Villón Bejar, Máximo.
HidroEsta, software para cálculos hidrológicos.
Tecnología en Marcha. *Vol. 18 N.° 2 Especial.*

HidroEsta, *software* para cálculos hidrológicos

Máximo Villón Bejar [1]

Palabras clave

Precipitación, caudales, variables aleatorias, distribuciones, parámetros estadísticos, media, varianza, hidrología.

Resumen

Este trabajo de investigación se orientó a la elaboración de una herramienta computacional bajo el título *HidroEsta,* software *para cálculos hidrológicos*, utilizando Visual Basic, el cual pretende ser una aplicación que permita facilitar y simplificar los cálculos laboriosos que se deben realizar en los estudios hidrológicos.

El *software* permite el cálculo de los parámetros estadísticos, cálculos de regresión lineal, no lineal, simple y múltiple, así como regresión polinomial, evaluar si una serie de datos se ajustan a una serie de distribuciones, calcular a partir de la curva de variación estacional o la curva de duración, eventos de diseño con determinada probabilidad de ocurrencia, realizar el análisis de una tormenta y calcular intensidades máximas, a partir de datos de pluviogramas, los cálculos de aforos realizados con correntómetros o molinetes, el cálculo de caudales máximos, con métodos empíricos y estadísticos, cálculos de la evapotranspiración y cálculo del balance hídrico.

En la investigación se probaron diferentes métodos numéricos para la solución de las ecuaciones, seleccionándose el más adecuado para cada situación.

El producto del trabajo proporciona al ingeniero civil, agrícola, agrónomo, hidrólogos y otros especialistas que trabajen en este campo, una herramienta que permite realizar cálculos, simulaciones rápidas, y determinar los caudales o precipitaciones de diseño.

Introducción

Los estudios hidrológicos requieren del análisis de cuantiosa información hidrometeorológica; esta información puede consistir en datos de precipitación, caudales, temperatura, evaporación, etc.

Los datos recopilados, solo representan una información en bruto, pero si estos se organizan y se analizan en forma adecuada, proporcionan al hidrólogo una herramienta de gran utilidad, que le

1 Ingeniero agrícola, especialista en ingeniería de recursos de aguas y suelos y computación. Catedrático, Escuela Ingeniería Agrícola. Instituto Tecnológico de Costa Rica. Cartago, Costa Rica. Teléfono: 550-2595. Correo electrónico: *mvillon@itcr.ac.cr.*

permite tomar decisiones en el diseño de estructuras hidráulicas (5).

Para realizar los cálculos, los hidrólogos tienen que afrontar una serie de problemas, debido a que:

- El procesamiento de la información que se tiene que realizar es bastante laboriosa.
- Las ecuaciones por solucionar, en la mayoría de los casos, son muy complejas, y para su solución se requiere del uso de métodos numéricos.
- Las simulaciones que se realizan manualmente consumen mucho tiempo, debido a los cálculos requeridos.

Por lo laborioso del proceso de la información y de los cálculos, se puede incurrir en errores, en razón de lo cual es necesario un *software* que brinde al hidrólogo una herramienta que le permita simplificar todos estos procesos, e inclusive posibilitarle simular sus resultados, facilitando con esto optimizar su diseño.

Objetivo

Elaborar un *software* que facilite y simplifique los cálculos laboriosos, y el proceso del análisis de la abundante información que se deben realizar en los estudios hidrológicos.

Fundamentos teóricos

Variables aleatorias

Una variable aleatoria (1) es una función X, definida sobre un espacio muestral S, que asigna un valor a esta variable, correspondiente a cada punto (o cada resultado) del espacio muestral de un experimento.

A una variable aleatoria se le conoce también como variable estocástica, sus valores son números reales, que no pueden predecirse con certeza antes de ocurrir el fenómeno; es decir, ocurren al azar.

Variable aleatoria discreta

Se dice que una variable aleatoria X es discreta cuando sus valores se restringen a un conjunto enumerable finito o infinito.

Ejemplo: Número de días de lluvias ocurridas en los meses de un año cualquiera.

Variable aleatoria continua

Se dice que una variable aleatoria X es continua, cuando sus valores se encuentran en un rango continuo y puede ser representado por cualquier número entero o decimal.

Ejemplo: El caudal diario registrado en una estación de aforo.

Distribuciones

El comportamiento de una variable aleatoria se describe mediante su ley de probabilidades, que, a su vez, se puede caracterizar de varias maneras. La más común es mediante la distribución de probabilidades de la variable aleatoria (4).

Notación:

X $\Rightarrow$ variable aleatoria de la función

x $\Rightarrow$ valor particular que toma la variable aleatoria

f(x) $\Rightarrow$ función de densidad de probabilidad (función de probabilidad, distribución de probabilidad de x)

F(x) $\Rightarrow$ función acumulada (función de distribución acumulada)

Medidas de las distribuciones

La media aritmética

Dada la muestra compuesta de *n* datos, $x_1, x_2, ..., x_n$, la media, se define como la suma algebraica de ellas, dividida entre el número de datos. Cuando se calcula la media para una población, esta se denota por μ, y cuando se trata de una muestra, por *x*.

Matemáticamente, la media de los datos no agrupados se representa por:

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

donde:

μ = media poblacional

$\bar{\chi}$ = media muestral

x_i = valor i-ésimo de la muestra

n = número de datos de la muestra o población

La mediana

Es un valor único de un conjunto de datos que mide el elemento central en los datos (5). Este único elemento de los datos ordenados es el más cercano a la mitad, o el más central en el conjunto de números. La mitad de los elementos quedan por encima de ese punto, y la otra mitad por debajo de él.

Sean x_1, x_2, x_3, ..., x_n datos ordenados por magnitud creciente o decreciente y *n* el número impar de datos, la mediana (Med) es el dato situado en el centro, es decir:

$$\text{Med} = X_{(n+1)/2} \text{ , para } n \text{ impar}$$

si *n* es par, la mediana es el promedio de los números centrales; es decir:

$$Med = \frac{x_{n/2} + x_{n/2+1}}{2} \text{ , para n par}$$

La moda

Es aquel valor que se repite más frecuentemente en un conjunto de datos, se denota por *Mo*.

Para datos agrupados en intervalos de clase, la moda, una vez determinada la clase modal, se calcula con la siguiente ecuación:

$$Mo = Lm + \frac{d1}{d1 + d2} w$$

donde:

Mo = moda

Lm =límite inferior de la clase modal

d_1 = diferencia entre la frecuencia de la clase modal y la premodal (clase anterior)

d_2 = diferencia entre la frecuencia de la clase modal y la posmodal (clase siguiente)

w = amplitud del intervalo de clase

En general, la clase modal es aquella que tiene la máxima frecuencia.

Varianza

La varianza poblacional (σ^2) se define como la suma de cuadrados de las desviaciones de los datos con respecto a la media, dividida entre el número total de datos, es decir:

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_1 - \mu)^2}{n}$$

La varianza muestral (S^2) (5) se obtiene dividiendo la suma de cuadrados de las observaciones de los datos con respecto a la media, entre el número total de datos menos uno; es decir:

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Desviación estándar

La desviación estándar se define como la raíz cuadrada positiva de la varianza; es decir:

$$\sigma = \sqrt{\sigma^2} \quad \text{(poblacional)}$$

$$S = \sqrt{S^2} \quad \text{(muestral)}$$

Coeficiente de variación

Es una medida relativa de dispersión, que relaciona la desviación estándar y la media; es decir:

$$C_v = \frac{S}{\bar{x}}$$

Sesgo

El sesgo es el estadístico que mide la simetría y asimetría (2).

El sesgo (γ) para datos poblacionales se obtiene con la siguiente ecuación:

$$\gamma = \frac{\mu 3}{\sigma^3}$$

donde:

$$\mu 3 = \frac{\sum_{i=1}^{n} (x_i - \mu)^3}{n}$$

$$\sigma = \sqrt{\frac{1}{n} \sum (x_1 - \mu)^2}$$

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

El sesgo para datos muestrales, se obtiene con:

$$\mathrm{Cs} = \frac{n^2 M_3}{(n-1)(n-2)S^3}$$

donde:

$$\mathrm{M}_3 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{n}$$

$$\mathrm{S} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Curtosis

Para datos poblacionales, el coeficiente de curtosis (k) se define mediante la siguiente ecuación:

$$k = \frac{\mu 4}{\sigma^4}$$

donde:

$$\mu 4 = \frac{\sum_{i=1}^{n} (x_i - \mu)^4}{n}$$

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2}$$

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$

El coeficiente de curtosis para datos muestrales (2) se define como:

$$\mathrm{C_k} = \frac{n^3 M_4}{(n-1)(n-2)(n-3)S^4}$$

donde:

$$\mathrm{M}_4 = \frac{\sum_{i=1}^{n} (x_1 - \bar{x})^4}{n}$$

$$S=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_1-\bar{x}\right)^2}$$

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$$

Estimación de parámetros

Los parámetros de una distribución teórica son variables que para cada conjunto de datos tiene un valor definido. Una vez que los parámetros quedan definidos, también queda definida la distribución teórica.

Por lo general, una función densidad o una función de distribución acumulada, puede escribirse como una función de la variable aleatoria y, en general, como una función de sus parámetros; así, por ejemplo, la función densidad de la distribución normal de variable aleatoria X, es:

$$f(x)=\frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

donde:

μ = parámetro de localización.

σ^2= parámetro de escala.

Para que la función *f(x)*, quede definida, debe calcularse los parámetros μ y σ^2.

Como normalmente no se conocen todos los valores de la variable aleatoria, la estimación de los parámetros se realiza a partir de una muestra (4).

Métodos de estimación de parámetros

Para determinar los valores numéricos de los parámetros de la distribución teórica, a partir de los datos muestrales, se utilizan varios métodos de estimación (5), siendo en orden ascendente de menor a mayor eficiencia, los siguientes:

- Gráfico
- Mínimos cuadrados
- Momentos
- Máxima verosimilitud

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.

Las pruebas estadísticas (4) 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 probabilidades.

Las pruebas de bondad de ajuste más utilizadas son:

- Ajuste gráfico
- Ajuste estadístico { Chi - cuadrado / Smirnov - Kolmogorov

Distribuciones teóricas

El hidrólogo generalmente tendrá disponible un registro de datos hidrometeorológico (precipitación, temperaturas, evapotranspiración, caudales, etc.), a través de su conocimiento del problema físico, escogerá un modelo probabilístico por usar, que represente en forma satisfactoria el comportamiento de la variable.

Distribución normal o gaussiana

Se dice que una variable aleatoria X tiene una distribución normal (2), si su función densidad, es:

$$f(x)=\frac{1}{\sqrt{2\Pi S}}EXP\left[-\frac{1}{2}\left(\frac{x-\bar{X}}{S}\right)^2\right]$$

ó

$$f(x)=\frac{1}{\sqrt{2\Pi S}}e^{-\frac{1}{2}\left(\frac{x-\overline{X}}{S}\right)}$$

para - ∞ < x < ∞

donde:

$f(x)$ = función densidad normal de la variable x

x = variable independiente

$\overline{X}$ = parámetro de localización, igual a la media aritmética de x

S = parámetro de escala, igual a la desviación estándar de x

EXP = función exponencial con base *e*, de los logaritmos neperianos.

Cuando la variable aleatoria *X* se distribuye normalmente con media $\mu = \overline{X}$ y varianza ($\sigma^2 = S^2$), se denota de la siguiente forma:

$$X \sim N\left(\overline{X},S^2\right)$$

Distribución log-normal de 2 parámetros

La variable aleatoria *X* es positiva y el límite inferior x_o no aparece.

La variable aleatoria: $Y = lnX$, es normalmente distribuida con media μ y y varianza σ^2 y

Se usan estos parámetros para especificar que la distribución es logarítmica, puesto que también puede usarse la media y la varianza de X.

Se dice que una variable aleatoria X tiene una distribución log-normal de 2 parámetros (5), si su función densidad es:

$$f(x)=\frac{1}{x\sqrt{2\Pi\sigma_y}}e^{-\frac{1}{2}\left[\frac{1nx-\mu_y}{\sigma_y}\right]^2}$$

para 0 < x < ∞

$$x \sim \log N\left(\mu_y,\sigma_y^2\right)$$

Donde μ_y, σ y, son la media y desviación estándar de los logaritmos naturales de *x*; es decir, de *lnx*, y representan, respectivamente, el parámetro de escala y el parámetro de forma de la distribución.

Distribución gamma de dos parámetros

Se dice que una variable aleatoria X, tiene una distribución gamma de 2 parámetros (5), si su función densidad de probabilidad es:

$$f(x)=\frac{x^{\gamma-1}e^{-\frac{x}{\beta}}}{\beta^{\gamma}\Gamma(\gamma)}$$

para:

$0 \leq x < \infty$

$0 < \gamma < \infty$

$0 < \beta < \infty$

siendo:

γ = parámetro de forma (+)

β = parámetro de escala (+)

$\Gamma(\gamma)$ = función gamma completa, definida como:

$\Gamma(\gamma)=\int_0^\infty x^{y-1}e^{-x}dx$ que converge si $\gamma > 0$

Distribución log-Pearson tipo III

Se dice que una variable aleatoria *X* tiene una distribución log-Pearson tipo II [5], si su función densidad de probabilidad es:

$$f(x)=\frac{\left(1nx-x_0\right)^{\gamma-1}e^{-\frac{1nx-x_0}{\beta}}}{x\beta^{\gamma}\Gamma(\gamma)}$$

para:

$x_o \leq x < \infty$

$-\infty < x_o < \infty$

$0 < \beta < \infty$

$0 < \gamma < \infty$

donde:

x_o = parámetro de posición

ß = parámetro de escala

γ = parámetro de forma

Distribución Gumbel

La distribución Gumbel (5) es una de las distribuciones de valor extremo, es llamada también Valor Extremo Tipo I, Fisher-Tippett tipo I o distribución doble exponencial.

La función de distribución acumulada de la distribución Gumbel, tiene la forma:

F(x) = EXP(-EXP(-(x - μ) /α))

ó

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

para: $-\infty < x < \infty$

donde:

$0 < \alpha < \infty$ es el parámetro de escala

$-\infty < \mu < \infty$ es el parámetro de posición, llamado también valor central o moda

Derivando la función de distribución acumulada con respecto a *x*, se obtiene la función densidad de probabilidad; es decir:

$$f(x) = \frac{dF(x)}{dx}$$

$$f(x) = \frac{1}{\alpha} EXP\left(-\frac{(x-\mu)}{\alpha} - EXP\left(-\frac{(X-\mu)}{\alpha}\right)\right)$$

ó

$$f(x) = \frac{1}{\alpha} e^{-\frac{x-\mu}{\alpha} - e^{-\frac{x-\mu}{\alpha}}}$$

para:

$-\infty < x < \infty$

Distribución log-Gumbel

La función de distribución acumulada de la distribución Gumbel (5) tiene la forma:

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

para: $-\infty < x < \infty$

donde:

$0 < \alpha < \infty$ es el parámetro de escala

$-\infty < \mu < \infty$ es el parámetro de posición, llamado también valor central o moda

si en la ecuación, la variable *x* se reemplaza por *lnx*, se obtiene la función acumulada de la distribución log-Gumbel, o distribución de Fréchet.

Medición de caudales (aforos)

La hidrometría es la rama de la Hidrología que estudia la medición del escurrimiento. Para este mismo fin, es usual emplear otro término denominado *aforo*. Aforar una corriente significa determinar, a través de mediciones, el caudal que pasa por una sección dada y en un momento dado (6).

Caudales máximos

Para diseñar (6):

- Las dimensiones de un cauce
- Sistema de drenaje
  - agrícola
  - aeropuerto
  - ciudad
  - carretera
- Muros de encauzamiento para proteger ciudades y plantaciones
- Alcantarillas
- Vertedores de demasías
- Luz en puentes

Se debe calcular o estimar el caudal de diseño que, para esos casos, son los caudales máximos.

Algunos métodos para el cálculo del caudal máximo, son:

- Método directo
- Métodos empíricos
- Método del número de curva
- Métodos estadísticos
- Métodos hidrológicos

Evapotranspiración

Son las pérdidas totales: evaporación de la superficie evaporante (del suelo y agua) + transpiración de las plantas.

El término ***evapotranspiración potencial*** (6) fue introducido por Thornthwaite y se define como la pérdida total del agua que ocurriría si en ningún momento existiera deficiencia de agua en el suelo para el uso de la vegetación.

Se define como el ***uso consuntivo*** la suma de la evapotranspiración y el agua utilizada directamente para construir los tejidos de las plantas.

Como el agua para construir los tejidos comparada con la evapotranspiración es despreciable, se puede tomar:

Uso consuntivo ≈ evapotranspiración

En los proyectos de irrigación interesa hacer cálculos previos de las necesidades de agua de los cultivos. Estas necesidades de agua, que van a ser satisfechas mediante el riego, viene a constituir la evapotranspiración o el uso consuntivo.

Para el cálculo de estas cantidades de agua, se han desarrollado métodos basados en datos meteorológicos, de los cuales los más conocidos son el Thornthwaite y el de Blaney –Clidde.

Materiales y métodos

Para el desarrollo del *software*, se utilizó *Visual Basic*, versión 6.0, la cual permite crear aplicaciones de 32 bits.

Para la solución de las ecuaciones, se utilizarán los métodos numéricos, tales como:

- algoritmo de Newton Raphson
- método de la secante
- integración gráfica
- interpolación de Lagrange
- algoritmo de Romberg

Para el cálculo de las distribuciones acumuladas teóricas como la Normal, Log-Normal, Gamma, y el cálculo de sus valores inversos, se utiliza el desarrollo de series.

Muchas soluciones de las ecuaciones se realizan mediante nomogramas y tablas. Para un proceso computacional, estas se deben transformar a ecuaciones, por lo que se investigarán las ecuaciones más adecuadas, utilizando:

- correlaciones lineales simples
- correlaciones no lineales, simples
- correlaciones lineales múltiples
- correlaciones no lineales múltiples

Resultados

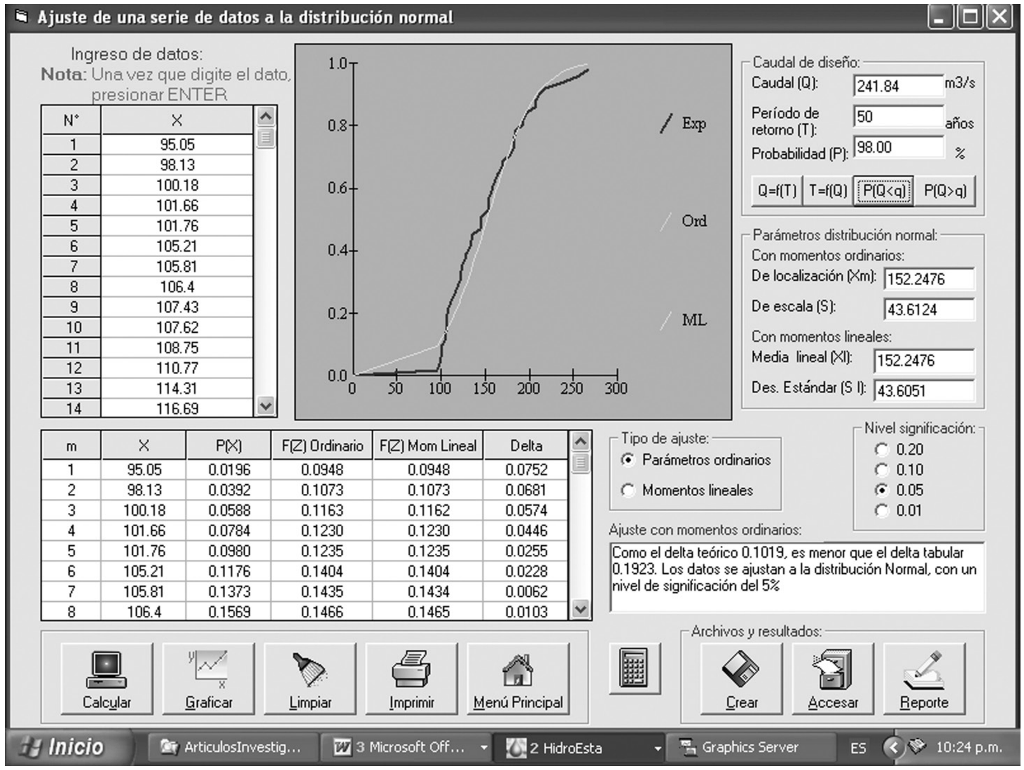
El producto de este trabajo es HidroEsta. Una de sus pantallas, correspondiente al cálculo de la distribución normal, se muestra en la figura de la siguiente página.

Discusión

Los resultados computacionales obtenidos con la aplicación son, en todos los casos, más aproximados que los obtenidos con los nomogramas.

HidroEsta representa una contribución para simplificar los estudios hidrológicos. Es importante porque:

- Proporciona una herramienta novedosa y fácil de utilizar para el ingeniero civil, ingeniero agrícola, ingeniero agrónomo y otros especialistas que trabajen en el campo de los estudios hidrológicos.
- Permite, simplificar el proceso de la abundante información y los



cálculos laboriosos.

- Permite a partir de la información proporcionada, simular los parámetros de diseño de las estructuras por construir.
- Reduce enormemente el tiempo de cálculo.
- Permite obtener un diseño óptimo y económico.

Agradecimiento

Un especial agradecimiento a la Vicerrectoría de Investigación y Extensión del Instituto Tecnológico de Costa Rica, que facilitaron el tiempo y el material para realizar este trabajo.

Referencias bibliográficas

1 Haan, Ch. Statistical methods in hydrology. Iowa State University Press, U.S.A., 1977.

2 Kite, G.W. Frequency and risk analyses in hydrology. Water Resources Publications, Michigan, U.S.A. 1977.

3 Maidment, David. Handbook of hydrology. Editorial McGraw-Hill. New Cork – USA. 1993.

4 Villón, Máximo. Diseño de capacidad de embalses por el método experimental-teoría de range. Tesis para optar por el grado de Magíster Scientiae en Ingeniería de Recursos de Agua y Tierra. Universidad Nacional Agraria “La Molina”. Lima – Perú, 1983.

5 Villón, Máximo. Hidrología estadística. Taller de Publicaciones. Instituto Tecnológico de Costa Rica. Cartago - Costa Rica. 2001.

6 Villón, Máximo. Hidrología. Taller de Publicaciones. Instituto Tecnológico de Costa Rica. Cartago – Costa Rica. 2002.