la infraestructura.

A la vista de los resultados de la aplicación de las pruebas de medias y varianzas, se decidió trabajar únicamente con las series B (1983-2017) para el cálculo.

D. Modelación probabilística

De los diversos modelos ajustados a las series de intensidades, los que presentaron mejor ajuste fueron el modelo Log Normal de tres parámetros y el modelo Doble Gumbel.

D.1. Ajuste del modelo Log Normal 3p a las series

En la Tabla VIII se presentan los parámetros determinados mediante el método de máxima verosimilitud.

TABLA VIII

PARÁMETROS DE LA DISTRIBUCIÓN LOG-NORMAL 3P AJUSTADA A CADA SERIE DE INTENSIDADES Y P_{MAX} DÍA

Parámetro	P_{mx} día	i10	i15	i30	i60	i120	i180	i360	i720	i1440
μ_y	2.78	2.76	2.66	2.44	2.16	1.63	1.37	0.00	0.00	0.00
σ_y	1.22	0.00	1.02	1.07	1.06	1.15	1.17	1.15	1.17	1.21
x_o	0.73	-0.38	-0.23	0.20	-0.03	0.34	0.25	0.06	0.03	0.04

D.2. Ajuste del modelo Doble Gumbel a las series

El ajuste de las series a modelos Doble Gumbel se inicia con la selección del parámetro p , a partir de la observación de los datos, considerando que agrupa a los eventos climatológicamente ordinarios, esto es, sin evidencia de presencia de El Niño. Luego de definidos todos los parámetros y evaluado el ajuste logrado, se podrá variar el valor del parámetro p para intentar alcanzar un mejor ajuste. La evaluación hecha arrojó que la mejor división de las series de estudio fue con 24 años normales y 10 años extraordinarios o con presencia de El Niño. Una vez divididas las series la determinación de los parámetros restantes se realiza de manera independiente en cada una como se explicó en II.B.2 (Tabla IX).

TABLA IX

PARÁMETROS DE LA DISTRIBUCIÓN DOBLE GUMBEL AJUSTADA A CADA SERIE DE INTENSIDADES Y P_{MAX} DÍA

Serie	P_{max} día	i10	i15	i30	i60	i120	i180	i360	i720	i1440
n1	10									
m1	84.0	54.5	51.7	45.9	34.2	23.4	18.8	12.1	6.8	3.5
s1	49.7	24.0	23.6	23.9	14.8	9.3	9.7	6.9	3.9	1.9
yn1	0.495									
Sn1	0.950									
a1	52.3	25.2	24.9	25.2	15.6	9.8	10.2	7.3	4.1	2.0
u1	58.0	42.0	39.4	33.4	26.4	18.6	13.7	8.5	4.8	2.5
n2	24									
m2	11.6	11.5	10.5	8.7	6.2	3.8	2.9	1.8	1.0	0.6
s2	6.3	7.1	6.6	5.3	3.7	2.0	1.5	0.9	0.5	0.3
yn2	0.530									
Sn2	1.086									
a2	5.8	6.6	6.1	4.9	3.4	1.8	1.4	0.9	0.5	0.3
u2	8.6	8.0	7.3	6.1	4.4	2.9	2.2	1.3	0.7	0.4

D.3. Bondad de ajuste de los modelos ajustados

Se aplicó el test de bondad de ajuste Kolmogorov-Smirnov para los ajustes realizados con los modelos probabilístico Log-Normal 3p y Doble Gumbel (Tabla X).

TABLA X
VALORES DE DELTA MÁXIMO PARA LA DETERMINACIÓN DE LA BONDAD DE AJUSTE CON EL TEST KOLMOGOROV – SMIRNOV

Delta máximo	Pmax día	i10	i30	i60	i120	i360	i720	i1440
LN 3	0.070	0.067	0.074	0.098	0.087	0.108	0.102	0.083
Doble Gumbel	0.079	0.085	0.091	0.059	0.057	0.105	0.084	0.080

Con lo observado en la Tabla X, se puede afirmar que los ajustes de las series a los modelos probabilísticos Log-Normal 3p y Doble-Gumbel son adecuados, pues en todos los casos se obtienen valores de delta máximo, Δ_{\max} , menores que el delta crítico, $\Delta_{\text{crítico}}$, con 5% de significancia, que para el tamaño de la muestra trabajada, corresponde a 0.233.

E. Curvas IDF

E.1. Determinación de la matriz de intensidades definitiva

Al observar que ambos tipos de modelo arrojan valores relativamente semejantes, se decidió promediar sus predicciones para formar una única matriz de intensidades (Tabla XI), que permitió graficar las curvas IDF requeridas (Fig. 7).

TABLA XI
MATRIZ DE PRECIPITACIONES MÁXIMAS (MM) E INTENSIDADES MÁXIMAS ANUALES (MM/H) EN FUNCIÓN DE LA DURACIÓN (MIN) Y EL PERÍODO DE RETORNO T (AÑOS)

T (años)	Pmax día (mm)	i10	i15	i60	i120	i360	i720	i1440
5	48.5	37.8	35.3	22.8	15.5	7.2	4.0	2.1
10	90.6	59.6	56.7	36.8	24.9	13.1	7.3	3.8
25	147.5	88.9	85.7	55.8	37.9	20.9	11.8	6.1
50	196.9	113.3	110.0	71.9	49.4	27.6	15.6	8.1
100	254.1	140.3	137.2	90.2	62.8	35.2	19.9	10.5
150	291.9	157.6	154.6	102.0	71.6	40.2	22.7	12.0
200	321.0	170.5	167.8	110.9	78.4	44.0	24.9	13.2

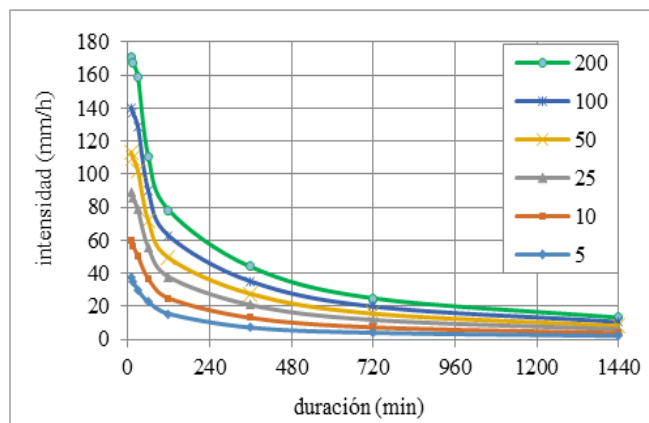


Fig. 7 Curvas IDF para la ciudad de Piura (1983-2017).

E.2. Determinación de la expresión matemática

Basados en [15] se propone, a partir de la matriz de intensidades obtenidas para la zona de estudio (Tabla XI), determinar una sola expresión para toda la familia de curvas IDF de la forma:

$$i_t^T = \frac{\alpha(T)}{(t+b)^c} \quad (30)$$

Para ello, se trabajó preliminarmente con la expresión (28), hallando los parámetros a , b y c , es decir trabajando por separado cada periodo de retorno. En este análisis se obtuvo una buena correlación (Tabla XII).

TABLA XII
VALORES DE LOS PARÁMETROS Y EL COEFICIENTE DE CORRELACIÓN PARA CADA PERÍODO DE RETORNO

Tr	a	b	c	R ²
5	1.933	58,82	0,930	1,000
10	3.274	67,90	0,919	0,998
25	5.131	73,47	0,915	0,997
50	6.890	77,59	0,916	0,996
100	9.077	81,97	0,919	0,996
150	10.605	84,68	0,921	0,996
200	11.820	86,68	0,923	0,996

Se puede observar que los valores de los parámetros b y c no muestran mayores diferencias, por lo cual se puede establecer un valor único para dichos parámetros. Por tal motivo, se decidió encontrar las expresiones apropiadas para el numerador, N , y el denominador, D , de la expresión (30).

En primer lugar se procedió a asignar un valor común a b , tomando el promedio de los valores hallados para cada periodo de retorno. Así, se llegó a establecer $b = 75$, y con este valor único de b se recalcularon los valores de c para cada periodo de retorno. Con el mismo criterio, se estableció un único valor de c con el promedio de los nuevos valores, obteniendo como resultado $c = 0.912$. Con los valores únicos establecidos de $b = 75$ y $c = 0.912$ para todos los periodos de retorno, el denominador de la ecuación (30) quedó expresado como:

$$D = (t+b)^c = (t+75)^{0.912} \quad (31)$$

Seguidamente se prosiguió a establecer la expresión para el numerador de la ecuación (30), N , en función del periodo de retorno, T . Esto se logra estimando valores de N como el producto de las intensidades de la Tabla XI i_{matriz} , por el denominador calculado con la expresión (31) y ajustando una expresión matemática a dichos valores:

$$N = i_{\text{matriz}} * D \quad (32)$$

La Fig. 8 grafica el periodo de retorno, T , y el numerador, N , para encontrar la expresión matemática buscada para el numerador de la expresión (30).

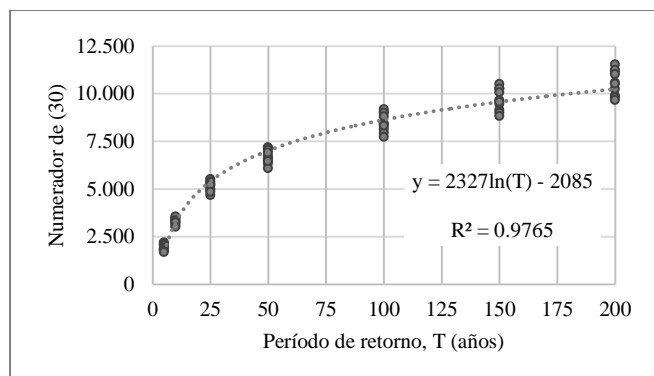


Fig. 8 Ajuste del numerador en función del periodo de retorno.

Finalmente, la ecuación queda expresada de la siguiente manera:

$$i_t^T = \frac{2327 \ln(T) - 2085}{(t + 75)^{0.912}} \quad (33)$$

Para dar validez a la ecuación (33) se correlacionaron las intensidades de la matriz (i_{matriz}) y las brindadas por la ecuación ajustada, dando como resultado un coeficiente de correlación bastante cercano a la unidad (Fig. 9), con lo que se da por apropiada la expresión obtenida de la intensidad de lluvia en función del período de retorno y la duración.

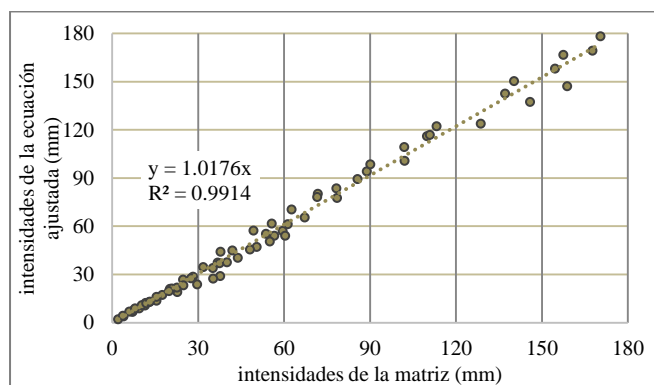


Fig. 9 Correlación entre las intensidades de la matriz y las obtenidas por la ecuación ajustada.

CONCLUSIONES

Las curvas IDF son una herramienta fundamental para la modelación hidrológica orientada a determinar los caudales de diseño de obras ubicadas en lugares sin registros hidrométricos que permitan modelar probabilísticamente los caudales históricos. Usualmente se determinan a partir de la modelación de las tormentas históricas, asumiendo que el clima tiene la misma tendencia a lo largo del tiempo, es decir que es estacionario, lo cual no es correcto ante las evidencias de cambio climático que se registran en diversas partes del mundo.

En el Perú, la presencia recurrente de eventos El Niño, costero y global, genera que en los últimos años, a partir de El Niño de 1983, los registros de precipitaciones de la ciudad de Piura, indiquen un cambio en el clima, con años lluviosos más frecuentes e intensos, contando con tres eventos extraordinarios, 1983, 1998 y 2017 y otros de moderada intensidad. Este cambio se ha verificado estadísticamente al aplicar pruebas de hipótesis de medias y

de varianzas. La serie histórica de 1972 a 2017 fue dividida en dos: 1972-1982 y 1983-2017, encontrando que estadísticamente podrían provenir de universos diferentes, esto quiere decir que la pluviometría de ambos períodos es diferente.

Se propone actualizar las curvas IDF existentes para incluir los últimos eventos y evaluar recortar apropiadamente las series históricas para trabajar con únicamente con datos más recientes, que reflejen la tendencia climática no estacionaria actual.

A las intensidades definidas para la ciudad de Piura, como caso de estudio, se les ajustó modelos Log-Normal 3p y Doble Gumbel, cuyas predicciones se han promediado para obtener las curvas actualizadas y una ecuación general que relaciona las tres variables de diseño para mayor practicidad.

REFERENCIAS

- [1] S. Fadhel, M. Rico-Ramirez y D. Han., "Uncertainty of Intensity–Duration–Frequency (IDF) curves due to varied climate baseline periods", *Journal of Hydrology*, vol. 547, pp. 600-612, Abril 2017.
- [2] G. Mirhosseini, P. Srivastava y L. Stefanova (2013). *Reg Environ Change* 13 (Suppl 1):25. <https://doi.org/10.1007/s10113-012-0375-5>
- [3] E. Varas, y P. Bois, "*Hidrología Probabilística*". Chile: Ediciones Universidad Católica de Chile, 1998.
- [4] P. Acosta, y L. Sierra, "Evaluación de métodos de construcción de curvas IDF a partir de distribuciones de probabilidad y parámetros de ajuste", *Revista Facultad de Ingeniería, UPTC*, Vol. 22, no 35, pp.25-33, Jul-Dic 2013.
- [5] M. Villón, "*Hidrología Estadística*". Lima: Villón, 2005.
- [6] R. Linsley, M. Khohler y J. Paulhus, "*Hidrología para ingenieros*". D. F. México: Programas Educativos. 1992
- [7] M. Farias, E. Dos Santos y C. Palacios S., "Estimación de Caudales de Diseño con la Distribución Doble Gumbel", en XIX Simposio Brasileiro de Recursos Hídricos, Maceió, Brasil, pp.1-10, Noviembre – Diciembre 2011.
- [8] J. Raynal y M. Raynal, "Cálculo de los límites de confianza de la Distribución de Probabilidad de valores extremos tipo I para dos poblaciones". *Información Tecnológica*, vol. 15, no 1, pp.87-94, 2004.
- [9] R. Pizarro, D. Aravena, K. Macaya, A. Abarza, M. Cornejo, M. Labra, M. Pavez y L. Román. "Curvas Intensidad – Duración – Frecuencia, para la zona centro sur de Chile". UNESCO, Editorial Universidad de Talca, Chile, Documentos Técnicos del PHI-LAC No7, 2007.
- [10] V. Chow, L. Mays y D. Maidment, "*Hidrología aplicada*". Santa Fe de Bogota, Colombia: Mc Graw-Hill. 1994
- [11] D. Pereyra, J. Pérez y L. Gómez, "Ecuaciones que estiman las curvas intensidad-duración-período de retorno de la lluvia", *GEOS*, vol. 24, no 1, pp.46-45, Agosto 2004.
- [12] G. Remenieras, "*Tratado de Hidrología Aplicada*". Madrid. España: Editores Asociados, 1971.
- [13] D. Koutsoyiannis, D. Kozonis y A. Manetas, "A mathematical framework for studying rainfall intensity-duration-frequency relationships", *Journal of Hydrology*, vol. 206, pp. 118-135, 1998.
- [14] ASCE. "*Hydrology Handbook*". Nueva York, Estados Unidos de America: ASCE. 1996.
- [15] E. Varas y M. Farias, "Estudio Comparativo de Relaciones Intensidad-Duración-Frecuencia". En Memorias XIX Congreso Latinoamericano de Hidráulica, Córdoba, Argentina, Tomo II, pp. 359-366. 2000.
- [16] C. Chen, "Rainfall intensity-duration-frequency formulas", *Journal of Hydraulic Engineering*, ASCE, vol 109, no 12, pp. 1603-1621. 1983.
- [17] C. Palacios, "Caudales de diseño en el Río Piura y su variación histórica ante El Fenómeno El Niño", Tesis para obtener el título de Ingeniero Civil, Universidad de Piura, Perú, 2010.

Determinación de ecuación Intensidad-Duración-Frecuencia en presencia de ENSO y Niño Costero. Caso: ciudad de Piura

Marina Farías de Reyes, Magister en Ingeniería¹ y Milagros Ruiz Tesen, Bachiller en Ingeniería¹

¹University of Piura, Perú, marina.farias@udep.pe, mruiz.t92@gmail.com

Abstract— Las curvas IDF son una herramienta fundamental para modelar los caudales de diseño en lugares sin registros hidrométricos. Usualmente se determinan a partir de la modelación de tormentas históricas, asumiendo que el clima tiene la misma tendencia a lo largo del tiempo, lo que no es correcto ante las evidencias de cambio climático. En el Perú, la presencia recurrente de eventos El Niño, costero y global, genera que en los últimos años, a partir de El Niño de 1983, los registros de precipitaciones indiquen un cambio en el clima, corroborado estadísticamente. Se propone actualizar las curvas IDF existentes para incluir los últimos eventos y evaluar recortar apropiadamente las series históricas para trabajar con series más recientes que refleje la tendencia climática actual. A las intensidades definidas para la ciudad de Piura, como caso de estudio, se les ha ajustado modelos Log-Normal 3p y Doble Gumbel, cuyas predicciones se han promediado para obtener las curvas actualizadas y una ecuación general que relaciona las tres variables de diseño.

Keywords-- Curvas IDF, Distribuciones de probabilidad, Log-Normal 3, Doble Gumbel, ENSO.

I. INTRODUCCIÓN

Para diseñar las obras hidráulicas que tienen por finalidad controlar la escorrentía y proteger la infraestructura es necesario realizar un estudio hidrológico previo, que permita determinar el caudal que se requerirá evacuar. Por lo general, no se cuenta con registros de caudales para modelarlos probabilísticamente, como es el caso en el diseño de reservorios o algunos puentes, de modo que en su ausencia es necesario recurrir a modelación hidrológica de la cuenca de interés. Así, los sistemas de manejo de aguas pluviales tendrán en las curvas de Intensidad-Duración-Frecuencia (IDF) una herramienta de diseño estándar.

Cabe resaltar que, los cuantiles de precipitación extrema representados por las curvas IDF estarán sujetos a alteraciones a lo largo del tiempo [1]. Esto debido a los cambios en el ciclo hidrológico generados por el aumento de los gases de efecto invernadero. Una forma de reducir la vulnerabilidad es cuantificar los efectos potenciales del cambio climático y adaptarse a ellos [2]. En la referencia [1] se indica que en la actualidad, un enfoque común es adoptar un único período de referencia para determinar curvas IDF, y comparan curvas IDF desarrolladas tomando datos históricos de diferentes períodos de tiempo, así como escenarios de clima futuros proyectados, encontrando que hay una diferencia significativa en las intensidades de lluvia. Ante esas evidencias, en este trabajo se destaca la importancia de diseñar las obras con curvas IDF elaboradas con datos de precipitaciones actualizados y que reflejen la

tendencia actual del comportamiento climático, frente al enfoque clásico que considera el clima invariable, y busca contar con series históricas lo más largas posibles, y por tanto antiguas, para representar un universo asumido en régimen permanente. Se presenta la metodología de determinación de las curvas IDF para la ciudad de Piura, en la costa norte peruana, considerando la presencia recurrente de Fenómenos El Niño (FEN), ya sea de carácter global (ENSO) o costero, llegando a obtener una ecuación general para la intensidad en función de la duración de la lluvia y el período de retorno en la ciudad.

II. MARCO TEÓRICO

El desarrollo del presente estudio involucra varios pasos diferentes, como se muestra en la Fig. 1 y se describen con mayor detalle a lo largo de esta sección. El enfoque metodológico tuvo como objetivo la obtención de las curvas IDF para la Ciudad de Piura con información de precipitación actualizada; y posteriormente la determinación de la expresión matemática de las mismas.

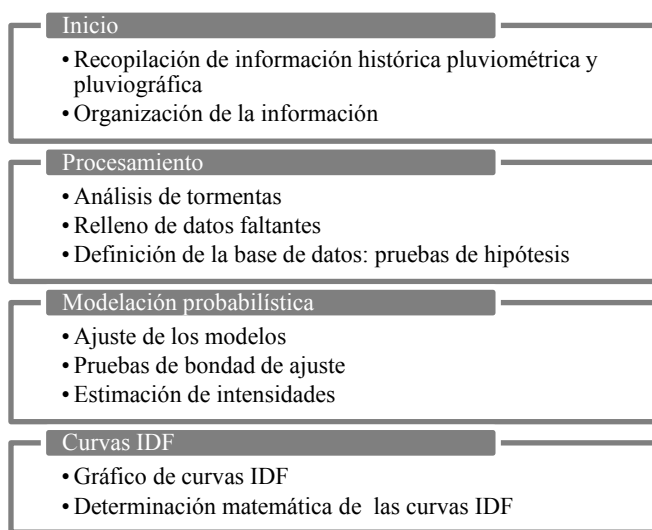


Fig. 1. Secuencia de trabajo empleada para el análisis

A. Pruebas de hipótesis

Estas pruebas de hipótesis se realizaron con la finalidad de evaluar si la media y la desviación estándar de las series A y B indican que el universo ha cambiado, es decir si hay evidencia estadística de un cambio climático en la serie.

A.1. Prueba T-Student para comparación de medias

En la referencia [3], se explica que esta prueba de hipótesis se utiliza para saber si las medias de dos muestras

son significativamente diferentes. Para tal fin, se deben distinguir dos casos: muestras grandes, por ejemplo, mayor de 50 valores, y muestras pequeñas.

Si los tamaños de ambas muestras son suficientemente grandes, se puede demostrar que el estadígrafo z definido por la expresión (1), tiene una distribución normal con promedio 0 y varianza 1; y la probabilidad de sobrepasar el valor del estadígrafo calculado, se obtiene de las tablas de la distribución normal [3].

$$z = \frac{m_1 - m_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (1)$$

En el caso en el que el tamaño de alguna o de ambas muestras sea más reducido, lo cual es una situación común en hidrología, entonces el estadígrafo t definido por la ecuación (2), sigue una ley de distribución de Student [3]:

$$t = \frac{m_1 - m_2}{s \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (2)$$

Con una varianza que tiene en cuenta los tamaños efectivos de las muestras:

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (3)$$

Si la hipótesis de que las muestras provienen de una misma distribución es verdadera, el estadígrafo t sigue una ley de distribución de Student con $n_1 + n_2 - 2$ grados de libertad. Por consiguiente, se debe obtener de la tabla de distribución de Student la probabilidad de sobrepasar el valor del estadígrafo calculado, bajo el supuesto de la hipótesis nula, es decir, que ambas muestras provienen de la misma población o universo [3].

A.2. Prueba F para comparación de dos varianzas

En esta prueba, se define el estadígrafo F como la razón entre la varianza mayor y la varianza menor. Si ambas varianzas provienen de la misma distribución, aproximadamente normal, el estadígrafo F sigue una distribución llamada ley de Fisher-Snedecor [3], con $g_1 = n_1 - 1$ y $g_2 = n_2 - 1$ grados de libertad.

$$F = s_1^2 / s_2^2 \quad (4)$$

Por consiguiente, se debe obtener de la tabla de distribución de Fisher la probabilidad de sobrepasar el valor del estadígrafo calculado, bajo el supuesto de la hipótesis nula, es decir, que ambas muestras provienen de la misma población o universo.

B. Modelación probabilística

Los datos sistematizados deben ser modelados probabilísticamente, con el fin de relacionar la magnitud y la frecuencia de los eventos. Cada modelo o función de distribución tiene sus propios parámetros. Es común el empleo de la media aritmética, relacionada con la magnitud de las variables. De igual manera, se halla la dispersión o variación de los datos de la muestra, utilizando la desviación estándar de cada muestra [4].

Existen diversos modelos probabilísticos que se pueden ajustar a las variables, a través de los cuales es posible predecir su comportamiento en función del período de retorno. Para realizar estimaciones con estos modelos se debe evaluar previamente su bondad, realizando una prueba de bondad ajuste [5].

Las distribuciones de probabilidad comúnmente utilizadas en Hidrología son entre otras:

- Distribución Normal, principalmente sus variantes Log-Normal de 2 ó 3 parámetros.
- Distribución Gamma, con sus variantes Exponencial, Gamma de 2 parámetros, Gamma de 3 parámetros o Pearson tipo III y Log-Pearson tipo III.
- Distribución Gumbel, con sus variantes Doble Gumbel y Log-Gumbel.

Para este análisis las distribuciones que se ajustaron de manera satisfactoria a las series de intensidades máximas diarias fueron la de Doble Gumbel y la Log-Normal de 3 parámetros.

B.1. Modelo probabilístico de Log-Normal de 3 parámetros

Las variables físicas de interés en hidrología (precipitación, caudal, evaporación y otras) son generalmente positivas, por lo cual es usual que presenten distribuciones de frecuencias asimétricas [3].

Algunos investigadores han propuesto aplicar una transformación logarítmica a la variable de interés y luego utilizar el modelo de distribución normal para la variable transformada. La distribución así obtenida se denomina Log-Normal de dos parámetros, que son la media y la desviación estándar de la variable transformada y , tal que [3]:

$$y = \ln(x) \quad (5)$$

Asimismo, hay una distribución Log-Normal de 3 parámetros, donde el tercer parámetro, denominado parámetro de posición, es el límite inferior x_o tal que la variable transformada es [5]:

$$y = \ln(x - x_o) \quad (6)$$

Función de densidad Log-Normal 3p

$$f(x) = \frac{1}{(x - x_o) \sigma_y \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left[\frac{\ln(x - x_o) - \mu_y}{\sigma_y} \right]^2\right) \quad (7)$$

Para $x_o \leq x < \infty$, donde:

μ_y : Parámetro de escala

σ_y : Parámetro de forma

x_o : Parámetro de posición

Ajustar un modelo a una serie de datos consiste en estimar los parámetros de la distribución. Un método para realizarlo en una muestra de longitud N es el método de máxima verosimilitud:

$$\mu_y = \frac{1}{N} \sum_{i=1}^N \ln(x_i - x_0) \quad (8)$$

$$\sigma_y = \left[\frac{1}{N} \sum_{i=1}^N [\ln(x_i - x_0) - \mu_y]^2 \right]^{\frac{1}{2}} \quad (9)$$

El parámetro de posición se calcula mediante un proceso iterativo de modo que el coeficiente de asimetría tienda a cero, es decir que indique que el histograma de la variable es simétrico.

B.2. Modelo probabilístico de Doble Gumbel

Fisher y Tippet encontraron que la distribución del valor máximo (o mínimo) seleccionado de muestras de tamaño n se aproxima a una distribución límite cuando el tamaño de la muestra se aumenta. Cuando las distribuciones iniciales representan colas que pueden aproximarse por funciones de tipo exponencial, se encuentra que los valores extremos tienden a tener distribución tipo I o Gumbel [6].

Fariás, Dos Santos y Palacios explican que mientras la función de valores extremos tipo I o Gumbel está pensada para una población, la distribución de Doble Gumbel permite modelar el comportamiento de dos poblaciones consideradas mutuamente excluyentes que intercalan su ocurrencia bajo una misma variable climatológica [7]. Es el caso de las precipitaciones en la costa norte de Perú, donde se presentan años normales y años con presencia de El Niño, con magnitudes totalmente diferentes a las observadas el resto de años.

En la referencia [8], se plantea una función de distribución de probabilidad para dos poblaciones:

$$F(x) = (1-p)F(x; \theta_1) + p.F(x; \theta_2) \quad (10)$$

Donde:

p : Probabilidad de que la variable x pertenezca a la segunda población, que agrupa a elementos sucedidos en condiciones climatológicas ordinarias.

Función de acumulada de Gumbel

$$F(x) = \exp \left(- \exp \left(- \frac{(x - \alpha)}{\beta} \right) \right) \quad (11)$$

Donde:

x : variable hidrológica

α : Parámetro de posición de la función

β : Parámetro de escala de la función

Estos parámetros se pueden estimar conociendo que la media y la desviación estándar de la función son respectivamente:

$$\mu = \alpha + \gamma\beta \quad (12)$$

$$\sigma = \beta \frac{\pi}{\sqrt{6}} \quad (13)$$

Siendo γ la constante de Euler-Mascheroni ($\gamma \approx 0.577216$).

De modo que si se tuviera una muestra de tamaño infinito se podrían estimar los parámetros de la función, α y β , a través de los estadígrafos de la muestra, con las expresiones (14) y (15), [7]:

$$\beta = s_x \frac{\sqrt{6}}{\pi} \quad (14)$$

$$\alpha = \bar{x} + \gamma\beta \quad (15)$$

Donde \bar{x} y s_x son la media y la desviación estándar muestrales, respectivamente.

Sin embargo, en la práctica se trabaja con muestras de tamaño finito, por tanto, el valor de los parámetros debe modificarse. Gumbel obtuvo valores modificados minimizando la sumatoria de cuadrados de los errores perpendiculares a la recta de ajuste de valores extremos. Las ecuaciones que obtuvo están en función del tamaño de la muestra, n , y de los estadígrafos [3]:

$$\beta = \frac{s_x}{s_n} \quad (16)$$

$$\alpha = \bar{x} + y_n \cdot \beta \quad (17)$$

Donde:

y_n : Valor esperado de la variable reducida y_m

s_n : Desviación estándar de la variable reducida y_m

Estos estadígrafos pueden obtenerse de la variable reducida ordenando la muestra de manera decreciente y calculando para cada x_i la correspondiente variable reducida y_m con la expresión (18) de Weibull:

$$y_m = -\ln \left[-\ln \left(\frac{n+1-m}{n+1} \right) \right] \quad (18)$$

Donde:

n : Tamaño de la muestra

m : Número de orden

En la referencia [3] se recomienda trabajar con tablas que entregan el valor de y_n y s_n directamente en función del tamaño de la muestra n , ya que éstos no dependen de los valores muestrales. Se propone el uso de las expresiones para evitar el uso de las tablas y facilitar el cálculo (19) y (20).

$$y_n = 0.0252 \ln(x) + 0.4477 \quad (19)$$

$$s_n = 0.0993 \ln(x) + 0.7635 \quad (20)$$

Función de distribución Doble Gumbel

Con base en la ecuación (10), la probabilidad de excedencia de Gumbel de dos poblaciones se expresa mediante la expresión (21), [7]:

$$F(X) = (1-p)e^{-e^{-\frac{(x-\alpha_1)}{\beta_1}}} + p.e^{-e^{-\frac{(x-\alpha_2)}{\beta_2}}} \quad (21)$$

Esta función de distribución cuenta con cinco parámetros, dos de cada población, α_1 , β_1 , α_2 y β_2 , mientras el quinto es la proporción de mezcla, p .

C. Pruebas de bondad de ajuste

Las pruebas de bondad de ajuste consisten en comprobar gráfica y estadísticamente si la frecuencia empírica de la serie analizada, se ajusta a una determinada

función de probabilidades teórica seleccionada a priori, con los parámetros estimados con base en los valores muestrales [5].

Las pruebas estadísticas tienen por objeto medir la certidumbre que se obtiene al hacer una hipótesis estadística sobre una población, es decir calificar el hecho de suponer que una variable aleatoria, se distribuya según una cierta función de probabilidad [5].

De uso frecuente son la prueba χ^2 , la de Kolmogorov-Smirnov y la del error cuadrático estándar. La primera trabaja con valores agrupados, por lo que su aplicación no es muy práctica; la segunda no requiere agrupar valores, se concentra en determinar el punto en el que estas dos distribuciones muestran la mayor divergencia, pero no evalúa las diferencias en su conjunto; mientras que la tercera es una mezcla de ambas pruebas, trabajando con los datos independientes y sumando las divergencias parciales entre las probabilidades de excedencia, teórica y empírica [7].

En este estudio se aplicó el test de Kolmogorov-Smirnov como medida de bondad del ajuste.

El test de Kolmogorov-Smirnov es una prueba no paramétrica que permite probar si dos muestras provienen del mismo modelo probabilístico. Este test consiste en comparar las diferencias absolutas entre dos funciones de distribución acumulada, la empírica y la teórica [3].

Por lo tanto, es necesario definir tanto la frecuencia observada o empírica acumulada, como la frecuencia teórica acumulada. En el primer caso, para determinar la frecuencia observada acumulada de la serie, se ordenan los datos de menor a mayor y se trabaja con la probabilidad de excedencia asignada a cada valor de la muestra, calculada con la fórmula de Weibull (22), como se cita en la referencia [9]:

$$P_{emp}(x) = \frac{m}{n+1} \quad (22)$$

Donde:

$P_{emp}(x)$: frecuencia empírica u observada acumulada.

m : número de orden

n : número total de datos

En el segundo caso, la frecuencia teórica acumulada, $F(x)_i$, se determina a través de la función probabilística teórica.

Cuando han sido obtenidas ambas frecuencias para cada evento de la serie analizada, se selecciona el valor máximo el valor absoluto, de la diferencia entre el valor observado y el valor de la recta teórica del modelo [5], es decir:

$$\Delta_{max} = \max |P_{emp}(x)_i - F(x)_i| \quad (23)$$

Una vez obtenido el valor del estadígrafo Δ_{max} , se compara con el $\Delta_{crítico}$, proveniente de tablas de valores críticos de esta prueba de bondad de ajuste, con un tamaño de muestra n y asumiendo un 95% de confianza o 5% de significancia. Luego se decide si el ajuste es o no adecuado, considerando las siguientes afirmaciones:

- Si $\Delta_{max} > \Delta_{crítico}$ se rechaza H_0 (el ajuste no es adecuado).
- Si $\Delta_{max} < \Delta_{crítico}$ no se rechaza H_0 (el ajuste es adecuado).

D. Determinación matemática de las curvas IDF

Las curvas IDF no sólo se construyen de manera gráfica, sino que también pueden expresarse como ecuaciones con el fin de evitar la imprecisión de lectura de la intensidad de lluvia de diseño en una gráfica [10].

El planteamiento de las primeras relaciones matemáticas, que modelan a dichas curvas, data desde el inicio de los años treinta del siglo pasado, y fueron realizadas por Sherman, en 1931 y Bernard, en 1932 [11]. Estos modelos han conducido a numerosos investigadores del mundo a generar diversas expresiones matemáticas que simulan las curvas IDF [12], algunas de las cuales se presentan a continuación:

Modelo propuesto por Bernard en [11]:
$$i_t^T = \frac{kT^m}{t^n} \quad (24)$$

Modelo propuesto por Chow [10]:
$$i_t^T = \frac{kT^m}{t^n + b} \quad (25)$$

Modelo propuesto por Koutsoyiannis et al. en [13]:
$$i_t^T = k \left\{ \frac{m - \ln[-\ln(1 - 1/T)]}{(t+b)^n} \right\} \quad (26)$$

Modelo propuesto por ASCE en [14]:
$$i_t^T = \frac{A}{t^B + C} \quad (27)$$

Modelo propuesto por ASCE en [14]:
$$i_t^T = \frac{a}{(t+b)^c} \quad (28)$$

Donde T es período de retorno, t es la duración e i_t^T la intensidad máxima de precipitación y A, B, C, a, b, c, k, m y n son parámetros adimensionales.

Los estudios para relacionar las lluvias de duraciones menores a un día, expresadas en valores horarios, con las lluvias diarias registradas, incluyen generalmente el periodo de retorno y el tiempo como variables [15].

Chen en [16], usando series anuales, propuso la siguiente fórmula para expresar la intensidad de una lluvia:

$$i_t^T = \frac{a \cdot i_t^{10} \cdot \log \left\{ 10^{2-x} \cdot [\ln(T/(T-1))] \right\}^{-(x-1)}}{(t+b)^c} \quad (29)$$

La relación es válida para periodos de retorno mayores a un año y para duraciones entre 5 minutos y 24 horas. Los parámetros a, b y c se obtienen de un gráfico presentado por dicho autor en función de la razón entre la lluvia de una hora y de 24 horas para el mismo periodo de retorno (K). El exponente x es la razón entre las lluvias de 1 hora de 10 y 100 años de periodo de retorno.

Varas y Farías [15] analizaron distintos trabajos regionales encontrados acerca de las relaciones IDF, los cuales se unificaron de acuerdo a la ecuación de Chen para series anuales, obteniendo una síntesis de las ecuaciones para determinar los parámetros regionales a, b y c empleados en dicha expresión según la región de aplicación, sobre la base del parámetro k. Dicha combinación de parámetros entregó valores parecidos entre los diferentes trabajos en el rango de duración de la lluvia entre 1 y 24 horas.

III. RECOPIACIÓN DE INFORMACIÓN HISTÓRICA

Se contó con información histórica de lluvias de 45 años entre las estaciones meteorológicas Miraflores y Universidad de Piura (UDEP), por ser las más representativas de la zona. Ambas se ubican en la ciudad de Piura, en la provincia y departamento del mismo nombre, a poco más de 2 km de distancia. La Tabla I presenta los detalles de ubicación de las estaciones.

TABLA I
UBICACIÓN DE LAS ESTACIONES PLUVIOGRÁFICAS DE ESTUDIO

Estación	Longitud W	Latitud S	Distrito	Categoría
Miraflores	80°36'59"	05°10'31"	Castilla	PLU
Udep	80°38'18"	05°10'14"	Piura	PLU-PG

La estación Miraflores empezó su operación en el año 1972 contando con pluviógrafo y pluviómetro, a cargo del Proyecto Especial Chira Piura (PECHP), hasta el año 1992; y en 1993 pasó a cargo del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI), hasta la actualidad, en que cuenta con un pluviómetro convencional.

La información histórica de lluvias de la estación Miraflores fue brindada por el departamento de Hidrometeorología del PECHP, consistente en las precipitaciones diarias máximas anuales de los años 1971 al 2017, intensidades máximas para diversas duraciones y años, y bandas pluviográficas de los años 1984, 1987, 1988, 1989 y 1990.

La estación meteorológica UDEP se encuentra en el Campus de la Universidad de Piura, a cargo del Laboratorio de Física General y empezó su operación en febrero de 1991. Esta estación meteorológica cuenta con un pluviógrafo basculante Campbell Scientific; modelo TE525 con precisión de 0.25 mm.

El Instituto de Hidráulica, Hidrología e Ingeniería Sanitaria de la Universidad de Piura (IHHS) proporcionó la base de datos pluviográficos de la estación meteorológica UDEP, años 1995-1998 y 2005-2017, consistente en una base de datos de lluvia en intervalos de 30 minutos, salvo los datos de 2017 que están configurados cada 10 minutos.

IV. PROCESAMIENTO DE LA INFORMACIÓN

A. Análisis de tormentas

Obtenida toda la información histórica de lluvias posible, se procedió a realizar el análisis de tormentas respectivo con la finalidad de obtener las intensidades máximas de cada año hidrológico (01 septiembre-31 de agosto) y así actualizar la base de datos de intensidades máximas). Se describe a continuación el procesamiento de la información recogida en cada estación meteorológica.

La base de datos de lluvias cada media hora de la estación Udep se procesó en hoja de cálculo Excel, determinando las intensidades de precipitación para periodos de duración de 30, 60, 120, 180, 240, 360, 720 y 1440 minutos a partir todos los datos. Se obtuvieron las intensidades máximas anuales para cada duración.

De las bandas pluviográficas de las máximas tormentas de la estación Miraflores se realizó el correspondiente análisis, a fin de extraer la información numérica a hojas de cálculo, y de esta manera obtener las intensidades máximas

para cada duración, tal como se realizó con los registros históricos de la estación UDEP, y para cada año hidrológico.

Como resultado del análisis de las dos estaciones se actualizó la matriz de precipitaciones máximas diarias e intensidades, obteniendo una versión preliminar, que contiene información de los años 1972 a 2017; con excepción del año 1993, del cual no se encontraron registros en ninguna de las dos estaciones. Esta matriz está completa en un 73% y se requiere completarla para elaborar las curvas IDF, sobre todo porque de algunos años solo se dispone de registros de precipitaciones máximas diarias.

B. Relleno de datos faltantes de intensidades

El primer paso fue completar las intensidades en los años que presentaban menor cantidad de datos faltantes. Así, se inició con el año hidrológico 2016-2017 determinando la intensidad de 15 minutos de duración, i_{15} , interpolando las intensidades adyacentes, i_{10} e i_{30} del mismo año, y luego la i_{240} en diversos años faltantes, interpolando del mismo modo los valores anuales correspondientes a la i_{180} e i_{360} (Fig. 2).

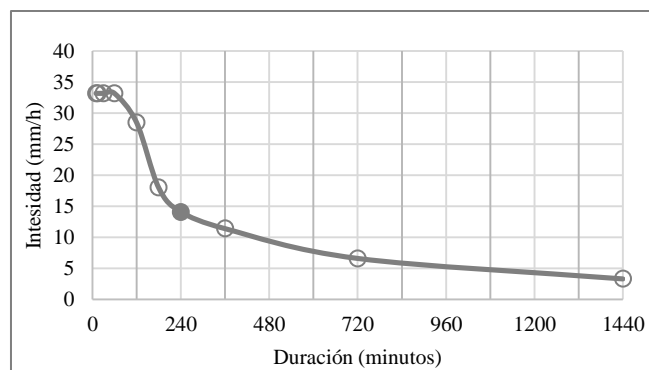


Fig. 2 Ejemplo de determinación de intensidad (i_{240}) por interpolación.

A continuación se completó la serie i_{60} , por presentar pocos datos faltantes y ser representativa del comportamiento pluviográfico, correlacionándola con la Pmax día (Fig. 3).

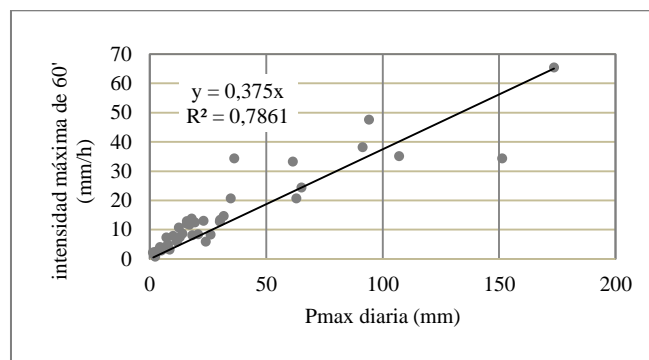


Fig. 3. Correlación de la serie i_{60} y Pmax día.

Finalmente, se evaluó la correlación de las series de intensidades restantes respecto a las series completas, Pmax día e i_{60} , efectuando el relleno con aquella que presentara mejor correlación (Fig. 4). Esto permitió completar todos los datos faltantes y obtener la matriz de intensidades máximas anuales para el período 1972-2017, exceptuando el año 1993 sin registros.

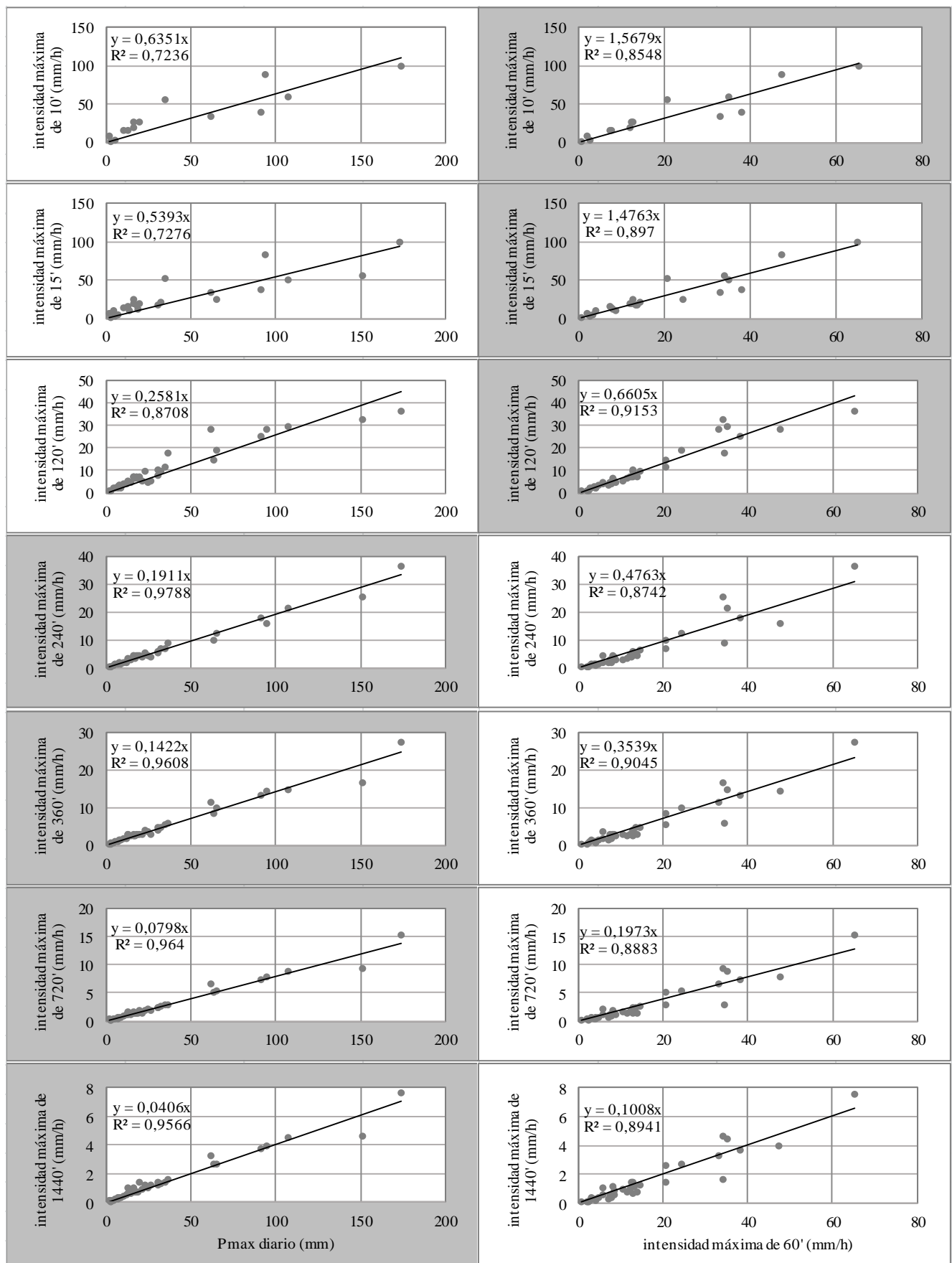


Fig. 4 Evaluación de la correlación de las diferentes series de intensidades máximas con la serie de Pmax día (izquierda) y con i60 (derecha) para el relleno de datos faltantes. Los recuadros sombreados indican las mejores correlaciones que se han usado para el relleno.

C. Definición de la base de datos

La Región Piura está ubicada en la zona árida del norte peruano, presenta años muy secos y algunos años de considerable cantidad de lluvia, producto de la presencia de Fenómenos El Niño (FEN), ya sea de carácter global (ENSO) o costero. La Fig. 5 muestra la variabilidad de las precipitaciones diarias máximas anuales para el período 1972-2017, donde se aprecia que a partir de 1983 el régimen hídrico de la zona podría haber aumentado.

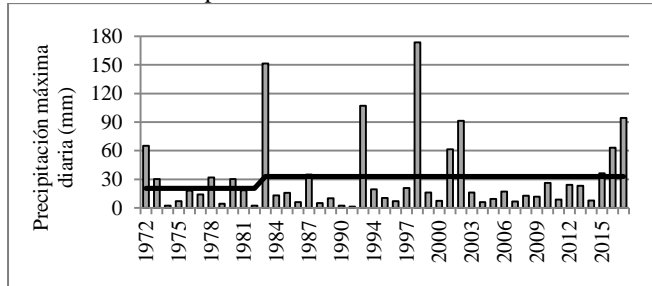


Fig. 5 Variabilidad de las precipitaciones máximas diarias en Piura y su media.

Existen evidencias de que los caudales máximos del río Piura (Fig. 6) presentan este mismo cambio en la serie aún más larga, entre 1926-2008 C. Su trabajo analiza estadísticamente esta hipótesis y propone una división de la serie hidrométrica, basada en que, en los últimos años, la intensidad de los FEN y sus respectivos caudales han incrementado, dejando prever que posiblemente la serie ha cambiado y por lo tanto las series 1926-1982 y 1983-2008 pueden no pertenecer al mismo universo.

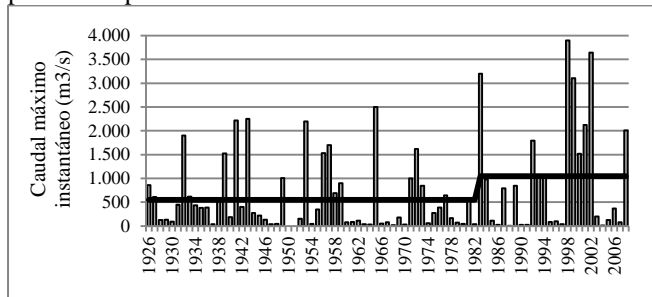


Fig. 6 Variabilidad de los caudales máximos anuales en el río Piura.

Basados en los hallazgos antes mencionados, se dividió la matriz de series de intensidades máximas anuales en dos grupos de series para analizar si todas las series pertenecen o no a un mismo universo; y decidir en base a ello qué series se propondría para la construcción final de las curvas IDF.

Al dividir la base de datos y analizar dichos grupos de series de datos, se observa que estas series tienen parámetros estadísticos diferentes. Sin embargo, y de acuerdo con las conclusiones de varios otros estudios a nivel mundial, se considera que estas diferencias no pueden ser directamente consideradas como la confirmación de cambios en las tendencias y parámetros principales de una cuenca; pero se puede analizar si existe diferencia en las medias y desviaciones estándar de las divisiones de las series mediante pruebas de hipótesis [17]

Para realizar el análisis se escogieron tres variables representativas de toda la matriz: Pmax día, i_{10} e i_{60} . Los datos de cada variable fueron divididos en dos series, A

(1972-1982) y B (1983-2017), de acuerdo a lo propuesto en [17].

C.1. Estadística descriptiva de las series

Para verificar y comprobar la calidad y confiabilidad de las series antes mencionadas, primero se evaluaron los principales estadígrafos para apreciar en qué medida los datos se agrupan o dispersan en torno a ese valor central (Tabla II y Tabla III). Los estadígrafos se aprecian muy diferentes entre las series. Las series B (1983-2017) superan notablemente los valores en diferentes parámetros de las series completa y A; por tal motivo se decidió continuar con las pruebas de hipótesis necesarias para comprobar si realmente difieren dichos parámetros de cada una de las series.

TABLA II
ESTADÍSTICA DESCRIPTIVA DE LA PMAX DIA

Parámetros	Serie total	Serie A	Serie B
Media	29.8	20.4	32.9
Mediana	16.0	18.1	16.0
Desviación estándar	38.4	18.7	42.7
Varianza	1474	348	1821
Curtosis	5.4	2.5	3.9
Coefficiente de asimetría	2.3	1.4	2.1
Rango	172.1	63.1	172.1
Mínimo	1.5	2.2	1.5
Máximo	173.6	65.3	173.6
Cuenta	45.0	11.0	34.0

TABLA III
ESTADÍSTICA DESCRIPTIVA DE LA i_{60}

Parámetros	Serie total	Serie A	Serie B
Media	13.2	9.6	14.4
Mediana	8.2	8.6	8.0
Desviación estándar	13.9	7.0	15.4
Varianza	193	48	237
Curtosis	3.8	0.5	2.6
Coefficiente de asimetría	1.9	0.7	1.7
Rango	64.6	23.5	64.6
Mínimo	0.7	0.8	0.7
Máximo	65.3	24.3	65.3
Cuenta	45.0	11.0	34.0

C.2. Aplicación de las pruebas de hipótesis a las series

Para la aplicación de la prueba de hipótesis de comparación de medias a las series elegidas se planteó como hipótesis nula: “El promedio de las series B es igual al promedio de las series A” y como hipótesis alternativa, que existe diferencia significativa entre los promedios.

Se prosiguió evaluando el tamaño de las muestras, menores que 50 en ambos casos, 11 y 34 años respectivamente. Por lo explicado, se aplicó la prueba de t-Student, obteniendo los resultados de la Tabla IV y la Tabla V para las variables Pmax día e i_{60} , respectivamente.

TABLA IV
PRUEBA T-STUDENT APLICADA A LAS SERIES DE PMAX DIA

Parámetros	Serie A	Serie B
Media	20.4	32.9
Varianza	347.9	1820.8
Observaciones	11.0	34.0
Varianza agrupada	1478.3	
Diferencia hipotética de las medias	0	
Grados de libertad	43.0	
Estadístico t	0.938	
P(T<=t) una cola	0.180	

Con los resultados de los estadígrafos t , se obtuvieron las probabilidades de excedencia de 0.18 y 0.16, respectivamente, lo que representa que, con esos niveles de significancia o mayores, 20% por ejemplo, se puede admitir que las medias de los dos grupos de series estudiadas pertenecen a universos diferentes.

TABLA V

PRUEBA T-STUDENT APLICADA A LAS SERIES DE LA VARIABLE I_{60}

Parámetros	Serie A	Serie B
Media	9.6	14.4
Varianza	48.4	237.0
Observaciones	11	34
Varianza agrupada	193.1	
Diferencia hipotética de las medias	0	
Grados de libertad	43	
Estadístico t	0.992	
$P(T \leq t)$ una cola	0.163	

Por otro lado, en C.1 se puede observar que los valores de las varianzas de las series B son notablemente mayores a los de las series A. Esta evidencia no es suficiente para rechazar la hipótesis de que las muestras tienen varianzas iguales, por lo que se aplicó la prueba de hipótesis F de comparación de dos varianzas, para lo cual se estableció como hipótesis nula que las muestras provienen de universos con varianzas iguales y como hipótesis alternativa que la varianza de la serie B es mayor a la varianza de las series A.

Los valores F calculados a partir de las muestras son mayores al F crítico, entonces podemos rechazar la hipótesis nula (Tabla VI y Tabla VII). Por lo tanto, se puede considerar aceptar la hipótesis de que las muestras no provienen de universos con la misma varianza, con un nivel de significancia del 5%.

TABLA VI

PRUEBA F APLICADA A LAS SERIES DE LA VARIABLE P_{MAX} DÍA

Parámetros	Serie A	Serie B
Media	32.91	20.40
Varianza	1820.8	347.9
Observaciones	34	11
Grados de libertad	33	10
F calculado	5.23	
$P(F \leq f)$ una cola	0.00427	
Valor crítico para F (una cola: 5%)	2.69	

TABLA VII

PRUEBA F APLICADA A LAS SERIES DE LA VARIABLE I_{10}

Parámetros	Serie C	Serie B
Media	14.42	9.64
Varianza	236.95	48.40
Observaciones	34	11
Grados de libertad	33	10
F calculado	4.90	
$P(F \leq f)$ una cola	0.00558	
Valor crítico para F (una cola: 5%)	2.69	

C.3. Base de datos final

Basados en el análisis de la estadística descriptiva, las pruebas de hipótesis y el trabajo realizado en [17], se concluye que las series A y B provienen de universos diferentes, esto es que se está ante un cambio en las características climatológicas de las lluvias en la ciudad de

Piura, que debe ser considerado para determinar las lluvias de diseño de la infraestructura.

A la vista de los resultados de la aplicación de las pruebas de medias y varianzas, se decidió trabajar únicamente con las series B (1983-2017) para el cálculo.

D. Modelación probabilística

De los diversos modelos ajustados a las series de intensidades, los que presentaron mejor ajuste fueron el modelo Log Normal de tres parámetros y el modelo Doble Gumbel.

D.1. Ajuste del modelo Log Normal 3p a las series

En la Tabla VIII se presentan los parámetros determinados mediante el método de máxima verosimilitud.

TABLA VIII

PARÁMETROS DE LA DISTRIBUCIÓN LOG-NORMAL 3P AJUSTADA A CADA SERIE DE INTENSIDADES Y P_{MAX} DÍA

Parámetro	P_{mx} día	i10	i15	i30	i60	i120	i180	i360	i720	i1440
μ_y	2.78	2.76	2.66	2.44	2.16	1.63	1.37	0.00	0.00	0.00
σ_y	1.22	0.00	1.02	1.07	1.06	1.15	1.17	1.15	1.17	1.21
x_o	0.73	-0.38	-0.23	0.20	-0.03	0.34	0.25	0.06	0.03	0.04

D.2. Ajuste del modelo Doble Gumbel a las series

El ajuste de las series a modelos Doble Gumbel se inicia con la selección del parámetro p , a partir de la observación de los datos, considerando que agrupa a los eventos climatológicamente ordinarios, esto es, sin evidencia de presencia de El Niño. Luego de definidos todos los parámetros y evaluado el ajuste logrado, se podrá variar el valor del parámetro p para intentar alcanzar un mejor ajuste. La evaluación hecha arrojó que la mejor división de las series de estudio fue con 24 años normales y 10 años extraordinarios o con presencia de El Niño. Una vez divididas las series la determinación de los parámetros restantes se realiza de manera independiente en cada una como se explicó en II.B.2 (Tabla IX).

TABLA IX

PARÁMETROS DE LA DISTRIBUCIÓN DOBLE GUMBEL AJUSTADA A CADA SERIE DE INTENSIDADES Y P_{MAX} DÍA

Serie	P_{max} día	i10	i15	i30	i60	i120	i180	i360	i720	i1440
n1	10									
m1	84.0	54.5	51.7	45.9	34.2	23.4	18.8	12.1	6.8	3.5
s1	49.7	24.0	23.6	23.9	14.8	9.3	9.7	6.9	3.9	1.9
yn1	0.495									
Sn1	0.950									
a1	52.3	25.2	24.9	25.2	15.6	9.8	10.2	7.3	4.1	2.0
u1	58.0	42.0	39.4	33.4	26.4	18.6	13.7	8.5	4.8	2.5
n2	24									
m2	11.6	11.5	10.5	8.7	6.2	3.8	2.9	1.8	1.0	0.6
s2	6.3	7.1	6.6	5.3	3.7	2.0	1.5	0.9	0.5	0.3
yn2	0.530									
Sn2	1.086									
a2	5.8	6.6	6.1	4.9	3.4	1.8	1.4	0.9	0.5	0.3
u2	8.6	8.0	7.3	6.1	4.4	2.9	2.2	1.3	0.7	0.4

D.3. Bondad de ajuste de los modelos ajustados

Se aplicó el test de bondad de ajuste Kolmogorov-Smirnov para los ajustes realizados con los modelos probabilístico Log-Normal 3p y Doble Gumbel (Tabla X).

TABLA X
VALORES DE DELTA MÁXIMO PARA LA DETERMINACIÓN DE LA BONDAD DE AJUSTE CON EL TEST KOLMOGOROV – SMIRNOV

Delta máximo	Pmax día	i10	i30	i60	i120	i360	i720	i1440
LN 3	0.070	0.067	0.074	0.098	0.087	0.108	0.102	0.083
Doble Gumbel	0.079	0.085	0.091	0.059	0.057	0.105	0.084	0.080

Con lo observado en la Tabla X, se puede afirmar que los ajustes de las series a los modelos probabilísticos Log-Normal 3p y Doble-Gumbel son adecuados, pues en todos los casos se obtienen valores de delta máximo, Δ_{\max} , menores que el delta crítico, $\Delta_{\text{crítico}}$, con 5% de significancia, que para el tamaño de la muestra trabajada, corresponde a 0.233.

E. Curvas IDF

E.1. Determinación de la matriz de intensidades definitiva

Al observar que ambos tipos de modelo arrojan valores relativamente semejantes, se decidió promediar sus predicciones para formar una única matriz de intensidades (Tabla XI), que permitió graficar las curvas IDF requeridas (Fig. 7).

TABLA XI
MATRIZ DE PRECIPITACIONES MÁXIMAS (MM) E INTENSIDADES MÁXIMAS ANUALES (MM/H) EN FUNCIÓN DE LA DURACIÓN (MIN) Y EL PERÍODO DE RETORNO T (AÑOS)

T (años)	Pmax día (mm)	i10	i15	i60	i120	i360	i720	i1440
5	48.5	37.8	35.3	22.8	15.5	7.2	4.0	2.1
10	90.6	59.6	56.7	36.8	24.9	13.1	7.3	3.8
25	147.5	88.9	85.7	55.8	37.9	20.9	11.8	6.1
50	196.9	113.3	110.0	71.9	49.4	27.6	15.6	8.1
100	254.1	140.3	137.2	90.2	62.8	35.2	19.9	10.5
150	291.9	157.6	154.6	102.0	71.6	40.2	22.7	12.0
200	321.0	170.5	167.8	110.9	78.4	44.0	24.9	13.2

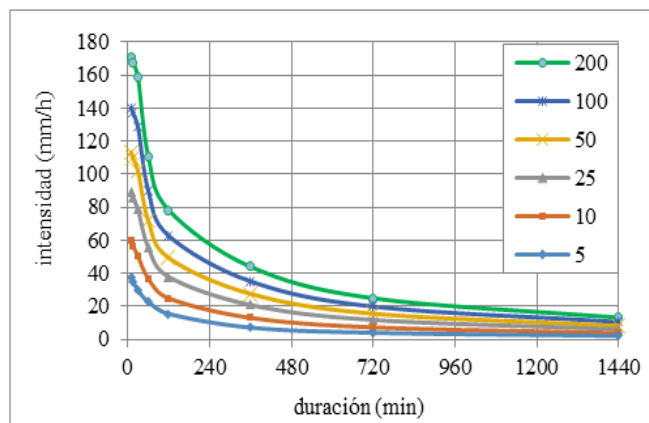


Fig. 7 Curvas IDF para la ciudad de Piura (1983-2017).

E.2. Determinación de la expresión matemática

Basados en [15] se propone, a partir de la matriz de intensidades obtenidas para la zona de estudio (Tabla XI), determinar una sola expresión para toda la familia de curvas IDF de la forma:

$$i_t^T = \frac{\alpha(T)}{(t+b)^c} \quad (30)$$

Para ello, se trabajó preliminarmente con la expresión (28), hallando los parámetros a , b y c , es decir trabajando por separado cada periodo de retorno. En este análisis se obtuvo una buena correlación (Tabla XII).

TABLA XII
VALORES DE LOS PARÁMETROS Y EL COEFICIENTE DE CORRELACIÓN PARA CADA PERÍODO DE RETORNO

Tr	a	b	c	R ²
5	1.933	58,82	0,930	1,000
10	3.274	67,90	0,919	0,998
25	5.131	73,47	0,915	0,997
50	6.890	77,59	0,916	0,996
100	9.077	81,97	0,919	0,996
150	10.605	84,68	0,921	0,996
200	11.820	86,68	0,923	0,996

Se puede observar que los valores de los parámetros b y c no muestran mayores diferencias, por lo cual se puede establecer un valor único para dichos parámetros. Por tal motivo, se decidió encontrar las expresiones apropiadas para el numerador, N , y el denominador, D , de la expresión (30).

En primer lugar se procedió a asignar un valor común a b , tomando el promedio de los valores hallados para cada periodo de retorno. Así, se llegó a establecer $b = 75$, y con este valor único de b se recalcularon los valores de c para cada periodo de retorno. Con el mismo criterio, se estableció un único valor de c con el promedio de los nuevos valores, obteniendo como resultado $c = 0.912$. Con los valores únicos establecidos de $b = 75$ y $c = 0.912$ para todos los periodos de retorno, el denominador de la ecuación (30) quedó expresado como:

$$D = (t+b)^c = (t+75)^{0.912} \quad (31)$$

Seguidamente se prosiguió a establecer la expresión para el numerador de la ecuación (30), N , en función del periodo de retorno, T . Esto se logra estimando valores de N como el producto de las intensidades de la Tabla XI i_{matriz} , por el denominador calculado con la expresión (31) y ajustando una expresión matemática a dichos valores:

$$N = i_{\text{matriz}} * D \quad (32)$$

La Fig. 8 grafica el periodo de retorno, T , y el numerador, N , para encontrar la expresión matemática buscada para el numerador de la expresión (30).

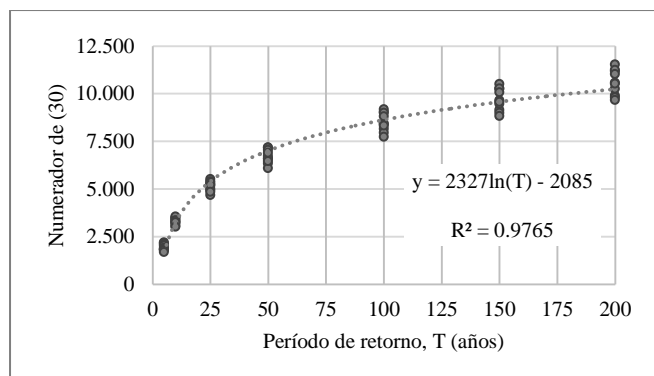


Fig. 8 Ajuste del numerador en función del periodo de retorno.

Finalmente, la ecuación queda expresada de la siguiente manera:

$$i_t^T = \frac{2327 \ln(T) - 2085}{(t + 75)^{0.912}} \quad (33)$$

Para dar validez a la ecuación (33) se correlacionaron las intensidades de la matriz (i_{matriz}) y las brindadas por la ecuación ajustada, dando como resultado un coeficiente de correlación bastante cercano a la unidad (Fig. 9), con lo que se da por apropiada la expresión obtenida de la intensidad de lluvia en función del período de retorno y la duración.

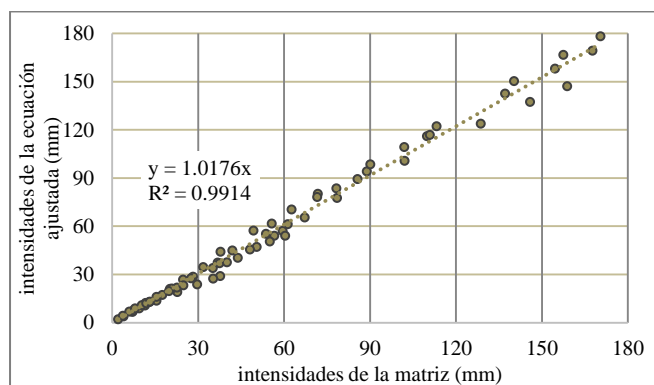


Fig. 9 Correlación entre las intensidades de la matriz y las obtenidas por la ecuación ajustada.

CONCLUSIONES

Las curvas IDF son una herramienta fundamental para la modelación hidrológica orientada a determinar los caudales de diseño de obras ubicadas en lugares sin registros hidrométricos que permitan modelar probabilísticamente los caudales históricos. Usualmente se determinan a partir de la modelación de las tormentas históricas, asumiendo que el clima tiene la misma tendencia a lo largo del tiempo, es decir que es estacionario, lo cual no es correcto ante las evidencias de cambio climático que se registran en diversas partes del mundo.

En el Perú, la presencia recurrente de eventos El Niño, costero y global, genera que en los últimos años, a partir de El Niño de 1983, los registros de precipitaciones de la ciudad de Piura, indiquen un cambio en el clima, con años lluviosos más frecuentes e intensos, contando con tres eventos extraordinarios, 1983, 1998 y 2017 y otros de moderada intensidad. Este cambio se ha verificado estadísticamente al aplicar pruebas de hipótesis de medias y

de varianzas. La serie histórica de 1972 a 2017 fue dividida en dos: 1972-1982 y 1983-2017, encontrando que estadísticamente podrían provenir de universos diferentes, esto quiere decir que la pluviometría de ambos períodos es diferente.

Se propone actualizar las curvas IDF existentes para incluir los últimos eventos y evaluar recortar apropiadamente las series históricas para trabajar con únicamente con datos más recientes, que reflejen la tendencia climática no estacionaria actual.

A las intensidades definidas para la ciudad de Piura, como caso de estudio, se les ajustó modelos Log-Normal 3p y Doble Gumbel, cuyas predicciones se han promediado para obtener las curvas actualizadas y una ecuación general que relaciona las tres variables de diseño para mayor practicidad.

REFERENCIAS

- [1] S. Fadhel, M. Rico-Ramirez y D. Han., "Uncertainty of Intensity–Duration–Frequency (IDF) curves due to varied climate baseline periods", *Journal of Hydrology*, vol. 547, pp. 600-612, Abril 2017.
- [2] G. Mirhosseini, P. Srivastava y L. Stefanova (2013). *Reg Environ Change* 13 (Suppl 1):25. <https://doi.org/10.1007/s10113-012-0375-5>
- [3] E. Varas, y P. Bois, "*Hidrología Probabilística*". Chile: Ediciones Universidad Católica de Chile, 1998.
- [4] P. Acosta, y L. Sierra, "Evaluación de métodos de construcción de curvas IDF a partir de distribuciones de probabilidad y parámetros de ajuste", *Revista Facultad de Ingeniería, UPTC*, Vol. 22, no 35, pp.25-33, Jul-Dic 2013.
- [5] M. Villón, "*Hidrología Estadística*". Lima: Villón, 2005.
- [6] R. Linsley, M. Khohler y J. Paulhus, "*Hidrología para ingenieros*". D. F. México: Programas Educativos. 1992
- [7] M. Farias, E. Dos Santos y C. Palacios S., "Estimación de Caudales de Diseño con la Distribución Doble Gumbel", en XIX Simposio Brasileiro de Recursos Hídricos, Maceió, Brasil, pp.1-10, Noviembre – Diciembre 2011.
- [8] J. Raynal y M. Raynal, "Cálculo de los límites de confianza de la Distribución de Probabilidad de valores extremos tipo I para dos poblaciones". *Información Tecnológica*, vol. 15, no 1, pp.87-94, 2004.
- [9] R. Pizarro, D. Aravena, K. Macaya, A. Abarza, M. Cornejo, M. Labra, M. Pavez y L. Román. "Curvas Intensidad – Duración – Frecuencia, para la zona centro sur de Chile". UNESCO, Editorial Universidad de Talca, Chile, Documentos Técnicos del PHI-LAC No7, 2007.
- [10] V. Chow, L. Mays y D. Maidment, "*Hidrología aplicada*". Santa Fe de Bogota, Colombia: Mc Graw-Hill. 1994
- [11] D. Pereyra, J. Pérez y L. Gómez, "Ecuaciones que estiman las curvas intensidad-duración-período de retorno de la lluvia", *GEOS*, vol. 24, no 1, pp.46-45, Agosto 2004.
- [12] G. Remenieras, "*Tratado de Hidrología Aplicada*". Madrid. España: Editores Asociados, 1971.
- [13] D. Koutsoyiannis, D. Kozonis y A. Manetas, "A mathematical framework for studying rainfall intensity-duration-frequency relationships", *Journal of Hydrology*, vol. 206, pp. 118-135, 1998.
- [14] ASCE. "*Hydrology Handbook*". Nueva York, Estados Unidos de America: ASCE. 1996.
- [15] E. Varas y M. Farias, "Estudio Comparativo de Relaciones Intensidad-Duración-Frecuencia". En Memorias XIX Congreso Latinoamericano de Hidráulica, Córdoba, Argentina, Tomo II, pp. 359-366. 2000.
- [16] C. Chen, "Rainfall intensity-duration-frequency formulas", *Journal of Hydraulic Engineering*, ASCE, vol 109, no 12, pp. 1603-1621. 1983.
- [17] C. Palacios, "Caudales de diseño en el Río Piura y su variación histórica ante El Fenómeno El Niño", Tesis para obtener el título de Ingeniero Civil, Universidad de Piura, Perú, 2010.