



Universidad Autónoma de Querétaro
Facultad de Ingeniería
Doctorado en Ingeniería

DESARROLLO, COMPILACIÓN Y SISTEMATIZACIÓN DE ALGORITMOS
MATEMÁTICOS PARA LA MODELACIÓN HIDROLÓGICA AVANZADA

Opción de titulación
TESIS

Que como parte de los requisitos para obtener el grado de
Doctorado en Ingeniería

Presenta:

M. en C. Israel Ruiz González

Dirigido por:

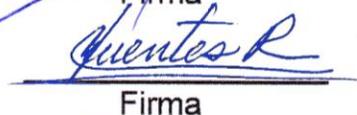
Dr. M. Alfonso Gutiérrez López

Dr. M. Alfonso Gutiérrez López
Presidente



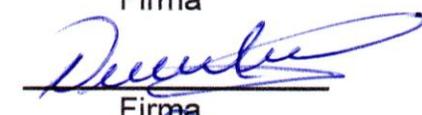
Firma

Dr. Carlos Fuentes Ruiz
Secretario



Firma

Dr. Eusebio Jr. Ventura Ramos
Vocal



Firma

Dr. Carlos Chávez García
Suplente

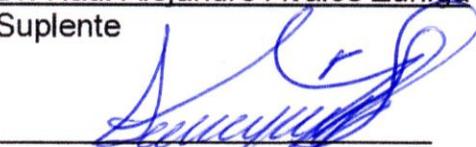


Firma

Dr. Raúl Alejandro Ávalos Zúñiga
Suplente



Firma



Dr. Aurelio Domínguez González
Director de la Facultad de Ingeniería



Dra. Ma. Guadalupe Flavia Loarca Piña
Director de Investigación y Posgrado

Centro Universitario
Querétaro, Qro.
Diciembre 2016

RESUMEN

En su parte central la Hidroinformática trata con la modelación de la información relacionada con la Hidrología. La escasa utilización de la modelación espacial de variables hidrológicas dentro de los modelos de predicción y pronóstico, es un campo de investigación y de innovación que debe ser explotado y desarrollado por la hidrología moderna. Es por esto que la tesis plantea que una plataforma computacional en la que se presenten algoritmos conocidos, se utilicen distribuciones mezcladas de probabilidad y se planteen nuevos algoritmos que contemplen las condiciones geomorfológicas, facilitará la modelación hidrológica y además permitirá tener datos más confiables y en tiempo real sobre la distribución espacial y temporal de las variables hidrológicas. Se tiene como objetivo desarrollar una plataforma computacional amigable y robusta que permita al usuario realizar estudios de modelación hidrológica mediante la compilación y sistematización de algoritmos matemáticos complejos, que además incluya rutinas para la mejora de los resultados de las metodologías actuales en las áreas de: Geoestadística, Hidrología estocástica e hidrología paramétrica. El software generado en la presente investigación está dividido en módulos según la disciplina a la que pertenezca, cada módulo tiene sus algoritmos particulares. Y se dividen en análisis de frecuencias, curvas IDT, lógica difusa, manejador de datos, series de tiempo y geoestadística. El análisis de frecuencias es un procedimiento para la estimación de eventos de diseño, asociados a un cierto período de retorno, basado en la observación de los eventos ya ocurridos, las curvas IDT son curvas que resultan de unir los puntos representativos de la intensidad media en intervalos de diferente duración, y correspondientes todos ellos a una misma frecuencia. La geoestadística es la ciencia que se encarga estudiar la información espacial. Una isolínea para una función de varias variables, es una curva que conecta los puntos en que la función tiene un mismo valor constante. Una serie de tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí. Ciertos módulos tienen comprobación de resultados con publicaciones anteriores, como el caso de las curvas IDT del país con las curvas IDT presentadas por la SCT; comparativo con software comercial actual, el módulo análisis de frecuencias con desarrollos como Ax o Easyfit; y aplicaciones predictivas o diagnósticas a datos ya medidos, módulos lógica difusa, series de tiempo e interpolación espacial.

(Palabras clave: Hidroinformática, tiempo real, isolínea, lógica difusa, geoestadística)

SUMMARY

Hydroinformatics focuses in hydrologic data modelling. The lack of use of measured hydrological data is a research field with big opportunities, thus it is essential a work in this topic. With observation, it is clear that Mexico do not have an informatics tool that obtains hydrometeorological real time data and process info in order to make forecast and analyze spatial behavior of variables. Therefore, this thesis propose that an informatics tool that shows known algorithms, applies mixed probability distributions and formulate new algorithms that underlay geomorphological conditions, ease the hydrological modelling and allow use of real time reliable data about spatial distribution of hydrological variables. The main goal of the work is to develop an informatics tool that allows users to do hydrological modelling based on compilation and systematization of complex known algorithms, besides it should include routines that improve actual results in areas such as: Geostatic, Stochastic and parametric Hydrology. The software made in this research is divided in: frequency analysis, IDF Curves, fuzzy logic, data management, time series and geostatics. Frequency analysis is a procedure for estimation of design events associated with certain return periods, based on phenomena that happened. Applying probability techniques, a probability function is selected and represents the relationship of an event with the observed frequency. IDF curves result by joining representative points with the mean rainfall intensity measured at non equals time intervals that each one correspond with the same frequency or return period. Geostatics is the science that focuses in spatial information; an isoline for a multi variable function is a curve that joins the points where the given function has constant value. A Time series is a sequence of observations measured data given time and ordered chronologically, equal separated and uniformly spaced; this way data is dependent itself. Certain units have validation with previous results published, such as official IDF curves published by SCT; or are compared with commercial software such as AX or Easyfit; and /or with predictive applications or forecasted from measured data.

(Key words: Hydroinformatics, real time, isolines, fuzzy logic, geostatics)

Marce, eres la persona más importante de mi vida

Para ti y contigo

AGRADECIMIENTOS

La tesis además del instrumento de acreditación y titulación de doctorado es el camino y el aprendizaje tanto académico como trascendental. En este viaje que ha durado a nivel tesis 3 años y a nivel titulación doctorado 4 y medio, me tocó pasar de la independencia al compromiso y a la vida familiar, antes que otra cosa, Marce creo que formamos un maravillosa familia en la que además de crecer intelectualmente, crecemos juntos como personas y podemos crear a diario un hogar maravilloso. Eres la primera persona a quien le debo agradecer no sólo tu participación en la tesis sino tu participación en este viaje.

Hablando del tránsito es indispensable agradecer a quienes me brindan apoyo incondicional y forman mi familia fuera de De Paula, Jesús, Guadalupe, Paco, Ricardo, Catia; gracias por sus buenas intenciones, consejos ayuda y apoyo para Marce y para mí.

Para con los maestros y sinodales que me han dado herramientas para la escritura y programación de esta tesis, tengo una deuda intelectual y profesional con ustedes y más importante, han sido excelentes amigos y compañeros de viaje; Alfonso, Eusebio, Primo Carlos, Carlos y Raúl han sido un ejemplo de cómo convivir, ayudar y trabajar con alumnos y colaboradores, les debo además una parte importante de mi formación como profesor y coordinador.

A los becarios y amigos que directamente trabajaron en este viaje conmigo, son la piedra sobre la que está edificado el viaje, son el medio de propagación y el medio de transporte para poder cruzar de manera pronta, segura y eficiente el camino. Jaime, Jorge, Raúl, Dionio, Luis Ángel, Porrás, Hugo, Carlos, Alejandro, Violeta, Raúl y todos los que me faltan, la deuda con ustedes si es verdaderamente impagable.

Amigos y compañeros, Roberto, Chucho, Noé, Vago, Arthur, Cuevas, Fut Inges, equipo de Posgrado y todos los que estoy omitiendo son culpables de las distracciones y de este combustible para cargar energías y llevar el barco a buen puerto.

El mejor y mayor logro de esta tesis es que ha sido el camino que me ha permitido el desarrollo profesional y tener las herramientas para mantener y mejorar mi hogar. Omar y Arturo, gracias por confiar en mí como maestro y permitirme formar parte de la facultad de Ingeniería UAQ y de la coordinación de Mecatrónica en la Universidad Anáhuac Querétaro. Luz, gracias por la confianza y apoyo para poder desarrollar mi trabajo. Alberto, Carina, Roberto, Pablo, Christian, Jazz, Alan y Javier, juntos seguimos en este viaje formando líderes de acción positiva y a los mejores ingenieros de Querétaro.

Finalmente, UAQ, eres mi casa, mi sustento y una institución tan grande que permite desarrollar verdaderas familias e historias de éxito como la mía.

ÍNDICE

1. INTRODUCCIÓN.....	1
1.1. Descripción del problema	3
1.2. Justificación.....	4
1.3. Hipótesis y Objetivos	8
1.3.1. Hipótesis	8
1.3.2. Objetivo general.....	8
1.3.3. Objetivos Particulares	8
2. REVISIÓN DE LITERATURA.....	10
2.1. Métodos numéricos	10
2.2. Curvas IDT	11
2.3. Lógica Difusa.....	18
3. METODOLOGÍA	22
3.1. Red CIAQ	22
3.1.1. Sistema de Adquisición Automatizado (SAA).....	23
3.1.2. Sistema de Acceso a Reportes y Herramientas (SARH)	25
3.2. Análisis de frecuencias.....	25
3.2.1. Distribuciones univariadas	28
3.2.2. Distribuciones bivariadas	29
3.2.3. Estimación de parámetros.....	32
3.2.4. Análisis de frecuencias a partir de distribuciones de probabilidad univariadas	35
3.2.5. Distribución Normal.....	35
3.2.6. Distribución Log-Normal.....	37
3.2.7. Distribución Exponencial.....	37
3.2.8. Distribución Gamma.....	38
3.2.9. Distribución Gumbel.....	39
3.2.10. Metodología propuesta seleccionando dos distribuciones univariadas diferentes.	40
3.3. Curvas IDT	43
3.4. Geoestadística	45

3.4.1.	Isoyetas	48
3.5.	Series de tiempo	50
3.5.1.	Proceso Autorregresivo de Orden 1: <i>AR1</i>	51
3.5.2.	Proceso Autorregresivo de Orden 2: <i>AR2</i>	52
3.5.3.	Proceso Autorregresivo de Orden 3: <i>AR3</i>	52
3.5.4.	Criterio de selección de Akaike	52
3.6.	Desarrollo de la interfaz gráfica de usuario MATLAB.....	53
4.	RESULTADOS	67
4.1.	Módulo Curvas IDT.....	67
4.2.	Módulo de Análisis de Frecuencias	88
4.3.	Módulo Interpolación Espacial	92
4.4.	Módulo de Lógica Difusa	99
4.5.	Módulo Manejador de Red CIAQ.....	104
4.6.	Módulo Series de Tiempo.....	110
5.	CONCLUSIONES Y RECOMENDACIONES	118
6.	REFERENCIAS	122

ÍNDICE DE TABLAS

Tabla 4-1 Registros Precipitación máxima en 24 horas en Calvillo	68
Tabla 4-2 Registros curva IDT Calvillo	69
Tabla 4-3 Ráster con valores de Intensidad de Lluvia	70
Tabla 4-4 Comparativo módulo IDT vs Curvas oficiales SCT	82
Tabla 4-5 Estaciones de Querétaro T=10 años	83
Tabla 4-6 Estaciones de Querétaro T=25 años	85
Tabla 4-7 Estaciones de Querétaro T=50 años	86
Tabla 4-8 Ejemplo Interpolación espacial.....	92
Tabla 4-9 Interpolación espacial en red CIAQ	96
Tabla 4-10 resultados interpolación espacial Red CIAQ	98
Tabla 4-11 Resultados tiempo de concentración Kirpich.....	101

ÍNDICE DE FIGURAS

Figura 3-1. Arquitectura básica de Red CIAQ.	22
Figura 3-2. Diagrama de clases del SAAA.	23
Figura 3-3. Diagrama de flujo principal.....	24
Figura 3-4. Zona de estudio Red CIAQ.	50
Figura 3-5. Acceso a un GUI.	54
Figura 3-6 Ventana de edición gráfica GUIDE.....	55
Figura 3-7 Insertar botones en la interfaz gráfica.	56
Figura 3-8 Componentes de una interfaz gráfica.....	57
Figura 3-9 Alineación de botones.....	58
Figura 3-10 Menú de propiedades elementos GUI.....	59
Figura 3-11 Modificación etiqueta de un elemento	60
Figura 3-12 Boceto de una interfaz gráfica.....	61
Figura 3-13 Ejemplo de una interfaz gráfica con ejes 3D	63
Figura 3-14 Acceso a callback de elemento gráfico	64
Figura 4-1 Curvas IDT en Calvillo	70
Figura 4-2 Interfaz de bienvenida Módulo IDT.....	72
Figura 4-3 Panel de Control con botones de radio	73
Figura 4-4 Selección de Estado y condiciones de tormenta	74
Figura 4-5 Selección de punto de análisis por geografía.....	75
Figura 4-6 Latitud y Longitud previamente definidas	76
Figura 4-7 Ventana para análisis puntual de una curva IDT	77
Figura 4-8 Mensaje de espera debido al tiempo de ejecución del módulo IDT	77
Figura 4-9 Gráfica con las curvas IDT 10, 25 y 50 años.....	78
Figura 4-10 Herramientas gráficas en la interfaz.....	79
Figura 4-11 Curva extrapolada.....	80
Figura 4-12 Cursos de datos	81
Figura 4-13 Comparativo módulo IDT vs Curvas SCT.....	83
Figura 4-14 Comparativo para T=10 años.....	84
Figura 4-15 Gráfica Error relativo en Baja California	85
Figura 4-16 Comparativo para T= 25 años.....	86

Figura 4-17 Comparativo para T=50 años.....	87
Figura 4-18 Interfaz gráfica Análisis de Frecuencias.....	88
Figura 4-19 Ajuste de Frecuencias por Gumbel a 15 registros.....	89
Figura 4-20 Límites de confianza en la plataforma.....	90
Figura 4-21 Límites de confianza a la par en la ventana gráfica.....	90
Figura 4-22 Rompimiento gráfico de serie.....	91
Figura 4-23 Serie combinada Gumbel-Gumbel.....	92
Figura 4-24 Interfaz gráfica Interpolación Espacial.....	93
Figura 4-25 Mapa de isolíneas.....	93
Figura 4-26 Mapa de isolíneas con relleno sólido.....	94
Figura 4-27 Herramientas gráficas interpolación Especial.....	95
Figura 4-28 Interpolación puntual espacial.....	96
Figura 4-29 Interpolación espacial en red CIAQ.....	97
Figura 4-30 Mapa Isolíneas interpoladas en red CIAQ.....	98
Figura 4-31 Módulo de Lógica Difusa.....	99
Figura 4-32 Cuadro de diálogo para lectura de archivo.....	100
Figura 4-33 Corrida del módulo Lógica Difusa.....	101
Figura 4-34 Gráfica comparativa del modelo de Lógica Difusa vs cálculo de concentración de Kirpich y el error porcentual del modelo.....	103
Figura 4-35 Módulo Manejador de Red CIAQ.....	104
Figura 4-36 Presentación de datos de la Estación Chulavista el día 18 de abril de 2016 a las 19:59 h en el Manejador Red CIAQ.....	105
Figura 4-37 Figura 4 35 Presentación de datos de múltiples estaciones el día 9 de abril de 2016 a las 13:34 h en el Manejador Red CIAQ.....	105
Figura 4-38 Estaciones en tiempo real en simultáneo.....	106
Figura 4-39 Interfaz para manejo gráfico de datos históricos.....	107
Figura 4-40 Datos históricos descargados desde Internet.....	108
Figura 4-41 Gráficos del mes en Chulavista.....	109
Figura 4-42 Datos gráficos en simultáneo para la estación Chulavista.....	109
Figura 4-43 Interfaz de bienvenida para módulo Series de Tiempo.....	110
Figura 4-44 Correlogramas sesgados e insesgados.....	111
Figura 4-45 Métodos de normalización de muestras.....	112
Figura 4-46 Normalización por el estadístico de Box y Cox.....	112

Figura 4-47 Normalización por doble potencia	113
Figura 4-48 Estadísticos de la muestra.	113
Figura 4-49 Tablas de coeficientes phi y valor de criterio de Akaike Modelo AR(1)	114
Figura 4-50 Coeficientes phi modelo AR(2).....	114
Figura 4-51 Coeficientes modelo AR(3)	115
Figura 4-52 Muestras sintéticas por procesos AR	116
Figura 4-53 Interfaz de generación de muestras	116
Figura 4-54 Caminata hacia el futuro con un modelo AR	117

1. INTRODUCCIÓN

Una breve presentación de dos posibles definiciones de la Hidroinformática puede apreciarse, por ejemplo, en Price (1997) y Abbot (1991): tecnología de la Información para el manejo del agua; disciplina básica de Hidroingeniería que soporta el desarrollo sostenible del medio acuático a través del uso de los computadores y las redes. En su parte central la Hidroinformática trata con la modelación de la información relacionada con la Hidroingeniería, así como también con las plataformas de proyectos distribuidos y soportados por la ingeniería de sistemas computacionales y por los procesos de trabajo en ingeniería para diseño, operación, manejo, consultoría y administración.

Cormen et al. (2009) define un algoritmo como una lista de instrucciones para resolver un problema abstracto, él plantea que un número finito de pasos convierten los datos de entrada de un problema en una solución o salida. Por otro lado, Dershowitz y Gurevich (2008) señalan que un algoritmo es cualquier cosa que funcione paso a paso, donde cada paso se pueda describir sin ambigüedad y sin hacer referencia a una computadora en particular, y además tiene un límite fijo en cuanto a la cantidad de datos que se pueden leer o escribir en un solo paso. Esta amplia definición abarca tanto a algoritmos prácticos como aquellos que solo funcionan en teoría, por ejemplo, el método de Newton y la eliminación de Gauss-Jordan funcionan, al menos en principio, con números de precisión infinita; sin embargo, no es posible programar la precisión infinita en una computadora, y no por ello dejan de ser algoritmos.

La Tierra, con sus diversas y abundantes formas de vida, que incluyen a más de 6,000 millones de seres humanos, se enfrenta en este comienzo de con una grave crisis del agua. Todas las señales parecen indicar que la crisis está empeorando y que continuará haciéndolo, a no ser que se emprenda una acción correctiva (WWDR, 2006).

La crisis del agua debe situarse en una perspectiva más amplia de solución de problemas y de resolución de conflictos. Para hacer frente a este problema, una mejor gestión nos permitirá encarar la creciente escasez de agua per cápita en muchas partes del mundo (Aparicio et al., 2006).

La UNESCO a través del Programa Hidrológico Internacional (PHI) promueve la realización de balances hídricos en cuencas de los diversos países de América del Sur. Debido a esto, en los últimos años, en muchos países se han sucedido reformas en las legislaciones y organizaciones orientadas a la gestión y el aprovechamiento del recurso agua. Sin embargo, la diversidad de culturas, climas, así como las actividades políticas y financieras en cada uno de estos países ha ocasionado que cada día se busquen nuevos y mejores procedimientos para llevar a cabo una correcta gestión integrada de los recursos hídricos (Aparicio et al., 2006).

En México uno de los problemas más importantes es la asignación del recurso hídrico por cuenca hidrológica, para ello se cuenta con una la Ley de Aguas Nacionales (LAN), toda esta normativa ha permitido sin duda reunir esfuerzos hacia la correcta administración del recurso, ya que el estudio de la disponibilidad en cantidad y calidad se vuelve prioritario en el marco del desarrollo sustentable de un país. (Aparicio et al., 2006).

Un problema común en hidrología y los recursos del agua es la imprecisa estimación de la precipitación para eventos de poca ocurrencia debido a la falta de datos. Además, en la mayoría de los análisis de frecuencias se considera que los datos provienen de una sola población. Por ejemplo, en Latinoamérica, la función de distribución de Gumbel ha sido ampliamente utilizada (Heo et al., 2001). Sin embargo, es muy frecuente encontrar que los fenómenos son el resultado de factores causantes distintos. Este hecho fue reconocido por Hazen (1930) quien planteó la existencia de poblaciones mezcladas en las muestras de datos. En estos casos, se dice que la distribución de la variable aleatoria es una combinación de dos o más distribuciones básicas (Varas y Lara, 1998) y que, además, los picos

máximos de las tormentas se encuentran mutuamente correlacionados con la cantidad de tormentas correspondientes (Yue, 2000). Por lo que su análisis ha requerido el desarrollo de la teoría estadística, en la aplicación para valores extremos en el diseño de recursos del agua y su manejo (Viessman et al., 2002).

1.1. Descripción del problema

La escasa utilización de la modelación espacial de variables hidrológicas dentro de los modelos de predicción y pronóstico, es un campo de investigación y de innovación que debe ser explotado y desarrollado por la hidrología moderna (Rajyalakshmi y Dutta, 2006).

Como lo mencionan Gutiérrez y Ramírez (2005), uno de los principales problemas a los que se enfrenta un ingeniero hidrólogo es, sin duda, el diseño de obras hidráulicas en zonas con poca o nula información climatológica e hidrométrica. Por otra parte, el cambio climático provoca una modificación en los patrones tradicionales del régimen de precipitación y escurrimiento, aumentando la incertidumbre en la estimación de eventos futuros.

Diversos organismos en el país administran las estaciones hidrometeorológicas se encuentran organismos de distintos orígenes tales como: Dependencias de gobierno (CONAGUA, CONAFOR, SEDEA, SMN), Universidades y Centros de investigación (CENAPRED, UNAM, CIAQ, UAQ) e inclusive la iniciativa privada (Fundación PRODUCE, fundación FEMSA). Cada uno de estos organismos tiene sus objetivos y necesidades propias, por lo que cada estación está configurada y enfocada de distinta manera.

En México el organismo que regula la disponibilidad y distribución del agua es la Comisión Nacional del Agua. La disponibilidad se calcula a partir de metodologías presentadas en la NOM-011-CONAGUA-2000, en estas metodologías se utilizan registros históricos y modelos de pronóstico para la estimación de los componentes de una ecuación de balance hídrico.

Generalmente los registros meteorológicos no están completos y tienen intervalos de medición irregulares; además los algoritmos mediante los que se completan y analizan los registros no consideran las condiciones geomorfológicas propias de las cuencas en el territorio mexicano.

No existe en México una herramienta computacional que integre las mediciones hidrometeorológicas en tiempo real y procese los datos para generar mapas de pronóstico y comportamiento espacial de las variables medidas.

Desde hace ya varios años en México como en otras partes del mundo, el apremio por resolver los problemas ha conducido a los especialistas a utilizar métodos empíricos con pocas bases científicas acerca de la dinámica hidrológica. Además de existir una tendencia a desarrollar modelos teóricos cada vez más complejos, sin datos de campo que los sustenten, o bien, fundamentados en mediciones escasas o de mala calidad, que difícilmente permiten verificar su validez (Garduño, 1990).

Finalmente, se encuentra que los algoritmos de interpolación geoespacial sugeridos para modelar la distribución espacial de los componentes del ciclo hidrológico, fueron desarrollados para condiciones geomorfológicas y climatológicas distintas a las encontradas en la República mexicana.

1.2. Justificación

Los datos con los que se modela el ciclo hidrológico provienen de fuentes diversas y no se encuentran homologados. Es necesario tener bases de datos confiables y en las que los datos estén estimados de forma correcta.

Un problema más es que los modelos no han sido aplicados satisfactoriamente a la distribución espacial de variables hidroclimatológicas ni meteorológicas. Por lo anterior, se requiere sistematizar algoritmos matemáticos para describir las características de los componentes de una ecuación de balance hídrico integral, y finalmente se desea que toda la ciencia en materia de hidrología

estocástica pueda ser empleada en forma espacial y no sólo temporal, ya que los problemas actuales requieren prioritariamente el estudio del componente espacial de variables hidrológicas. La puesta en marcha de la plataforma a desarrollar, puede ser una importante herramienta que puede ser ofrecida como una alternativa a la modelación tradicional. Estos procedimientos pueden traducirse en proyectos contratados con la CONAGUA y otros organismos nacionales e internacionales.

Los recursos hídricos accesibles para su aprovechamiento por el hombre tienen su origen en la precipitación pluvial que, al ocurrir sobre tierra firme, se divide en dos fracciones:

- Cerca de 70% del volumen de agua precipitado retorna a la atmósfera por evaporación.
- La fracción complementaria escurre superficialmente por las redes de drenaje natural -arroyos y ríos- hasta desembocar al mar o a cuerpos interiores de agua, o se infiltra y circula a través de acuíferos, que a su vez descargan a cuerpos y cursos superficiales, a través de manantiales o subterráneamente al mar.

Donde el agua no es desviada de manera artificial desde las fuentes hasta sus salidas al mar, a la parte baja de una cuenca interna o a la frontera interior de una unidad hidrogeológica, se desarrolla un sistema natural o virgen. Antes de que el hombre alterara el equilibrio hidrológico para satisfacer sus necesidades, el escurrimiento virgen sustentaba a ecosistemas. Por ello, la naturaleza puede ser visualizada como el primer y natural usuario del agua.

El equilibrio natural fue gradualmente afectado conforme el hombre fue aumentando la derivación artificial agua para satisfacer sus necesidades personales (uso doméstico), para la producción de alimentos (uso agropecuario) y para el desarrollo de procesos económicos (uso industrial).

Hasta el siglo XIX el aprovechamiento creciente del agua por el hombre con la consecuente reducción gradual de los escurrimientos naturales, en general, no

causó daños graves al ambiente. Sin embargo, en el transcurso del siglo XX la derivación del agua para diversos usos creció de modo acelerado, especialmente durante su segunda mitad, al grado que ahora existen porciones importantes de la superficie continental del planeta, en las cuales el ambiente ha sufrido daños graves; en casos extremos, irreparables. En vista de lo anterior, es de suma importancia tomar conciencia de que sólo una fracción de los escurrimientos naturales, superficiales o subterráneos, debe ser aprovechada por el hombre. Además de requerimientos del ambiente, existen limitaciones de índole técnica que reducen aún más la proporción de escurrimientos naturales aprovechables.

La porción accesible de los escurrimientos naturales de una cuenca, cuya infraestructura de regulación ha sido plenamente desarrollada, en la mayoría de los casos no supera el 70%, a la que hay que deducir los requerimientos del ambiente para determinar la cantidad de agua que puede destinarse a los diversos usos humanos.

Por otra parte, con base en los estudios que ha realizado la Comisión Nacional del Agua, se ha detectado que, en diversas regiones, entidades federativas y localidades del país, los volúmenes de agua concesionados superan el escurrimiento y la recarga de los acuíferos, situación que genera escasez del recurso, conflictos entre los usuarios y diversos efectos perjudiciales.

Considerando todo lo anterior y que la Ley de Aguas Nacionales y su Reglamento dispone que para el otorgamiento de asignaciones y concesiones se tomará en cuenta la disponibilidad media anual de agua, es necesario establecer las especificaciones para determinar con una metodología consistente, a nivel nacional, la disponibilidad media anual de aguas nacionales superficiales y subterráneas, como base técnica para regular su uso, de manera racional y equitativa.

En México, la LAN establece que para otorgar los títulos de concesión o asignación se tomará en cuenta la disponibilidad media anual de agua de la cuenca

hidrológica o acuífero en el que se haya de realizar el aprovechamiento. La CONAGUA tiene la obligación de publicar las disponibilidades, para lo cual se creó la norma NOM-011-CONAGUA-2000 "Conservación del Recurso Agua". Que establece las especificaciones y el Método para determinar la disponibilidad media Anual de las Aguas Nacionales" (SEMARNAT, 2002).

Asimismo, es necesario tener una herramienta matemática-hidrológica para evaluar la incertidumbre y variabilidad de dichos componentes. Además, los efectos de la distribución espacial en cada uno de los componentes deben tomarse en cuenta, en el sentido que la NOM prevé una asignación del recurso netamente espacial (volúmenes comprometidos aguas abajo).

De esta forma surge la necesidad de utilizar herramientas estocásticas, automáticas, programables, utilizables en la modelación de la mayoría de los procesos del ciclo hidrológico, estos modelos se aplican lo mismo a problemas relacionados con la modelación del flujo subterráneo (Palacios, 2004), como a problemas relacionados con el régimen de lluvia y escurrimiento de una región.

De la misma forma, el desconocimiento de la formación de las regiones hidrológicamente homogéneas está relacionado con la incomprensión del significado de las características fisiográficas de una región que pueden ser utilizadas para formar grupos o para ser utilizadas en procedimientos multivariados complejos. Este concepto de conformar regiones homogéneas sin duda es importante aplicarlo en un estudio de balance hídrico, pues con una simple ecuación lineal, no es posible explicar el comportamiento espacial de cada componente. Con estas regiones homogéneas conformadas para cada componente, será posible mejorar las políticas de gestión regionales. Finalmente, podemos decir que la combinación de modelos estocásticos para generar series sintéticas y de pronóstico; con modelos de regionalización hidrológica espacial constituye un invaluable avance en la modelación hidrológica moderna.

1.3. Hipótesis y Objetivos

1.3.1. Hipótesis

Una plataforma computacional en la que se presenten algoritmos conocidos, se utilicen distribuciones mezcladas de probabilidad y se planteen nuevos algoritmos que contemplen las condiciones geomorfológicas, facilitará la modelación hidrológica y además permitirá tener datos más confiables y en tiempo real sobre la distribución espacial y temporal de las variables hidrológicas.

1.3.2. Objetivo general

Desarrollar una plataforma computacional amigable y robusta que permita al usuario realizar estudios de modelación hidrológica mediante la compilación y sistematización de algoritmos matemáticos complejos, que además incluya rutinas para la mejora de los resultados de las metodologías actuales en las áreas de: Geoestadística, Hidrología estocástica e hidrología paramétrica.

1.3.3. Objetivos Particulares

1. Desarrollar un método que obtenga las curvas IDT del país.
2. Evaluar las curvas IDT generadas en esta plataforma con las curvas IDT oficiales publicadas por la SCT.
3. Generar una herramienta que combine distribuciones de probabilidad y ajuste datos para estimar eventos extremos en una serie.
4. Instalar y automatizar una red meteorológica para generar datos a tiempo real que alimenten la plataforma.
5. Programar una herramienta geoestadística para la interpolación espacial en tiempo real de datos hidrológicos.
6. Desarrollar herramientas para la descarga, análisis y almacenamiento de variables generadas en tiempo real en una red meteorológica.

2. REVISIÓN DE LITERATURA

2.1. Métodos numéricos

La historia de los métodos numéricos tiene origen a mediados del siglo XVI A. C. con el Papiro de Ahmes, también conocido como Papiro Rhind, un documento que contiene diversas cuestiones matemáticas, donde propone encontrar raíces de ecuaciones sencillas. Posteriormente Euclides un matemático griego, también conocido como el padre de la geometría, desarrolla un método para aproximar áreas de distintas figuras geométricas, con esto obtiene un valor muy aproximado al número PI. Siglos después son tomados en cuenta escritos de Al-Jwarizmi y Al-Karayi, dos matemáticos árabes, que implementaron el uso de métodos algebraicos dando origen a métodos algorítmicos para la resolución de problemas. En el siglo XVII, Gottfried Leibniz matemático alemán, desarrolla metodologías matemáticas dando origen al cálculo diferencial e integral. Contemporáneo a la época Isaac Newton, físico y matemático inglés, a quien también es atribuida la invención del cálculo infinitesimal, desarrolla una gran cantidad de métodos para realizar numéricamente procedimientos matemáticos, entre ellos uno de los más conocidos la interpolación polinomial. Años más tarde, tomando las aportaciones de Newton y Leibniz, Leonhard Euler otro físico matemático de origen suizo, desarrolla un método para encontrar soluciones aproximadas a problemas de ecuaciones diferenciales, con lo que da inicio a los métodos de integración numérica. A mediados del siglo XX surgen los primeros computadores, los cuales revolucionan la forma del estudio y análisis numérico, que años más tarde con el auge de las computadoras personales dieron origen a software especializado.

En la actualidad se ha vuelto de gran importancia el estudio de los métodos numéricos ya que estos se implementan en diferentes disciplinas dentro de la ingeniería, debido a que el uso de éstos permite la solución de problemas que no son fáciles de resolver analíticamente. Los Métodos Numéricos pueden ayudar a

encontrar una aproximación de la solución de un problema, sin embargo, no tenemos aún soluciones a todos ellos (Burden, 2001).

Los métodos numéricos constituyen técnicas mediante las cuales es posible formular problemas matemáticos, de tal forma que puedan resolverse utilizando operaciones aritméticas. Aunque existen muchos tipos de métodos numéricos, éstos comparten una característica común: invariablemente requieren de un buen número de tediosos cálculos aritméticos. No es raro que con el desarrollo de computadoras digitales eficientes y rápidas, el papel de los métodos numéricos en la solución de problemas en ingeniería haya aumentado de forma considerable en los últimos años (Chapra, 2006).

2.2. Curvas IDT

La lluvia se define mediante 3 variables: magnitud, duración y frecuencia. La magnitud de la lluvia es la precipitación total ocurrida durante la tormenta y se mide con unidades de longitud, milímetros o pulgadas; la frecuencia se expresa por el período de retorno de la lluvia o su intervalo de recurrencia, que es el tiempo en el que el evento puede ser igualado o excedido por lo menos en una ocasión. (Campos, 1990).

Por lo general los estudios sobre precipitaciones intensas se orientan alternativamente a la estimación de: (i) curvas intensidad-duración-frecuencia, i-d-f, o (ii) de tormentas máximas probables, PMP. Para estudios ligados a cuencas urbanas, obras viales y obras hidráulicas menores la característica habitual es el empleo de curvas i-d-f. Una solución ante la falta de estas curvas es el empleo de métodos de regionalización hidrológica. En este trabajo se abordan distintas metodologías de regionalización para la obtención de la intensidad de la precipitación en una determinada localidad. Las estimaciones son comparadas con el valor provisto por las curvas i-d-f del lugar (Bertoni et al., 2004).

Cuando se necesita determinar el caudal de diseño para una obra hidráulica se recurre al empleo de modelos de transformación de lluvia en caudal. Sin embargo, a veces, especialmente en cuencas pequeñas y/o urbanas, no se cuenta con la información pluviográfica necesaria para emplear esta herramienta.

Los registros temporales son necesarios para la confección de curvas Intensidad-Duración-Frecuencia (I-D-F) y son los datos de entrada en los modelos de lluvia-escorrentía. Consecuentemente, cuando esta situación se presenta es común recurrir al empleo de métodos alternativos que permitan obtener la variación temporal de las lluvias sobre la base de información pluviométrica diaria. De la comparación entre las curvas medidas y calculadas, se establece que tipo de metodología ajusta mejor a la distribución temporal real, para su empleo como un patrón regional en cuencas sin datos pluviográfica (Mattar et al., 2004). En Chile, hay una metodología que consistió en muestreos sobre los pluviogramas respectivos, con el fin de obtener las intensidades máximas por año, para las duraciones de 24, 12, 8, 6, 4, 2 y 1 hora. Los datos se llevaron a intensidad horaria y para cada duración se ajustó la función de Gumbel, que dio excelentes medidas de bondad de ajuste. Luego, se construyeron las curvas IDF para los períodos de retorno 5, 10, 20, 30, 40, 50, 60, 75, y 100 años. Posteriormente se construyen tablas de uso práctico para la extrapolación de las curvas hacia zonas sin pluviógrafos (Pizarro, 2001).

Estudiar las precipitaciones y conocer su distribución temporal es motivo de interés para diversos fines, por ejemplo, meteorológicos y edafológicos, como también hidrológicos, al tiempo de lo cual se pueden proporcionar índices para realizar estudios de crecidas o permitir la alimentación de modelos precipitación-escorrentía que permitan mejorar la información disponible, para un adecuado diseño y dimensionamiento de las obras civiles. Ahora bien, la disponibilidad de datos de caudal es imprescindible para el diseño y planificación de actividades físicas. Pero, muchas veces no se dispone de registros de caudales, o éstos no tienen la suficiente duración como para hacer los análisis de frecuencia requeridos;

debe entonces usarse la información pluviométrica para estimar crecidas de cierta frecuencia. Es, por lo tanto, muchas veces necesario presentar la información pluviométrica correspondiente a una tormenta o lluvia en formas de intensidades, a partir de los registros de las estaciones pluviográficas en estudio (Pizarro, 2001).

También existen estudios como éste realizados en otras partes de Sudamérica, algunos de los más importante han sido realizados en Argentina buscando resultados semejantes a los obtenidos en ese estudio mediante la obtención de grandes cantidades de información y filtración de la misma para poder calcular algunos de los parámetros más importantes de las curvas IDF y estimar algunos otros.

Argentina es un país con un territorio extenso y poca población lo cual sumado a una generalizada falta de interés por la medición hidrometeorológica, hace que la escasez de datos sea casi una constante a la hora de encarar proyectos de ingeniería civil. Esta situación es particularmente crítica cuando de precipitaciones intensas se trata (Devoto, 2009).

Devoto (2009) mismo propone un método para la estimación de las curvas IDF en la República Argentina para aquellas localidades sin información. Se apoya en 4 mapas con isolíneas de medias y coeficientes de variación para las precipitaciones intensas de 1 y 12 horas de duración y adopta a la función de distribución de Gumbel para la caracterización estocástica de las lluvias extremas.

El método es de aplicación en el cálculo de las crecidas de diseño de obras de drenaje urbano y el dimensionamiento de alcantarillas y de puentes viales o ferroviarios emplazados sobre cursos fluviales de cuencas pequeñas sin registros de aforos. Como consecuencia de ello, los valores que resultan son menos dependientes de la localización y de consideraciones climatológicas particulares, lo cual favorece su regionalización.

La regionalización hidrológica explora al máximo las informaciones existentes, buscando estimar las variables hidrológicas en lugares carentes de datos o bien donde los existentes resultan insuficientes por cantidad o por calidad. Esta técnica permite explorar mejor las muestras puntuales y, en consecuencia, mejorar las estimaciones de las variables; verificar la consistencia de las series hidrológicas; identificar la falta de puestos de observación, etc. (Gananacias, 2009).

La regionalización de caudales se realiza en la cuenca alta del Río Cuarto, un sistema hidrológico serrano representativo de la provincia de Córdoba, Argentina. En la zona existen 7 estaciones hidrométricas en un radio de 100 km, las cuales poseen registros de 22 a 43 años. Las características de homogeneidad regional fueron evaluadas siguiendo el procedimiento de los Trazos Multidimensionales propuestos por Andrews en 1972, confirmándose que las regiones son hidrológicamente homogéneas.

Fueron aplicados diferentes métodos de regionalización: Avenida Índice, Estaciones Año, Box-Cox, Regresión y Correlación Lineal y Técnica de Momentos Estandarizados de Probabilidad Pesada. La comparación de resultados entre los caudales correspondientes a distintas recurrencias obtenidos por análisis de frecuencia puntual en cada una de las estaciones y por regionalización mostró comportamientos dispares, dependientes de la calidad de la información básica. En varios casos el error de estimación no superó el 5%, considerado satisfactorio para los fines perseguidos (Gananacias, 2009).

Es importante mencionar, que las curvas IDF no sólo son de mucha utilidad en nuestro país, sino que en otros países son de relevancia para el diseño y planificación de estudios hidrológicos. En este sentido, Chen (1983) propone una fórmula general para representar la relación intensidad-duración-frecuencia para los Estados Unidos. Este método tiene gran potencial para la aplicación en el diseño de drenajes urbanos, y es actualmente utilizado en dicho país por la "Federation Highway Administration", siendo necesario para la utilización de este método, las

lluvias de 10 años de período de retorno y 1 hora de duración, lluvias de 24 horas con 10 años de período de retorno y lluvias de 1 hora con 100 años de período de retorno. La fórmula propuesta por Chen es la siguiente:

$$I_t^T = \frac{aI_t^{10} \log(10^{2-x} T^{x-1})}{(t+b)^c} \quad (2.1)$$

donde:

I_t^T = Intensidad de lluvia en mm/h, correspondiendo a T años de período de retorno.

R_t^{100} y R_t^{10} = Precipitaciones asociadas a una duración de t horas y 100 ó 10 años de período de retorno.

x = Relación entre las frecuencias R_t^{100} y R_t^{10}

I_t^{10} = Intensidad de lluvia asociada a 10 años de período de retorno y duración de t horas.

Por otro lado, en India, la asignación de un periodo de retorno de acuerdo a una duración e intensidad es necesaria para realizar estimaciones de inundaciones (Kothyari, 1992), para lo cual fue necesario utilizar los registros de lluvia menores a 24 horas de duración y para regiones hidrológicas muy similares. A partir de ellos, se planteó una fórmula general para describir la relación IDF, la cual es representada por la simple ecuación de Bernard (1932),

$$I_t^T = \frac{a_0 T^{a_1}}{t^{a_2}} \quad (2.2)$$

donde:

I_t^T = intensidad de precipitación para una duración t y un período de retorno o recurrencia de T años

a_0, a_1 y $a_2 =$ constantes.

Otros estudios se han llevado a cabo en India, y variados son los autores que han propuesto distintas fórmulas para la estimación de la intensidad de lluvia. Es en este sentido, que Bell (1969) citado por Kothyari (1992), propuso la siguiente fórmula:

$$R_t^T = (0.21 \ln T + 0.52)(0.54t^{0.25} - 0.50)R_1^{10} \quad (2.3)$$

donde:

R_t^T = Intensidad de precipitación para un período de retorno de T años y una duración de t horas.

R_1^{10} = Lluvia de 1 hora de duración y un período de retorno de 10 años.

Años más tarde, usando los datos de 39 estaciones pluviográficas "Rain Babu", Tejwani et al. (1979) propusieron diferentes relaciones en distintas localidades geográficas de India. Dichas relaciones produjeron resultados con un error de (+/-) 5%. Posteriormente se fueron incorporando nuevas estaciones de control lo que derivó en una nueva fórmula planteada por Chen (1983), que pudo representar la relación IDF, y obtener a la vez mejores resultados que los arrojados por el estudio anterior. Chow et al. (1994), en la ciudad de Denver en Estados Unidos, a partir de una curva Intensidad-Duración-Frecuencia, desarrolló un hietograma de diseño utilizando incrementos de tiempo de 10 minutos para una tormenta de 10 años y 2 horas.

En México, también se han hecho gran cantidad de estudios, no sólo para regionalizar sino también en escalas espaciales mucho menores al territorio nacional, escalas estatales, en donde se buscan las relaciones de las curvas IDF

para diseñar obras estatales como diques o drenes a partir de la información de los pluviógrafos del Estado.

Ramirez-Cerda (2011) propone la aplicación de las curvas IDF que ocurren en la zona teniendo dos objetivos fundamentales, el primero consiste en contrastar en 10 pluviógrafos de la República Mexicana, un procedimiento propuesto recientemente para construir curvas IDF, basado en la fórmula de Chen. Y esto se realiza utilizando la información sobre isoyetas de intensidades y pluviométrica disponible en la República Mexicana. Habiendo verificado la capacidad reproductiva y la aproximación de tal enfoque para representar curvas IDF, el segundo objetivo se desarrolló aplicando tal procedimiento en más de 50 localidades del país, las capitales de los estados y algunas de sus ciudades más importantes para presentar los parámetros de la fórmula de Chen que definen en cada sitio sus curvas IDF, con duraciones que varían de 5 minutos a 24 horas y periodos de retorno de 5, 10, 25, 50 y 100 años.

Los resultados dan a entender que el procedimiento utilizado emplea la información de isoyetas de intensidades, disponibles en la República Mexicana (SCT 1990), haciendo uso de los registros pluviométricos para estimar las predicciones necesarias de precipitación máxima en 24 horas. Por lo tanto, puede ser aplicado en cualquier sitio o localidad de la República Mexicana que cuente con una estación pluviométrica, la cual permita obtener con cierta confiabilidad los valores citados.

Por otra parte, Pilgrim (1993), desarrolló un método de análisis de hietogramas basado en el ordenamiento de los intervalos de tiempo de una tormenta, teniendo en cuenta la profundidad de precipitación que ocurre en cada uno de ellos y repitiendo ese procedimiento en muchas tormentas en la región.

Esta metodología es la estándar en el diseño hidrológico en Australia.

La construcción de las curvas IDF es dada por Sherman mediante la ecuación:

$$i = \frac{k Tr^M}{(d+c)^N} \quad (2.4)$$

donde:

i = Intensidad de precipitación en mm/h.

Tr = Período de retorno.

$c = 0$. C es únicamente un parámetro gráfico para obtener resultados más legibles dentro de las gráficas de intensidad-duración.

Cuando se requiera construir curvas IDF, en caso de contar con un registro de pluviógrafo cercano, se recomienda seguir el procedimiento descrito en la parte de registros pluviográficos. De no ser así, será necesario emplear algún registro cercano de lluvias máximas diarias, el más largo y confiable posible, utilizando las distribuciones de probabilidad mencionadas. Si no se dispone de ningún registro pluviométrico confiable o de suficiente amplitud, se pueden obtener las magnitudes de la lluvia de planos de curvas isomáximas de la Secretaría de Recursos Hidráulicos. Posteriormente se define el coeficiente R con los valores reportados en el cuadro 2, con lo cual, mediante las fórmulas de Bell y de Chen, pueden obtenerse lluvias de determinado período de retorno para duraciones diferentes (Campos, 1990).

2.3. Lógica Difusa

En la actualidad, los fenómenos del mundo real resultan ser muy complejos. Lo anterior porque la mayoría de los fenómenos que encontramos día con día son imprecisos, es decir; son fenómenos que tienen implícito un cierto grado de imprecisión en la descripción de su naturaleza. (Kauffman, 1977). Ésta imprecisión

puede estar asociada con su forma, posición, momento, color, textura, o incluso con en la semántica que describe lo que son.

Además, en muchos casos el mismo concepto puede tener diferentes grados de imprecisión en diferentes contextos o tiempo (Calviño, 2003). La complejidad al tratar de estudiar y explicar este tipo de fenómenos en toda su extensión, se presenta en forma de incertidumbre y ambigüedad (Freeman et al., 1991).

El ser humano ha tenido que desarrollarse en medio de este tipo de fenómenos durante todo el transcurso de su historia, aun así, las computadoras diseñadas por él no son capaces de manejar la complejidad y ambigüedad (Fu, 1994). En este sentido, según Sekuler (1994), cuando razonamos acerca de un sistema complejo, explicamos una aproximación de su comportamiento, y mantenemos un entendimiento general acerca del problema. En consecuencia, dicho autor propone que esta generalización con todo y ambigüedad, es suficiente para que comprendamos los sistemas complejos para aplicarlos a determinados conjuntos difusos, comúnmente denominados conjuntos borrosos.

Sin embargo, según los estudios propuestos por Chin-Teng (1995) la complejidad de un sistema decrece conforme aprendemos más acerca de él, incrementando al mismo tiempo nuestro entendimiento del mismo. Esta afirmación resulta cierta al analizar que al decrecer la complejidad se obtiene una mayor precisión en el modelado de un determinado sistema (Crespo, 2002).

La teoría de la lógica difusa como tal, fue inventada en 1965 por Lofti Zadeh guiado por el principio de que las matemáticas pueden ser usadas para encadenar el lenguaje con la inteligencia humana. A raíz de los resultados de Zadeh otros investigadores comenzaron a aplicar la lógica difusa en diversos rubros. En 1975, Mamdani, aplicó por primera vez la lógica difusa para controlar una máquina de vapor, que no podía ser controlada con técnicas convencionales, posteriormente en 1980 Smith aplica esta técnica al control de hornos rotativos en una cementera.

Algunos conceptos pueden ser mejor definidos con palabras, los conjuntos difusos ayudan a construir mejores modelos de la realidad (Rutkowska, 2002). La matemática de los conjuntos difusos trabaja con conjuntos que no tienen límites perfectamente definidos, es decir, la transición entre la pertenencia y no pertenencia de una variable a un conjunto es gradual (Jang et al., 1997).

La teoría difusa, se considera como una herramienta importante para poder controlar procesos industriales, así como para la electrónica de entretenimiento y hogar, aunque también se utiliza en sistemas de diagnóstico y otros sistemas expertos (Kosko, 1992; Martín del Brio y Sanz, 2001).

Si consideramos aplicaciones simples de lógica difusa (Klir y Folger, 1988) y según Zadeh (1965), la transición entre “pertenecer a un conjunto” y “no pertenecer a un conjunto” es gradual y esta transición suave es caracterizada por una función de pertenencia.

Los conjuntos definidos de forma imprecisa desempeñan un papel importante en el pensamiento humano, particularmente en los dominios del reconocimiento de patrones, de la comunicación de la información y de la abstracción. Sin embargo, recientemente Rutkowska (2002), define a un conjunto borroso como aquel razonamiento con lógica difusa que no es exacto sino en cierta forma imprecisa (función de membresía).

En Japón la investigación sobre lógica difusa es apoyada ampliamente con un presupuesto enorme. En Europa y Estados Unidos se están realizando esfuerzos para alcanzar al tremendo éxito japonés. Por ejemplo, la NASA emplea lógica difusa para el complejo proceso de maniobras de acoplamiento (Guijarro-Mata *et al.*, 2006, 2007). Recientemente se obtuvieron resultados muy satisfactorios al aplicar la metodología de los conjuntos borrosos al registro de avenidas máximas en México (Soto y Escalante, 1977).

Se ha empleado también en empresas de producción de artículos eléctricos y electrónicos como una herramienta de control, se ha utilizado para el desarrollo de procesadores y computadoras (Tanaka et al., 1982; Klir y Folger, 1988; Yo Chang y Ayyub, 2001). En general, la lógica difusa se aplica en cualquier campo donde sea muy difícil o casi imposible crear un modelo, en sistemas controlados por expertos humanos, en sistemas donde se tienen entradas y salidas que son continuas y complejas, en sistemas que utilizan observaciones humanas como entradas o reglas básicas, y en cualquier sistema en el cual se trabaje con conceptos imprecisos (Calviño, 2003).

3. METODOLOGÍA

El software a generar en la presente investigación está dividido en módulos según la disciplina a la que pertenezca, cada módulo tiene sus algoritmos particulares. La teoría y matemática a resolver para cada módulo se presenta en este capítulo de manera ordenada e independiente.

3.1. Red CIAQ

La arquitectura de Red CIAQ se encuentra implementada en dos sistemas de cómputo: Sistema de Adquisición Automatizado y Alertas (SAAA) y Sistema de Acceso a Reportes y Herramientas (SARH), dos bases de datos: Base de Datos Meteorológicos (BDM) y la Base de Datos del Sistema de Acceso a Reportes y Herramientas (BDSARH) y siete Estaciones Meteorológicas Automatizadas (EMA) Davis Vantage Pro2 distribuidas geográficamente en la zona metropolitana del estado de Querétaro. Los sistemas y bases de datos se encuentran distribuidos en 2 servidores de la siguiente manera: el servidor de aplicación contiene el SAAA y BAM y el Servidor Web contiene SARH y BDSARH. Ambos intercomunicados a través de red de área local (LAN) con acceso a internet. Estos servidores se encuentran de manera virtual en un servidor físico alojado en la Universidad Autónoma de Querétaro.

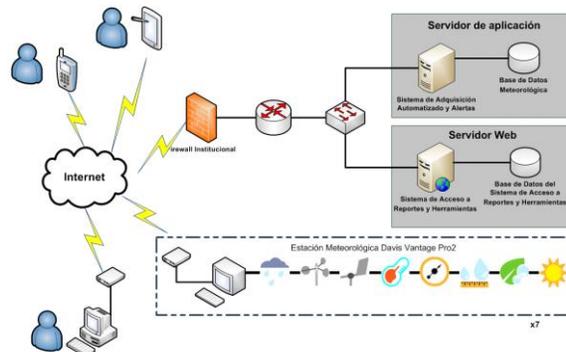


Figura 3-1. Arquitectura básica de Red CIAQ.

En la Figura 3-1 se representa la arquitectura previamente descrita, incluyendo el acceso a usuarios con distintos dispositivos vía internet a Red CIAQ.

3.1.1. Sistema de Adquisición Automatizado (SAA)

El sistema de Adquisición Automatizado y Alertas está desarrollado en lenguaje JAVA siguiendo el paradigma de desarrollo orientado a objetos y usando la técnica de hilos de programación con el objetivo que la conexión a las EMAs se realice en paralelo, de esta forma se reduce los el tiempo de adquisición a un lapso máximo de 8 segundos en total. En la imagen siguiente se muestra el diagrama de clases del SAA.

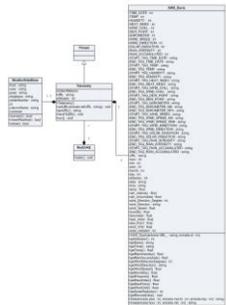


Figura 3-2. Diagrama de clases del SAAA.

El SAA se compone por la clase principal –Red CIAQ- la cual inicia el proceso de adquisición. El método *main* de dicha clase contiene un ciclo que realiza la creación de los objetos de tipo *Telemetry* dentro de un arreglo y ejecuta el método *start* de cada uno de estos para comenzar la consulta a la estación remota.

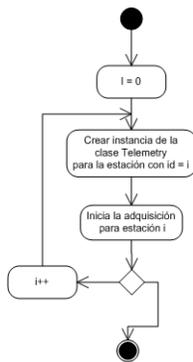


Figura 3-3. Diagrama de flujo principal.

La clase *Telemetry* es la responsable de administrar la conexión a la estación para la descarga de los datos meteorológicos y la base de datos donde se almacenará la información descargada. Contiene las instancias hacia la clase encargada de conectarse a la base de datos y una del tipo de estación a la cual se conectará. Esto es debido a que cada estación es diferente y esta arquitectura da la posibilidad de conectarse a distintas marcas y modelos de estaciones meteorológicas, lo cual da un valor agregado al SAA ya que se puede tener una arquitectura híbrida de EMAs y un mismo sistema de adquisición.

En este particular caso se utilizó estaciones de marca DAVIS por ese motivo, se programó una clase específica para esta marca. La clase *AWS_Davis* contiene los métodos los cuales descifran el código HTML, que es la forma la cual DAVIS nos entrega la información meteorológica. Contiene dicha clase los números de línea de código y se substraen el dato climático. Éste se formatea al tipo de dato que debe ser y se envía a la clase encargada de conectarse a la base de datos.

La clase *WeatherDataBase* es la encargada de realizar la conexión a la BDM, en ella se encuentran los datos básicos de conexión: usuario, *password* y *host*. Existe el método *insertWeather* que dedica a insertar a la BDM los datos recibidos desde la clase *AWS_Davis* en este caso.

3.1.2. Sistema de Acceso a Reportes y Herramientas (SARH)

El Sistema de Acceso a Reportes y Herramientas (SARH) consiste de un portal web donde se encuentran los reportes meteorológicos, herramientas hidrometeorológicos e información básica para que los usuarios aprendan más sobre los eventos climatológicos. El SARH está dividido en dos secciones principales: la sección de libre acceso y la sección de acceso restringido. La sección de libre acceso se dispone del reporte meteorológico a tiempo real, una imagen del satélite e información de los participantes y antecedentes del proyecto.

El reporte meteorológico está constituido por s cuales incluyen los datos de Temperatura, Precipitación Pluvial, Velocidad del viento, Dirección del viento, Radiación solar, Punto de rocío, humedad relativa, Presión atmosférica, índice de calor, Intensidad de lluvia y sensación térmica.

El reporte está constituido por 8 módulos básicos: Temperatura, Precipitación Pluvial, Velocidad del viento, Dirección del viento, Radiación solar, punto de rocío, humedad y Presión atmosférica.

Las 24 horas anteriores a la hora de acceso del usuario a consultar. Es decir, si el usuario consulta el reporte el día 10 de abril del 2013 a las 13:34h, el reporte le entregará:

- Base de Datos Meteorológicos (BDM)
- Base de Datos del Sistema de Acceso a Reportes y Herramientas (BDSARH)
- Estaciones Meteorológicas Automatizadas Davis Vantage Pro2
- Ubicación Geográfica de las Estaciones

3.2. Análisis de frecuencias

El análisis de frecuencias es un procedimiento para la estimación de eventos de diseño, asociados a un cierto período de retorno, basado en la observación de los eventos ya ocurridos. Aplicando técnicas de probabilidad y

estadística se selecciona una función de probabilidad que represente la relación del evento con su frecuencia observada (Izenman, 2008).

Haciendo referencia al campo de la ingeniería civil encontramos una aplicación por demás importante que es el análisis de frecuencia de eventos extremos es decir, se estima la probabilidad de ocurrencia de estos eventos asociados a un valor anual que puede ser máximo, mínimo o máximo proveniente de dos poblaciones. Eventos como la fuerza máxima del viento sobre las construcciones o el esfuerzo mínimo de ruptura de un material (Kinnison, 1985; Campos, 2008).

De forma más particular, podemos mencionar la aplicación del análisis de frecuencias a los eventos hidrológicos extremos (precipitación, escurrimiento, evaporación, temperatura, etcétera), con el objetivo de caracterizar los datos históricos de los eventos y asegurar la ocurrencia o no ocurrencia de dichos eventos en un futuro y de esta forma diseñar de manera óptima y confiable, obras hidráulicas o de aprovechamientos hidráulicos (Suárez, 2011).

El análisis de frecuencias es una herramienta fundamental para la predicción de eventos extremos. Cada espacio muestral tiene su propia función de distribución o de densidad de probabilidad que normalmente no se conoce a priori. Cuando de ese espacio se extrae un grupo de datos (muestra) al azar, es razonable esperar que su función de distribución de probabilidad sea similar a la del espacio completo, en particular si la muestra es grande (Esparza, 2005).

Parte importante en el análisis de frecuencias son los datos que se utilizan, debido a que deben cumplir con varias condiciones: ser completamente aleatorios, es decir, insesgados, independientes y homogéneos. Las mediciones y registros normalmente involucran errores humanos e instrumentales. Tales errores pueden ser de dos tipos, llamados errores accidentales y sistemáticos, aunque a veces es difícil distinguir entre ambos. Muchos errores son la combinación de los dos. Los errores accidentales son normalmente debidos al observador y en algunas

ocasiones debido a la naturaleza del instrumento de medición. Por otra parte, los errores sistemáticos pueden aparecer ya sea en el observador o en la medición y tales errores no son aleatorios (Chow, 1994).

Además de contar con datos confiables, es esencial considerar el número de años con registro, ya que, en algunas ocasiones, las series de tiempo son relativamente cortas y considerar estas para un análisis puede arrojar resultados erróneos. Por tal motivo, es importante conocer la longitud mínima de datos para obtener valores adecuados. En este sentido, algunos autores han realizado recomendaciones sobre la longitud mínima de registro de variables hidrológicas como se muestra a continuación.

- Ott (1972), recomienda usar como mínimo 20 años de longitud de registro.
- La IACWD (Interagency Advisory Committee on water Data), recomienda utilizar como mínimo 25 años de longitud de registro (IACWD, 1982).
- El USBR (United States Bureau of Reclamation), recomienda utilizar como mínimo 25 años de longitud de registro (USBR, 1981).
- El UKFSR (United Kingdom Flood Studies Report), recomienda utilizar como máximo periodo de retorno $2N$, donde N es la longitud del registro (UKFSR, 1975).
- Viessman recomienda utilizar como mínimo 10 años de longitud de registro para un periodo de retorno de $2N$, donde N es la longitud de registro (Viessman et al, 2002).

En el caso de México, Esparza (2005), realizó un estudio de la influencia de la longitud del registro en la estimación de gastos de diseño. Se abordaron diferentes escenarios (registros más antiguos, más recientes, máximos y mínimos de la estación), y los mismos escenarios considerando límites de confianza de 95%. Finalmente, se realizó una identificación de regiones a partir de los datos de las estaciones e interpolaron con el método IDW (Inverse Distance Weighting) cada una

de ellas. Con esto se puede determinar una longitud mínima de registro según la región donde se pretenda desarrollar un análisis de series de tiempo.

Funciones de distribución de probabilidad y metodologías más utilizadas en el análisis de frecuencias de variables hidrológicas.

Uno de los problemas más importantes en hidrología consiste en la interpretación de datos históricos de eventos hidrológicos en términos de probabilidades futuras de ocurrencia. Este problema aparece en los estimados de frecuencia de eventos, sequías, lluvias entre otros; y el proceso involucrado se conoce como análisis de frecuencias.

3.2.1. Distribuciones univariadas

Existen diversas funciones de distribución de probabilidad para el análisis de frecuencias. Sin embargo, los valores obtenidos no son iguales, por lo que es necesario escoger de esas funciones, la o las que se adapten mejor al problema bajo análisis. Comúnmente en hidrología, se han utilizado diferentes familias de distribución como son la familia Normal, Log-Normal, la familia Gumbel, Weibull y General de Valores extremos (GVE) y la familia de la Exponencial, Pearson, Log-Pearson tipo 3 (Maidment, 1992).

Sin duda, en diversas partes del mundo se han realizado estudios donde se aplica la teoría estadística al análisis de datos. Sin embargo, para series de tiempo de variables hidrológicas no todas son adecuadas debido a la naturaleza de estas. Por tal motivo, es importante conocer los estudios más relevantes desarrollados en esta área, además de las funciones de distribución y métodos de parámetros más adecuados, dependiendo de las características fisiográficas y climatológicas del lugar.

Por ejemplo, Daniel Gellens (2002) realizó un estudio para diferentes valores de precipitación diaria k ($k = 1, 2, 3, 4, 5, 7, 10, 15, 20, 25$ y 30) en la red climatológica de Bélgica con 165 estaciones, donde se evalúa la distribución general

de valores extremos (GVE). El método propuesto consistió en la combinación de valores regionales de los parámetros de la distribución GVE obtenidos a partir del método de momentos L.

En otro estudio realizado por Adlouni et al. (2010) se analizó la función de distribución Halphen que es poco conocida y la GVE que es ampliamente utilizada en hidrología. Se realizó una comparación con las distribuciones Gumbel y Fréchet donde las propiedades estadísticas de estas distribuciones indican que en algunos casos las distribuciones Halphen pueden presentar un mejor comportamiento que la distribución GVE, Gumbel y Fréchet. O bien, son distribuciones que pueden adaptarse a variables de valores extremos.

Cuando se considera que las distribuciones son limitadas, el análisis univariado parece proporcionar estimaciones razonables en valores extremos de duración menor a 10 horas, ya que los resultados muestran un comportamiento muy similar entre la Gumbel y Gamma. Sin embargo, cuando se utiliza la Gamma inversa, se obtienen mejores resultados respecto de la Gumbel (Kao, 2008). Por ejemplo, en Estados Unidos varias agencias federales han utilizado la distribución Log-Pearson tipo 3 (LP3) en estudios hidrológicos, explorado las características de la distribución en el espacio real y log-real. Además, de la relación entre estos, ya que en datos de gastos la distribución LP3 provee modelos razonables de la distribución de series de gasto anual de cuencas hidrográficas no reguladas (Griffis et al., 2007).

3.2.2. Distribuciones bivariadas

Por mucho tiempo se ha asumido que cualquiera de las variables tiene el mismo tipo de distribución marginal o que las variables tienen una distribución normal. Sin embargo, en la práctica, las variables de precipitación pueden tener diferentes distribuciones. Debido a esto, en estudios relacionados con los recursos hídricos se requiere de distribuciones de probabilidad conjunta de variables de precipitación. Tales como, la intensidad de la precipitación, lámina de lluvia y

duración (Zhang, 2007). Por tal motivo, se han desarrollado una serie de trabajos a partir de esta metodología, destacando algunos por su contribución a la ciencia.

Por ejemplo, Yue (2000a) aplica una distribución bivariada para representar la distribución conjunta de intensidad y lámina de lluvia máxima. En esta, utiliza la transformación de Box-Cox para normalizar las marginales originales. En otros estudios, Yue (1999, 2000b, c) aplica la distribución Log Normal bivariada, Gumbel Logística y Gumbel Mixta o Mezclada para el análisis de frecuencias multivariado de precipitación. Por su parte, Herr y Krzysztofowicz (2005) obtuvieron una forma genérica de la distribución de probabilidad bivariada utilizando la distribución meta-Gaussiana. En ésta se caracteriza la dependencia estocástica de la cantidad de precipitación entre dos estaciones. Además, se discute esta estimación a partir de la longitud de los datos.

Sin duda, una de las variables más relacionadas con los recursos hídricos, es el gasto. Por ejemplo, Carlos Escalante (2007) utilizó el modelo logístico para la distribución bivariada de Weibull o de valores extremos tipo 3 y la Weibull mixta de los marginales. Su aplicación se realizó en la región noreste de México con medición de 16 estaciones, obteniendo una mejora significativa a través del uso de una prueba de bondad de ajuste cuando se utiliza una distribución bivariada en lugar de una univariada. Los resultados muestran la importancia de tener en cuenta la distribución Weibull Mixta en el análisis bivariado cuando se analiza el gasto generado por dos poblaciones.

Otra metodología empleada en el análisis de frecuencia de gasto, es a partir de de cópulas. Aunque su aplicación en hidrología es corta y queda mucho por hacer, especialmente en el ámbito de la comprensión de las características físicas de estas, se han desarrollado muchos trabajos al respecto. Por ejemplo, Salvadori y De Michele (2004) presentan aspectos teóricos del análisis de frecuencia basado en cópulas. Además, mostraron varios casos de estudio para conocer el poder del análisis basado en cópula. Por su parte, Favre et al. (2004) desarrollaron una

metodología para la modelación de valores extremos usando cópulas, se utilizaron cópulas elípticas, arquimedeanas y con sección cuadrática. Estas fueron probadas en gasto pico y volúmenes de varias cuencas en Quebec, Canadá. En un caso muy similar de aplicación, De Michele et al. (2005) determinaron una distribución bivariada de valores extremos a partir de la distribución GVE para modelar gasto pico y volumen, usando el concepto cópula.

La aplicación de la distribución bivariada a partir del principio de cópulas, también se ha utilizado para evaluar la adecuación o cambios necesarios de un vertedero, ya que Grimaldi et al. (2005) utilizaron una cúpula de 3 dimensiones para calcular el hietograma de diseño y obtener los periodos de retorno que se necesitan para el estudio hidrológico. Además, las distribuciones obtenidas fueron probadas utilizando datos de gasto del río Amite en Baton Rouge, Luisiana, y el río Ashuapmushuan en Saguenay, Quebec, Canadá. Finalmente, se compararon con la Gumbel Doble y box-Cox-bivariada, encontrando que las distribuciones a partir de cópula obtienen mejores estimaciones de frecuencia basado en el trazado de la posición que con otras distribuciones (Zhang y Singh, 2006).

En otro estudio posterior Zhang y Singh (2007) obtuvieron la distribución de frecuencias bivariada de precipitación, donde se examinaron y compararon cuatro cúpulas de Arquímedes. Sin duda, este método presentó ventajas al no ser necesario suponer que las variables de lluvia son independientes, tienen comportamiento normal o presentan el mismo tipo de distribuciones marginales. Por lo tanto, las distribuciones bivariadas resultaron adecuadas para determinar los periodos de retorno de manera conjunta y condicional, a partir de datos de precipitación de la cuenca del río Amite en Luisiana, Estados Unidos. Por su parte Kao y Govindaraju (2007) realizaron un análisis bivariado a partir de datos de precipitación por hora a partir de cópulas empíricas. Los eventos de precipitación extrema fueron elegidos con base en tres criterios diferentes: volumen máximo anual (VMA), intensidad máxima anual (IMA), y la probabilidad máxima anual acumulada (PMA). Además, se analizaron características de precipitación, como

lámina total, duración e intensidad máxima para describir la dependencia entre las variables de precipitación y construir de manera conjunta eventos de lluvia extrema. Como resultado se tiene que para el análisis bivariado de eventos extremos es más adecuada la probabilidad máxima acumulada.

Debido a que es de gran importancia tomar en cuenta la dependencia de colas en el análisis de frecuencias bivariado a partir de cópulas Poulin et al. (2007) realizaron una comparación de tres estimadores no paramétricos de coeficientes de dependencia de cola a partir de simulaciones con siete familias de cópulas. Se eligieron los dos estimadores que más se adaptan al análisis de frecuencia bivariado de gasto máximo anual y el volumen de flujo correspondiente al hidrograma del río Loira (Francia). Los resultados muestran que cinco sobrestiman los periodos de retorno de eventos extremos. Estos resultados llevan a primer plano la importancia de tener en cuenta dependencia de cola con el fin de estimar el riesgo adecuadamente.

3.2.3. Estimación de parámetros

Una distribución de probabilidad es una función que representa la probabilidad de ocurrencia de una variable aleatoria. Mediante el ajuste a una distribución de un conjunto de datos hidrológicos, una gran cantidad de información probabilística en la muestra puede resumirse en forma compacta en la función y en sus parámetros asociados. El ajuste de las distribuciones puede llevarse a cabo por diferentes métodos dentro de los que destacan los de momentos o el de máxima verosimilitud (Chow, 1994). En la literatura, otros tantos métodos se han utilizado para la estimación de parámetros en diferentes distribuciones como son el de máxima entropía, momentos de probabilidad pesada, momentos L, momentos LH.

Uno de los principales métodos es el de momentos, desarrollado por primera vez por Karl Pearson en 1902, consideró que unos buenos estimativos de los parámetros de una función de probabilidad son aquellos para los cuales los momentos de la función de densidad de probabilidad alrededor del origen son

iguales a los momentos correspondientes de la información de la muestra (Chow, 1994)

En la literatura, algunos de los métodos se han utilizado para la estimación de parámetros en diferentes distribuciones, por ejemplo, en la distribución Log Pearson tipo 3 se han aplicado los métodos de momentos (MOM), máxima verosimilitud (MLE) y el método de momentos mixtos (MXM). El método MXM presenta buenos resultados en comparación con MOM y MLE cuando se tiene ausencia de información regional. Sin embargo, un análisis de Monte Carlo demuestra que las estimaciones con MOM de registro de MOM con información regional sesgada es más recomendado, y cuanto más precisa sea el sesgo de la información regional se tendrán estimaciones de flujo más precisas (Griffis et al., 2007).

El método de momentos L es un sistema alternativo al método tradicional de momentos para describir las formas de las funciones de distribución (Hosking, 1990) y surgen como combinaciones lineales de los momentos ponderados probabilísticamente (MPP) introducidos por Greenwood et al. (1979). Los momentos ponderados probabilísticamente de la variable aleatoria x con una función de distribución $F(x)$ son utilizados para estimar los parámetros de las distribuciones que pueden ser expresadas en forma inversa $x = x(F)$ explícitamente (Álvarez et al., 1999). Este método puede ser más útil que los momentos convencionales para discriminar entre diferentes distribuciones, porque la estimación de ubicación, escala y forma son casi imparcial, independientemente de la distribución de probabilidad de los datos que surgen (Hosking y Wallis, 1997; Saf 2009).

Una variedad de estudios basados en el método de momentos L han sido empleadas en el análisis de frecuencia regional de precipitaciones extremas (por ejemplo, Sveinsson et al., 2002; Burgueño et al., 2005; Endreny y Pashiardis, 2007; Gaal y Kysely, 2009). En el análisis regional, el método momentos L se puede aplicar en cuatro etapas del análisis de frecuencia regional, los cuales son: (i) el

análisis de los datos, (ii) la identificación de regiones homogéneas, (iii) la elección de una distribución de frecuencias, y (iv) la estimación de la frecuencia en la distribución (Kysely, 2006). Otros parámetros también se pueden calcular por el procedimiento de la probabilidad como un estudio reciente (Campos, 2008; Ellouze Abida, 2008; Escalante, 2009).

Wang (1997) desarrolló el método de momentos LH como una generalización de los momentos L. Demostró que el uso de esta técnica reduce influencias indeseables que los eventos pequeños de la muestra pueden tener sobre la estimación de periodo de retorno de grandes eventos en comparación a la utilización de momentos L. El estudio de Wang (1997) informó de la mejora del rendimiento de momentos LH en comparación con los momentos L, pero se observa que existe limitación al utilizar los momentos LH en el análisis de frecuencia de crecidas. Esto se debe a la dificultad de la formación de la región homogénea y el desarrollo para otras distribuciones de uso común en el análisis de frecuencia de crecidas. Meshgi y Khalili (2007a, b) realizaron el análisis de frecuencia de crecidas de las cuencas hidrográficas regionales Kharkhe, ubicado en el oeste de Irán, utilizando el método de momentos LH. En su estudio, utilizaron distribuciones GEV, GLO del PAM. También Bhuyan et al. (2010) utilizaron esta técnica para el análisis regional de frecuencia de la misma región, así como una comparación entre los los métodos de momentos.

Por su parte Ashkar (2006) investigó los parámetros en el método de momentos generalizado (GM) y la estimación de cuantiles en el modelo log-logístico de dos parámetros (LL2). Los resultados de la simulación muestran que el método de GM supera las estimaciones obtenidas por otros. También se demuestra que una mezcla de momentos de orden positivo y negativo que se necesita para la estimación óptima en un modelo LL2, puede implementar mediante el método de GM.

Otro método ampliamente utilizado en la estimación de parámetros es el de máxima verosimilitud. Sin embargo, la estimación de la distribución mezclada de Weibull por el método de máxima verosimilitud y otros métodos a menudo es difícil debido a las estimaciones inestables derivadas de datos limitados. Las técnicas bayesianas pueden estabilizar estas estimaciones a través de los priores, pero no hay familia conjugada de forma cerrada para la distribución de Weibull. Este estudio ha reducido de 5 a 2 el número de integraciones numéricas necesarias para el uso de la estimación bayesiana de la distribución mixta de Weibull, por lo que es un enfoque más viable para el usuario (Touw, 2009). Otra modalidad de este método es la función logarítmica de verosimilitud es la mostrada en el trabajo realizado por Gutiérrez y Ramírez (2005). Esta fue utilizada para el análisis del método de la avenida índice a partir de la función de distribución Gumbel Doble de gastos máximos.

3.2.4. Análisis de frecuencias a partir de distribuciones de probabilidad univariadas

Una distribución de probabilidad es una función que representa la probabilidad de ocurrencia de una variable aleatoria. Mediante el ajuste a una distribución de un conjunto de datos hidrológicos, una gran cantidad de información probabilística en la muestra puede resumirse en forma compacta en la función. Por lo tanto, la selección y aplicación de diferentes distribuciones de probabilidad univariadas es de vital importancia para un análisis adecuado. Debido a esto, en este estudio se aplican la distribución Normal, Log-Normal, Exponencial, Gamma y Gumbel que se han ampliamente para ajuste de variables hidrológicas.

Con el fin de conocer un poco de estas distribuciones, se hace una breve descripción de las funciones de distribución aplicadas en el estudio. Además, se presentan algunas de sus características y ecuaciones.

3.2.5. Distribución Normal

Su gráfica, la curva normal, describe en forma aproximada muchos fenómenos en la naturaleza de mediciones físicas en diversas áreas. Esta ecuación

fue desarrollada en 1733 por Abraham De Moivre. Sin embargo, Karl Frirdriek Gauss también derivó su ecuación a partir de un estudio de errores en mediciones repetidas de la misma cantidad (Walpone y Myers, 1992; Kreiszig, 1973).

La distribución normal surge del teorema del límite central, el cual establece que si una secuencia de variables aleatorias x_i son independientes y están idénticamente distribuidas con media μ y varianza σ^2 , entonces la distribución de la suma de n de estas variables aleatorias, $Y = \sum_{i=1}^n x_i$, tiende hacia la distribución normal con media μ_n y varianza σ_n^2 a medida que n aumenta. El punto importante es que esto es cierto sin importar cuál es la función de distribución de probabilidad de x (Chow et al., 1994).

La variable aleatoria continua que denota la distribución en forma de campana se le conoce como variable aleatoria normal. Las variables hidrológicas, como la precipitación anual, calculada como la suma de los efectos de muchos eventos independientes tienden a seguir la distribución normal, aunque su desventaja cuando se trata de describir variables de este tipo es que ésta varía en un rango continuo $(-\infty, \infty)$, mientras que la mayoría de las variables de tipo hidrológico no son negativas, además de que es simétrica alrededor de la media, mientras que la información hidrológica tiende a ser asimétrica (Chow, 1994).

Su función de densidad de probabilidad se muestra en la ecuación 3.1.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.1)$$

donde x representa la variable, μ es la media de la muestra y σ^2 la varianza. Además, su función de distribución de probabilidad para la distribución normal es:

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad (3.2)$$

3.2.6. Distribución Log-Normal

Si la variable aleatoria $Y = \log x$ está normalmente distribuida, entonces se dice que x está distribuida en forma Lognormal. Chow (1954) llegó a la conclusión de que esta distribución se aplica a variables hidrológicas formadas como productos de otras variables. La distribución log normal tiene las ventajas sobre la distribución normal de que está limitada ($X > 0$) y de que la transformación Log tiende a reducir la asimetría positiva comúnmente encontrada en información hidrológica, debido a que al tomar logaritmos se reducen en una proporción mayor los números grandes que los números pequeños. Algunas de las limitantes de esta distribución son, por un lado, que tiene solamente dos parámetros y, por otro lado, que requiere que los logaritmos de los datos sean simétricos alrededor de su media (Chow et al., 1994).

Su función de densidad de probabilidad se muestra en la ecuación 3.3.

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log x - \mu_y)^2}{2\sigma_y^2}} \quad (3.3)$$

donde x representa la variable, μ es la media de la muestra y σ^2 la varianza.

La función de distribución de probabilidad para la distribución normal es:

$$F(x) = \int_0^x \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{x\beta} e^{-\frac{1}{2}\left(\frac{\ln x - \alpha}{\beta}\right)^2} dx \quad (3.4)$$

3.2.7. Distribución Exponencial

Algunas secuencias de eventos hidrológicos, como la ocurrencia de precipitación, pueden considerarse como procesos de Poisson, en los cuales los eventos ocurren instantánea e independientemente en el tiempo, o a lo largo de una línea. El tiempo entre tales eventos, o tiempo de interarribo, está descrito por una

distribución exponencial cuyo parámetro λ es la tasa media de ocurrencia de los eventos. En la distribución exponencial es fácil estimar λ a partir de la función observada (Chow et al., 1994).

Esta distribución también se conoce como distribución exponencial negativa. El tiempo de ocurrencia de fenómenos como la precipitación, pueden considerarse como Procesos de Poisson ya que los eventos ocurren instantánea e independientemente en un horizonte de tiempo, a lo largo de la línea. El tiempo entre dichos eventos lo describe la distribución exponencial cuyo parámetro λ es la tasa media de ocurrencia de los eventos. La ventaja de la distribución exponencial radica en que es fácil estimar a partir de la información observada y que la distribución exponencial se adapta muy bien a estudios teóricos. Su desventaja es que se requiere que la ocurrencia de cada evento sea completamente independiente (Chow et al., 1988).

Su función de densidad de probabilidad se muestra en la ecuación siguiente.

$$f(x) = \lambda e^{-\lambda x} \quad (3.5)$$

donde x representa la variable, λ representa la ocurrencia media de eventos.

La función de distribución de probabilidad para la distribución normal es

$$F(x) = 1 - e^{-\lambda x} \quad (3.6)$$

3.2.8. Distribución Gamma

El tiempo que toma la ocurrencia de un número β de eventos en un proceso de Poisson está descrito por la distribución gamma, la cual es la distribución de una suma de β variables aleatorias independientes e idénticas, distribuidas exponencialmente. La distribución gamma tiene una forma que varía suavemente y es muy útil para la descripción de variables hidrológicas asimétricas sin el uso de la

transformación log (Chow et al., 1994). Esta distribución es de gran importancia en la teoría de colas y en los problemas de confiabilidad (Walpole y Myers, 1992). Sin embargo, la distribución gamma de dos parámetros tiene como límite inferior cero, lo cual es una desventaja para la aplicación a variables hidrológicas con un límite inferior que es superior a cero (Chow et al., 1994).

$$f(x) = \frac{\lambda^\beta x^{\beta-1}}{\Gamma(\beta)} e^{-\lambda x} \quad (3.7)$$

donde x representa la variable, λ es un parámetro de escala, β es un parámetro de forma y $\Gamma(\beta)$ es la función matemática gamma.

La función de distribución de probabilidad para la función gamma es

$$F(x) = \frac{1}{\alpha \Gamma(\beta)} \int_{\delta}^x \left(\frac{x-\delta}{\alpha}\right)^{\beta-1} e^{-\frac{x-\delta}{\alpha}} dx \quad (3.8)$$

donde α , β y δ son los parámetros. Además $\Gamma(\beta)$ es la función matemática gamma.

3.2.9. Distribución Gumbel

La distribución de Valores Extremos Tipo I o Distribución Gumbel nace con Fisher y Tippet, en los años veinte con la teoría de Valores Extremos. En los años treinta, Gumbel realiza aplicaciones prácticas usando los estadísticos de valores extremos de distribuciones del tiempo de duración de la vida humana y en 1941 publica algunos artículos con respecto a sus aplicaciones en el análisis de frecuencias de gastos máximos y mínimos. Jenkinson (1955) demostró que esta es un caso especial de la Distribución General de valores Extremos (Ven Te Chow et al., 1996). Donde su función de densidad de probabilidad es:

$$f(x) = \frac{1}{\beta} e^{-\frac{x-\alpha}{\beta}} e^{-\frac{x-\alpha}{\beta}} \quad (3.9)$$

Su función de distribución acumulada es:

$$F(x) = e^{-e^{\frac{x-\alpha}{\beta}}} \quad (3.10)$$

donde x es la variable, α y β son los parámetros de la función.

3.2.10. Metodología propuesta seleccionando dos distribuciones univariadas diferentes.

3.2.10.1. *Ajuste de datos*

Una vez que se tienen los registros de precipitación máxima anual, se procede al ajuste de los datos a una función de distribución. Para esto se sigue la secuencia de cálculo expuesta en sus trabajos originales de Gumbel (1954 y 1956) donde propone desarrollarla como se muestra a continuación.

- Ordenamiento de datos.
- Determinación de periodo de retorno.
- Valores de variable reducida.
- Valor de probabilidad.
- Ajuste a función de distribución y graficado de datos.

3.2.10.2. *Ordenamiento de los datos*

Al contar con la serie de tiempo de la estación meteorológica, el primer paso consiste en ordenar los datos, es decir, los valores de la variable aleatoria x , en orden creciente debido a que se trata de valores máximos anuales. Al cumplir con este requerimiento, implica alterar el orden cronológico del registro. Inmediatamente se asignó a cada dato un número de orden i , teniendo en cuenta que el número 1 corresponde al suceso menos probable, es decir, la secuencia resulta ser la siguiente:

$$x_1 > x_2 > \dots > x_i > \dots x_n \quad (3.11)$$

3.2.10.3. *Determinación de periodo de retorno*

En hidrología, normalmente se prefiere trabajar con periodo de retorno en lugar de probabilidades, pues es un concepto que resulta más claro ya que tiene las mismas unidades (tiempo) que la vida útil de las obras. Por lo tanto, el siguiente paso consiste en asignar a cada dato un valor de periodo de retorno o intervalo de recurrencia (Ecuación 3.12), que representa el número de años en que un evento puede ser igualado o excedido.

$$Tr = \frac{n+1}{m} \quad (3.12)$$

donde n y m son el número de datos y número de orden, respectivamente.

3.2.10.4. *Valores de la variable reducida*

Como alternativa, es conveniente el graficado de los puntos a partir de la variable reducida. A continuación, se presenta la formulación de la variable reducida a partir de la distribución Gumbel en términos del periodo de retorno.

Una variable reducida y puede definirse como:

$$y = \frac{x-a}{b} \quad (3.13)$$

Sustituyendo la variable reducida en la ecuación 3.10

$$F(x) = e^{-e^y} \quad (3.14)$$

Resolviendo para y :

$$y = -\ln \left[\ln \left(\frac{1}{F(x)} \right) \right] \quad (3.15)$$

Luego

$$F(x) = \frac{Tr-1}{Tr} \quad (3.16)$$

Luego,

$$y = -\ln \left[\ln \left(\frac{Tr}{Tr-1} \right) \right] \quad (3.17)$$

ó

$$y = -\ln \left[-\ln \left(\frac{Tr-1}{Tr} \right) \right] \quad (3.18)$$

A partir de la ecuación final se obtiene cada uno de los valores de variable reducida y que serán graficados posteriormente.

3.2.10.5. Valor de probabilidad

Cuando la variable aleatoria considerada es una magnitud relacionada con algún fenómeno natural (caudales, precipitación, velocidad de viento, etc.), es conveniente referirse a periodos de retorno en lugar de probabilidades de ocurrencia. Si p es la probabilidad que una variable x supere un dado valor X en un cierto lapso (por lo general un año), el periodo de retorno Tr representará el número de unidades de tiempo que se transcurrirán en promedio entre dos oportunidades en que la variable supere dicho valor, es decir:

$$p = P(x \geq X) = \frac{1}{Tr} \quad (3.19)$$

Por lo tanto, es equivalente especificar un periodo de retorno o recurrencia de 100 años o una probabilidad anual de 0.01.

3.2.10.6. Ajuste a función de distribución

Un punto importante de este análisis es la determinación del ajuste más adecuado a la serie de tiempo. Por tal razón, se hace uso del criterio de error mínimo estándar de ajuste SEF (por sus siglas en inglés), definido por Kite (1998) como se muestra en la ecuación 3.20. En este criterio se realiza una comparación entre los

valores reales (medidos) y los valores calculados por alguna función de distribución definida.

$$SEF_J = [\sum_{i=1}^n (g_i - h_i)^2 / (n - q)]^{1/2} \quad (3.20)$$

donde $g_i, i = 1, \dots, n$ son los eventos registrados, $h_i, i = 1, \dots, n$ son las magnitudes calculadas de los eventos a partir de la distribución de probabilidad de las probabilidades obtenidas de las filas ordenadas de $g_i, i = 1, \dots, N$, q es el número de parámetros estimados para la distribución marginal, n es la longitud de registro, y J es el número de la estación analizada.

3.3. Curvas IDT

La intensidad de la lluvia se define como la cantidad de agua de lluvia que cae en un punto, por unidad de tiempo y ésta es inversamente proporcional a la duración de la tormenta.

La duración de la lluvia es el tiempo que transcurre desde que inicia la precipitación de la tormenta hasta que ésta cesa. De acuerdo con Smith (1993) el análisis de frecuencia de la precipitación es usado ampliamente para diseño de obras de ingeniería que controlan los escurrimientos generados por las tormentas. Además, incluye los sistemas de alcantarillado municipales, vertedores, puentes y sistemas de drenaje agrícola. Willems (2000) menciona también que la relación Intensidad-Duración-Frecuencia es ampliamente usada para estimar las avenidas de diseño en los sitios donde se construirán las obras hidráulicas.

El período de retorno está íntimamente ligado a la importancia de la obra. El periodo de retorno (Tr) es el número de años que en promedio y a la larga un evento puede ser igualado o excedido. Para n datos anuales de una muestra el periodo de retorno se puede calcular con la fórmula de Weibull (Bedient y Hubert, 2002).

$$Tr = \frac{n+1}{m} \quad (3.21)$$

Donde Tr es el periodo de retorno en años; n el número total de datos de la muestra a analizar; y m el valor de rango de cada valor. El valor de rango se obtiene de los datos de altura de precipitación o de intensidad que han sido acomodados de mayor a menor, si es análisis de máximos, o de menor a mayor, si es de mínimos de mayor a menor.

Las Curvas IDT son curvas que resultan de unir los puntos representativos de la intensidad media en intervalos de diferente duración, y correspondientes todos ellos a una misma frecuencia o período de retorno (Témez, 1978).

Las Curvas IDT resultan ser la representación gráfica de la relación existente entre la intensidad, la duración y la frecuencia o período de retorno de la precipitación (Benítez, 2002).

Willems (2000) define a las Curvas Intensidad-Duración-Periodo de retorno (IDT) como la relación que existe entre la intensidad de lluvia media, su duración y la frecuencia de ocurrencia (inverso del período de retorno); estas curvas son herramientas ampliamente utilizadas en la ingeniería hidrológica para fines de planeación, diseño, revisión y operación de los proyectos hidráulicos, así como para la protección de obras de ingeniería contra avenidas máximas (Koutsoyiannis et al, 1998).

Considerando la relación duración-intensidad dentro del margen de una precipitación, podemos establecer que las tormentas de mayor intensidad no son necesariamente las más frecuentes en áreas con una alta pluviometría anual, las tormentas de gran intensidad suelen cubrir poco terreno y su duración generalmente es corta y las tormentas que cubren grandes zonas son raramente de gran intensidad, pero pueden durar varios días. Ahora bien, el planteamiento de las primeras relaciones matemáticas sobre las Curvas IDT data desde el inicio de la

década de los treinta, del siglo pasado, con los desarrollos y aplicaciones de Sherman (1931) y Bernard (1932).

Desde entonces varios conjuntos de relaciones han sido desarrollados para varias cuencas hidrológicas del mundo. Las curvas de IDT también pueden expresarse como ecuaciones, con el fin de evitar la lectura de la intensidad de la lluvia de diseño, en una gráfica (Chow et al., 1964).

Una expresión general para la intensidad de lluvia sobre una zona sería la ecuación de Sherman (1931) propuesta para la ciudad de Boston, Massachusetts, USA (Raudkivi, 1979); la intensidad de la tormenta se reduce conforme aumenta la duración del evento.

$$i = \frac{\lambda + Tr^\psi}{(d + \theta)^\eta} \quad (3.22)$$

Donde i es la intensidad de la lluvia en mm/h; Tr es el periodo de retorno en años; d es la duración de la tormenta en min; y λ , ψ y η son parámetros que se determinan al ajustar los datos.

La tormenta es aquel evento que se presenta en un punto y para definirla se utilizan los datos puntuales de altura de precipitación o intensidades máximas de lluvia registradas en una estación. Los métodos más comunes que se usan para definir su magnitud son: el análisis simultáneo de las variables de intensidad, duración y periodo de retorno (IDT), a través de una regresión múltiple; y el análisis independiente de las variables de intensidad y periodo de retorno (IDT) a partir de funciones de distribución de probabilidad (National Weather Service 2012).

3.4. Geoestadística

La geoestadística es la ciencia que se encarga estudiar la información espacial. Matheron (1971), fue el pionero en esta área, la define como “la aplicación del formalismo de las funciones aleatorias al reconocimiento y estimación de fenómenos naturales”.

Woodcock (1988a), analizó imágenes simuladas con la idea de modelar su variación espacial en función de los parámetros físicos de la escena y del sensor utilizado. Encontró que la herramienta idónea para esta tarea era el variograma y que sus características estaban asociadas a las de los objetos en la escena. La altura de la meseta del variograma estaba relacionada con la densidad de los objetos y el rango con su tamaño. Al incrementar la varianza del tamaño de los objetos la forma del variograma adoptaba una proporción más redondeada en la zona cercana a la meseta. El efecto de aumentar el tamaño del pixel se reflejaba en una disminución de la altura de la meseta y un incremento del valor del rango.

El método de interpolación por Krigage o “kriging” es definido como un método de interpolación por medias ponderadas donde el conjunto de los pesos asignados a los puntos de muestreo minimiza la varianza de estimación, que es calculada en función de un modelo de variograma y de los emplazamientos relativos de los puntos, unos con respecto a otros, así como con relación al punto de estimación (Banerjee et al., 2004).

El método reposa sobre el principio que dice que la variación espacial de una propiedad (lo que se llama una variable regionalizada) es muy irregular para ser modelada por una simple función matemática, pero que puede ser descrita por una superficie estocástica. La interpolación debe pues, primeramente, explorar y seguidamente modelar el aspecto estocástico de la variable regionalizada. La información que resulta es entonces utilizada para estimar los pesos de la función de interpolación.

La teoría de las variables regionalizadas (Matheron, 1971) asume que la variación espacial de una variable cualquiera puede ser obtenida por la adición de tres componentes mayores:

- a) la componente estructural, expresada con un valor medio constante o con una tendencia constante

- b) una componente aleatoria espacialmente correlacionada
- c) un ruido aleatorio o error residual.

Los componentes b) y c) se presumen constantes en toda la zona de muestreo.

Sea x la posición en 1, 2 ó 3 dimensiones, el valor de la variable regionalizada Z en x es:

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon'' \quad (3.23)$$

donde $m(x)$ es una función determinística que describe la componente estructural de Z en la posición x , $\varepsilon'(x)$ es el término que denota los residuos de $m(x)$ que son de naturaleza estocástica, localmente variables y espacialmente correlacionadas y ε'' es un ruido residual, gaussiano, espacialmente independiente teniendo por media cero y varianza s^2 .

En general, el Kriging predice valores en puntos y bloques en espacios d -dimensionales y permite la incorporación de la anisotropía. Varias extensiones al método permiten una gran flexibilidad para su aplicación (Gutiérrez, 2005).

El Cokriging es una técnica de interpolación que permite una mejora en la estimación de los valores si la distribución de una segunda variable es mejor conocida. En el caso de que la variable de estudio sea difícil o costosa en su medición, entonces el cokriging permite mejorar la interpolación, introduciendo el semivariograma de la segunda variable en el proceso.

Las principales cualidades del Kriging están en la calidad estadística de sus predicciones y en la habilidad de predecir la distribución espacial de la incertidumbre (Matheron, 1963).

3.4.1. Isoyetas

Una isolínea (también llamada, isopleta, curva de nivel, isógrama o isaritma), para una función de varias variables, es una curva que conecta los puntos en que la función tiene un mismo valor constante (Stein et. al. 2002). En esta plataforma se desarrollan algoritmos para el trazo de isoyetas (isolíneas de precipitación) que generan un módulo completo de la plataforma.

Las isolíneas que se representan en un mapa son líneas, rectas o curvas, que describen la intersección de una superficie real o hipotética con uno o más planos horizontales. La configuración de estas curvas permite a los lectores del mapa inferir el gradiente relativo de la variable o parámetro y estimar un valor en un lugar determinado.

El gradiente de la función es siempre perpendicular a la isolínea. Cuando las líneas están muy cerca, la longitud del gradiente es grande: la variación es acentuada. Si las isolíneas adyacentes tienen el mismo grosor de línea, la dirección del gradiente no puede determinarse y por ello se emplean diferentes grosores o se rotulan o etiquetan numéricamente: de este modo la dirección del gradiente puede ser fácilmente apreciada.

El uso más habitual de las isolíneas es en cartografía y en meteorología. Un mapa topográfico (o mapa de curvas de nivel) utiliza isolíneas que unen puntos de igual altitud y muestra, así, la forma de los valles y las colinas, y la pendiente de las laderas (Zimmerman et al., 1998).

La idea de utilizar líneas que unen puntos de igual valor fue redescubierta en varias ocasiones a lo largo de la historia. En 1701, Edmond Halley utilizó ese tipo de líneas (isogonas) en un gráfico que recogía la variación magnética. En 1727, el ingeniero holandés Nicholas Cruquius dibujó el fondo del río Merwede con líneas de igual profundidad (isóbatas), en intervalos de 1 braza; y en 1737, Philippe Buache, utilizó un intervalo de 10 brazas para elaborar un mapa del Canal de la Mancha, que fue publicado en 1752.

El uso de estas líneas (la curva de nivel) para describir la superficie de la tierra fue estudiado teóricamente por Ducarla en 1771, y Charles Hutton lo utilizó en el cálculo del volumen de una colina en 1777. En 1791, J. L. Dupain-Triel realizó un mapa de Francia, en que reflejó las curvas de nivel cada 20 metros de intervalo. En 1801, el también francés y jefe del Cuerpo de Ingenieros, Haxo, utilizó la línea de nivel a mayor escala, 1:500 en una planta para su proyecto para la Rocca de Aufo (Skelton, 1958).

En torno a 1843, cuando en Gran Bretaña e Irlanda la Ordnance Survey comenzó a recoger en sus mapas con regularidad las líneas de nivel, éstas ya eran usadas con generalidad en los países europeos. Las isóbatas no se generalizaron en las cartas náuticas hasta que fueron publicadas en 1834 las de Rusia y en 1838 las de Gran Bretaña (Close, 1926).

Cuando los mapas con curvas de nivel o isolíneas se hicieron habituales, la idea de utilizar este tipo de representación se propagó a todo tipo de aplicaciones. Tal vez los de más reciente desarrollo son los mapas de curvas de calidad del aire y contaminación acústica, que aparecieron por primera vez en los EE. UU., aproximadamente hacia 1970, en gran parte como resultado de la legislación nacional que requería la delimitación espacial de esos parámetros.

El uso de nombres específicos está muy extendido en Meteorología, en que se usan en un mismo mapa o diagrama (generalmente, isobaras e isotermas) para presentar una imagen de los principales factores que afectan al tiempo. Las isolíneas empleadas en Meteorológica se basan en la generalización de los datos recibidos desde la estación meteorológica. Las estaciones rara vez se colocan exactamente en una isolínea (cuando lo están, indica una medición precisa igual al valor de la isolínea). En lugar de ello, las líneas son dibujadas con la mejor aproximación de los valores exactos, basados en la dispersión de la información disponible. Los mapas del tiempo pueden presentar datos recogidos, como la presión de aire en un momento dado, o una extrapolación de esos datos, como la

presión media durante un período, o los pronósticos de tiempo, como la predicción de la presión de aire en algún momento futuro. En este capítulo lo más importante es hacer una descripción corta de los resultados, haciendo uso de cuadros donde se agrupe la información obtenida. La zona a estudiar se delimita mediante las estaciones de la Red CIAQ, ver Figura 3-4.



Figura 3-4. Zona de estudio Red CIAQ.

3.5. Series de tiempo

Una serie de tiempo es una secuencia de observaciones, medidos en determinados momentos del tiempo, ordenados cronológicamente y, espaciados entre sí de manera uniforme, así los datos usualmente son dependientes entre sí (Box y Jenkins, 1970). Los modelos de series de tiempo se construyen bajo la premisa que éstas tienen una historia estadística recurrente particular que se puede modelar y explotar para fines de pronóstico. Ésta técnica fue sistematizada por Box y Jenkins a fines de la década de los 60s.

De acuerdo a Box y Jenkins (1970), las etapas del análisis de series de tiempo son:

1. Identificación,
2. Estimación,
3. Verificación y,
4. Pronóstico (Predicción)

Si la serie es débilmente estacionaria, se procede a la etapa 1, de lo contrario se requiere procesar los datos a fin de ser transformada en realizaciones estacionarias. La etapa de identificación tiene el objetivo de determinar el tipo de modelo a aplicar y el orden del parámetro “p”.

Los modelos autorregresivos de orden p , $AR(p)$ consideran una media cero cuando

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \varepsilon_t \quad (3.24)$$

ó

$$\varphi_p(L)(X_t) = \varepsilon_t \quad (3.25)$$

donde X_t el valor actual de la serie que se explica en función de p valores pasados $X_{t-1}, X_{t-2}, \dots, X_{t-p}$, donde p determina el número de rezagos necesarios para pronosticar el valor actual; L es el operador de rezagos; $\varepsilon_t \sim RB(0, \sigma^2)$, es decir, un proceso de ruido blanco con media 0 y varianza constante σ^2 ; y φ_p es el polinomio autorregresivo de orden p descrito como $\varphi_p = 1 - \varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p$, donde $z \in \mathbb{C}$ y φ_i son los parámetros del modelo.

3.5.1. Proceso Autorregresivo de Orden 1: $AR(1)$

En los procesos $AR(1)$ la variable X_t está determinado únicamente por el valor pasado X_{t-1} . Por lo tanto, se utiliza la ecuación siguiente.

$$X_t = \varphi X_{t-1} + \varepsilon_t \quad (3.26)$$

donde ε_t es un proceso de ruido blanco con media 0 y varianza constante σ^2 , φ es el parámetro del modelo. Suponemos que el futuro no incluye en el pasado, es decir, el proceso es no anticipante.

Por lo que un modelo $AR(1)$ es una versión restringida de un modelo general de medias móviles. Entre sus características se encuentran que éste es siempre invertible, estacionario siempre que se cumpla $|\varphi| < 1$ y su representación gráfica de la función de autocorrelación, el correlograma, tiene un comportamiento amortiguado hacia cero con todos los valores positivos, en caso de que $\varphi > 0$, o bien alternando el signo, comenzando con negativo si $\varphi < 0$ (Giraldo Gómez, 2006).

3.5.2. Proceso Autorregresivo de Orden 2: $AR(2)$

En los procesos $AR(2)$ la variable X_t está determinado por el valor pasado X_{t-1} y el anterior a éste. Por lo tanto, se utiliza la ecuación siguiente.

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varepsilon_t \quad (3.27)$$

donde ε_t es un proceso de ruido blanco con media 0 y varianza constante σ^2 , φ_1 y φ_2 son los parámetros del modelo. (Giraldo Gómez, 2006)

3.5.3. Proceso Autorregresivo de Orden 3: $AR(3)$

Homólogamente, los procesos $AR(3)$ la variable X_t está determinado por el valor tres valores pasados. Por lo tanto, se utiliza la ecuación siguiente.

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varphi_3 X_{t-3} + \varepsilon_t \quad (3.28)$$

donde ε_t es un proceso de ruido blanco con media 0 y varianza constante σ^2 , φ_1 , φ_2 y φ_3 son los parámetros del modelo. (Giraldo Gómez, 2006)

3.5.4. Criterio de selección de Akaike

En series de tiempo, se selecciona el modelo a utilizar mediante un criterio de selección. El más utilizado es el criterio de selección de Akaike (AIC). Debido a que es un estimador insesgado asintótico, entre un modelo candidato ajustado y el

verdadero modelo. Éste considera el número de rezagos de cada una de las variables y castiga los coeficientes extras (los correspondientes a las variables rezagadas adicionales) por lo que, la especificación elegida corresponderá a aquél con el criterio más pequeño. El criterio de selección de Akaike se define como:

$$AIC(M) = -2 \log(L) + 2M \quad (3.29)$$

donde L se corresponde a la función de verosimilitud y M es el número de parámetros del modelo (Akaike, 1974; Urquiola et al., 2010).

3.6. Desarrollo de la interfaz gráfica de usuario MATLAB

GUIDE proporciona un conjunto de herramientas para crear una interfaz gráfica de usuario (GUI – Graphical User Interface en inglés) en MATLAB. Estas herramientas simplifican el proceso de diseño y creación de la interfaz gráfica. En la ventana gráfica se pueden disponer paneles, botones, cuadros de texto, barras de desplazamiento, etc. Te permite generar automáticamente, dos archivos uno con extensión .fig que contiene la información sobre el aspecto visual de la interfaz y otro archivo .m en el que se codifica la respuesta a las acciones del usuario sobre los controles.

1. Para iniciar GUIDE se escribe la instrucción `guide` en la ventana de comandos de MATLAB, ver Figura 3-5.

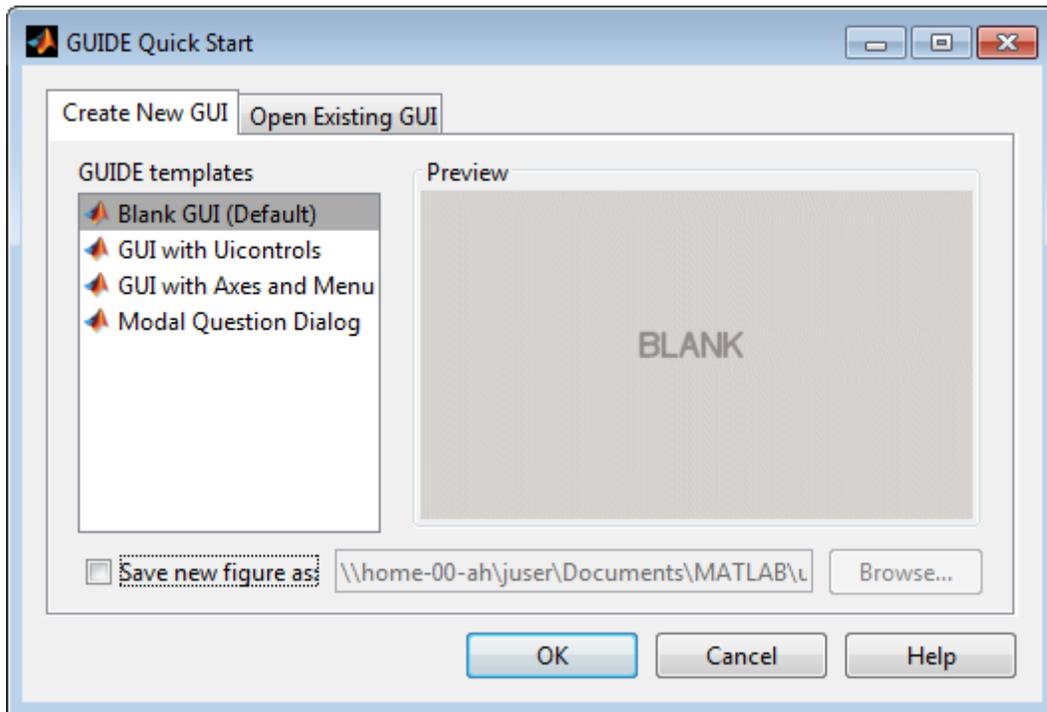


Figura 3-5. Acceso a un GUI.

2. En el cuadro de diálogo que aparece se selecciona la opción por defecto, Blank GUI (Default) y se da clic en OK. En la parte izquierda se observan la paleta de controles y en la parte derecha se sitúan el bosquejo (Figura 3-6).

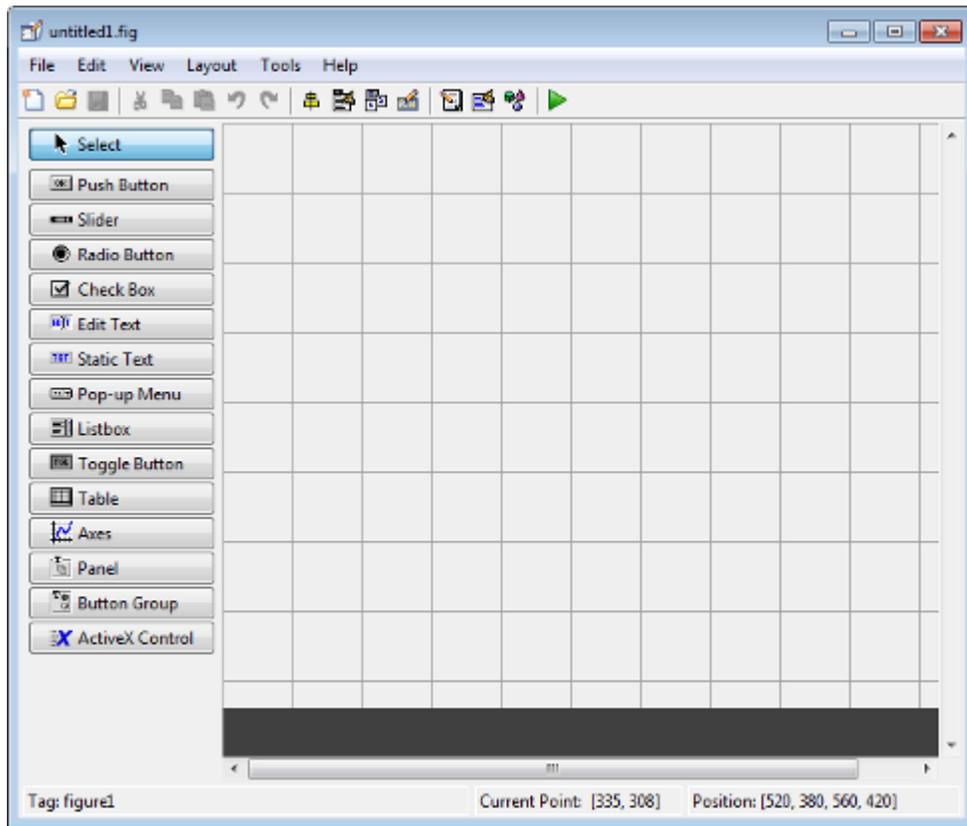


Figura 3-6 Ventana de edición gráfica GUIDE

3. En la paleta de controles se pueden agregar los botones y elementos al seleccionar los mismos (Push Button) del lado izquierdo y arrastrarlos al área de bosquejo, ver Figura 3-7.

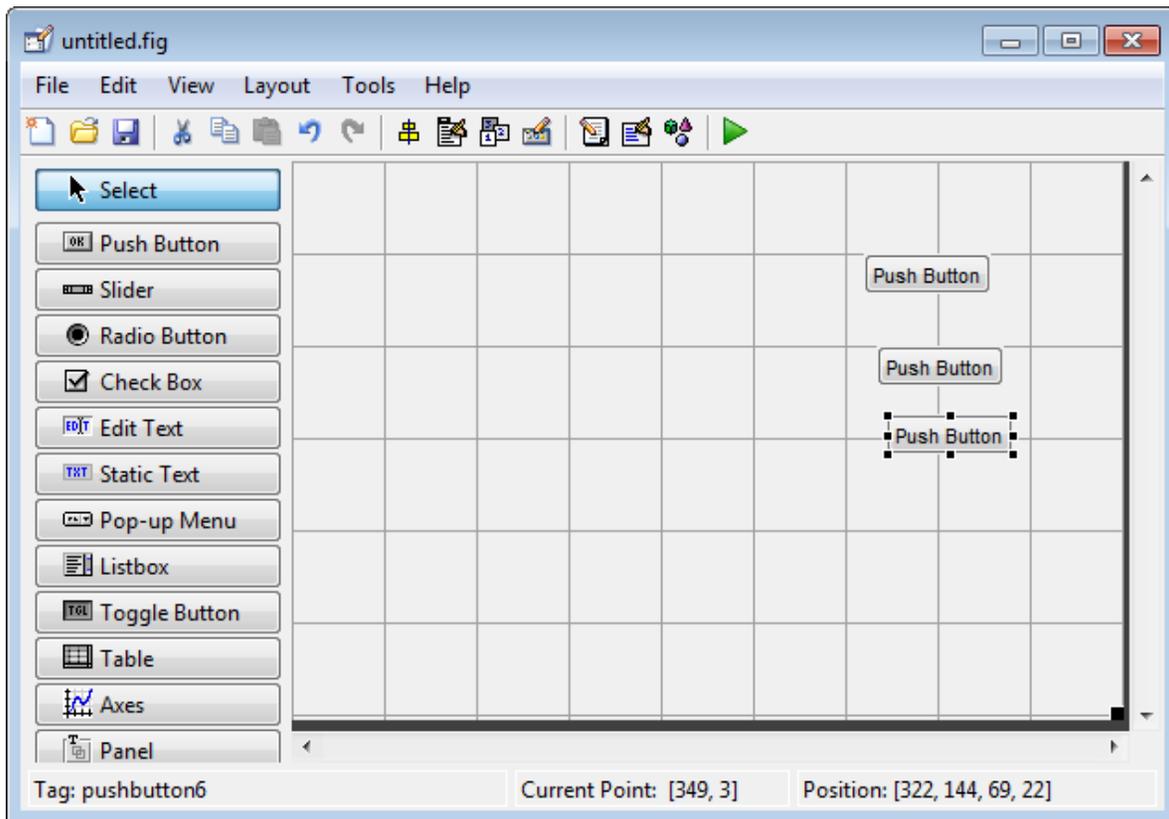


Figura 3-7 Insertar botones en la interfaz gráfica.

4. De igual manera, se pueden agregar el resto de los componentes de la interfaz gráfica a desarrollar, como un área estática de texto (Static Text), un menú desplegable (Pop-up Menu) y ejes (Axes), éstos últimos se pueden redimensionar al tamaño deseado (ver Figura 3-8).

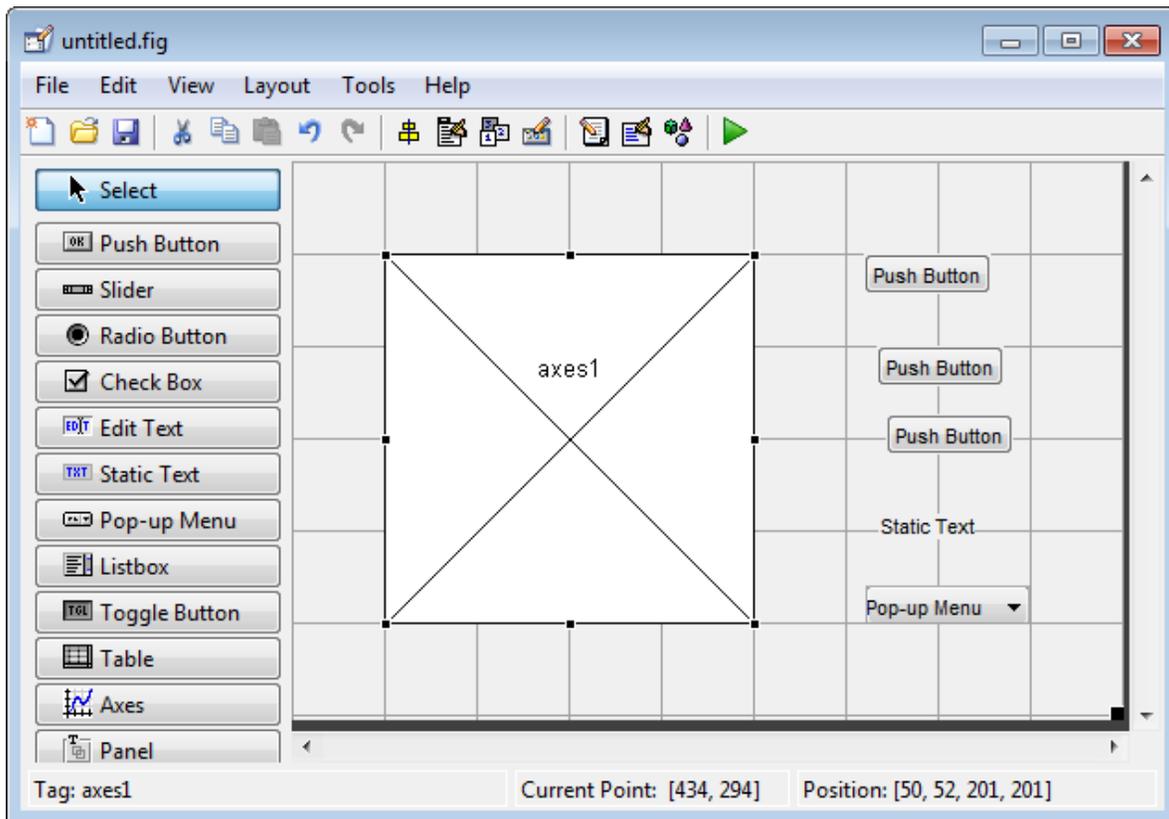


Figura 3-8 Componentes de una interfaz gráfica.

5. Los botones se pueden alinear con el siguiente procedimiento (Ver Figura 3-9):
 - a. Seleccionar los mismos, presionando Ctrl mientras se da clic en cada uno.
 - b. Seleccionar Tools > Align Objects.
 - c. En ajustes se requiere definir alineación izquierda en la dirección horizontal y 20 pixeles de espacio entre botones en la dirección vertical.
 - d. Dar clic en OK.

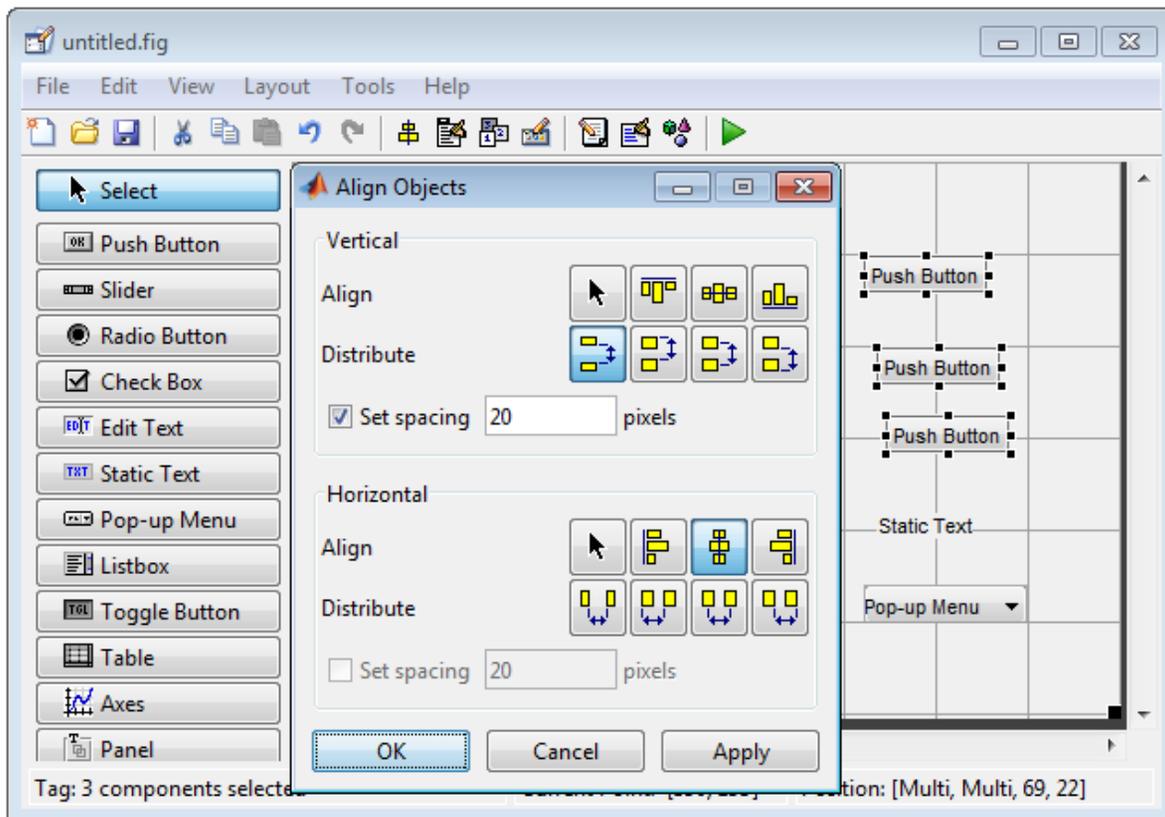


Figura 3-9 Alineación de botones.

6. Para etiquetar cada botón con el nombre adecuado, se requiere hacer lo siguiente (Figura 3-10):
 - a. Seleccionar **View > Property Inspector**.
 - b. En el área de bosquejo dar clic en el botón superior.
 - c. En el *Property Inspector*, seleccionar la propiedad *String*, y reemplazar el valor existente con la palabra deseada. Dar clic fuera del campo *String*, para que cambie la etiqueta a la palabra deseada.

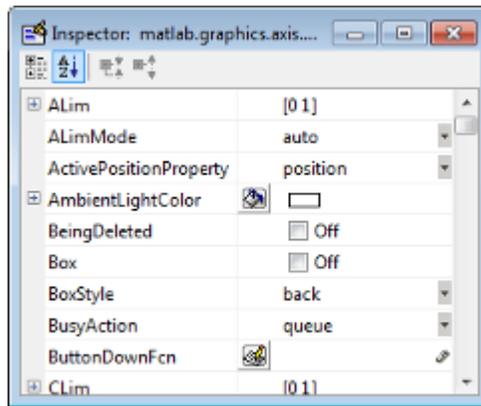


Figura 3-10 Menú de propiedades elementos GUI

- d. Dar clic en cada uno de los botones faltantes y repetir el paso anterior para cambiar las etiquetas de los mismos.
7. Para definir los objetos del menú desplegable, se requiere hacer lo siguiente:
- a. En el área de bosquejo, dar clic en el menú desplegable.
 - b. En el Property Inspector, dar clic en el botón ubicado junto a String, para mostrar el cuadro de diálogo del String.
 - c. Reemplazar el texto existente con los nombres de los objetos a desplegar. Presionar Enter para moverse entre líneas de texto.
 - d. Al terminar de editar los objetos, dar clic en OK. De manera que el primer objeto de la lista aparecerá en el menú desplegable en el área de bosquejo.

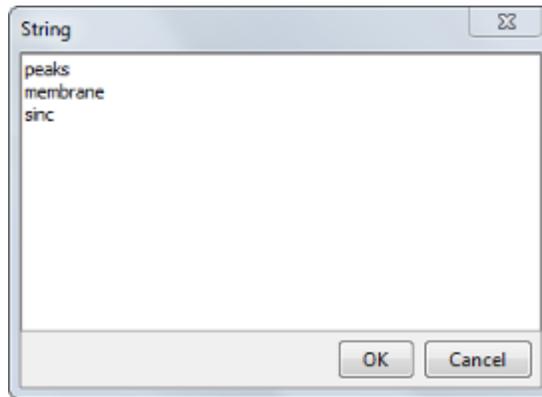


Figura 3-11 Modificación etiqueta de un elemento

8. Guardar el bosquejo seleccionando Tools > Run, de manera que, tras definir el nombre del archivo, GUIDE crea automáticamente los dos archivos con el mismo nombre y extensión. fig y .m que contienen la descripción del bosquejo en un archivo binario y las funciones de MATLAB para controlar las acciones del usuario sobre los controles, respectivamente. Posteriormente, corre el programa y abre el archivo con el código en el editor.

La aplicación creada se abre en una ventana nueva (ver Figura 3-12). Se observa que no contiene la barra de menú y las herramientas que se muestran por defecto en las ventanas de figuras de MATLAB. Se pueden agregar los menús propios, así como herramientas requeridas utilizando GUIDE.

Aquí se ha generado la interfaz gráfica, pero hasta no anexar el código correspondiente que indique qué hacer al hacer clic en los botones y seleccionar una opción en menú desplegable, no servirán los mismos.

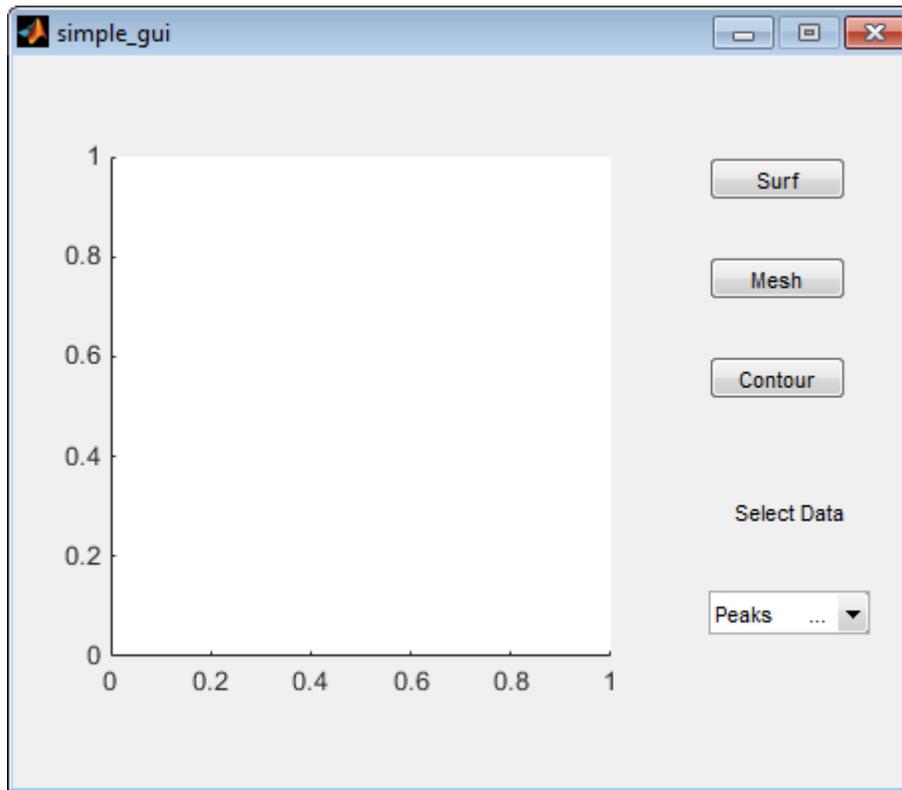


Figura 3-12 Boceto de una interfaz gráfica

9. Para generar datos para graficar se requiere abrir el Editor de MATLAB.
 - a. En la ventana EDITOR, en la sección NAVIGATE, dar clic en Go To, y seleccionar nombre_del_archiv_OpeningFcn.

Esto mueve el cursor a la función de abertura que contiene el código:

```
% --- Executes just before simple_gis made visible.
function simple_gui_OpeningFcn(hObject, eventdata, handles, varargin)
% This function has no output args, see OutputFcn.
% hObject   handle to figure
% eventdata reserved - to be defined in a future version of MATLAB
% handles   structure with handles and user data (see GUIDATA)
% varargin  command line arguments to simple_g(see VARARGIN)

% Choose default command line output for simple_gui
handles.output = hObject;
```

```

% Update handles structure
guidata(hObject, handles);

% UIWAIT makes simple_gwait for user response (see UIRESUME)
% uiwait(handles.figure1);

```

- b. Crear la información para graficar al ingresar el código que se muestra tras el comentario que comienza con % varagin...

```

% Create the data to plot.
handles.peaks=peaks(35);
handles.membrane=membrane;
[x,y] = meshgrid(-8:.5:8);
r = sqrt(x.^2+y.^2) + eps;
sinc = sin(r)./r;
handles.sinc = sinc;
% Set the current data value.
handles.current_data = handles.peaks;
surf(handles.current_data)

```

Las primeras seis líneas de texto ejecutable crean la información utilizando las funciones `peaks`, `membrane` y `sinc` de MATLAB. Además de guardar la información en la estructura de un apuntador, la cual puede ser llamada utilizando puntos de llamada por los botones previamente creados.

Las últimas dos líneas crean los valores de la información para definir los picos y desplegar los gráficos de curvas de nivel para los mismos, como se muestra en la Figura 3-13.

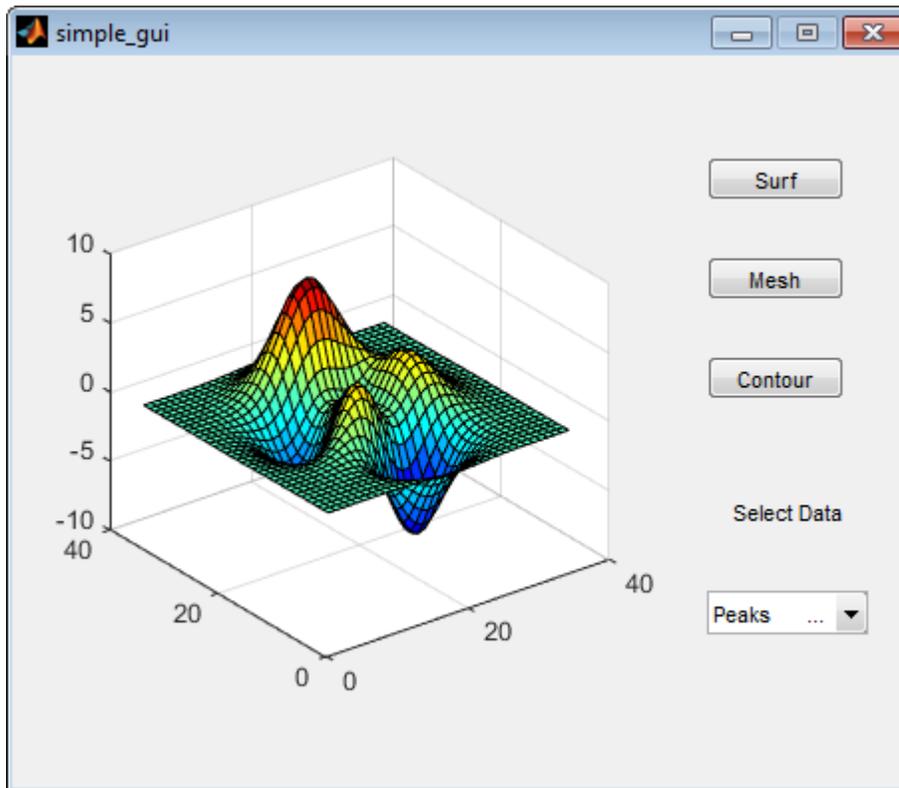


Figura 3-13 Ejemplo de una interfaz gráfica con ejes 3D

10. Codificar el comportamiento del menú desplegable.
 - a. Mostrar el menú desplegable en el punto de llamada del editor de MATLAB al seleccionar View Callbacks > Callback en el editor del bosquejo en GUIDE, el proceso se ilustra en la Figura 3-14.

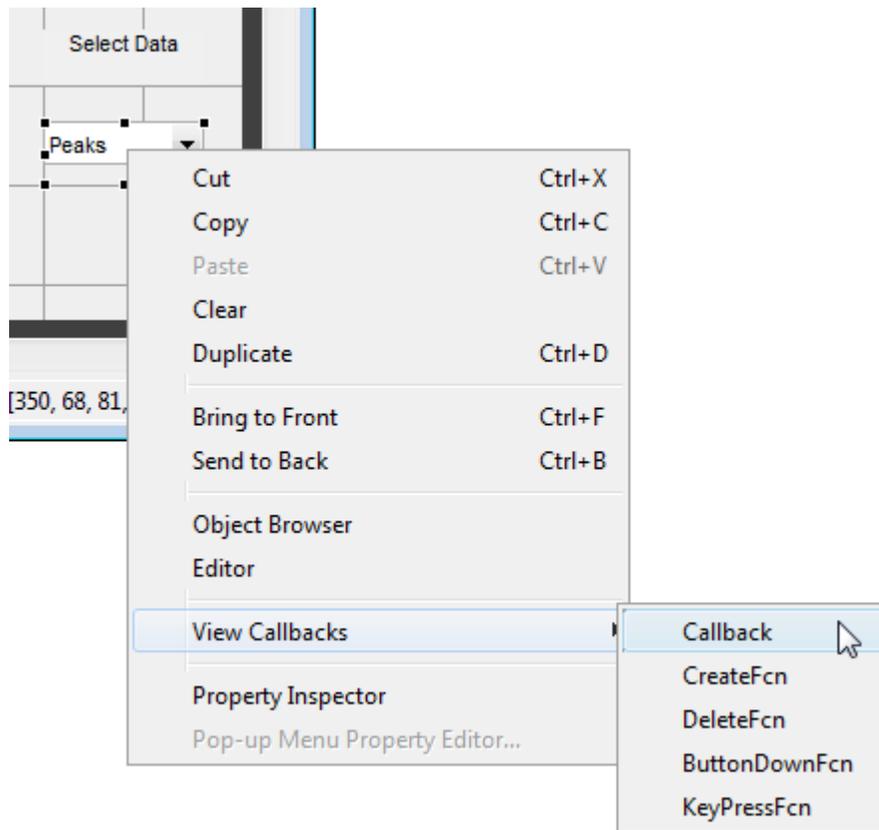


Figura 3-14 Acceso a callback de elemento gráfico

GUIDE mostrará el archivo de código en el editor y moverá el cursor al punto de llamada del menú desplegable, que contiene el código:

```
% --- Executes on selection change in popupmenu1.
function popupmenu1_Callback(hObject, eventdata, handles)
% hObject    handle to popupmenu1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
```

- b. Agregar el siguiente código en el popupmenu1_Callback después del comentario que comienza con % handles...

Este código primero recupera las propiedades: String y Value del menú desplegable, usa una declaración para cambiar la información a utilizar y, finalmente, guarda los cambios a la estructura del apuntador.

```
% Determine the selected data set.
str = get(hObject, 'String');
val = get(hObject, 'Value');
% Set current data to the selected data set.
switch str{val};
case 'peaks' % User selects peaks.
    handles.current_data = handles.peaks;
case 'membrane' % User selects membrane.
    handles.current_data = handles.membrane;
case 'sinc' % User selects sinc.
    handles.current_data = handles.sinc;
end
% Save the handles structure.
guidata(hObject,handles)
```

11. El comportamiento de los botones requiere mandar llamar cada botón para tomar la información de la estructura del apuntador para graficar la misma.
 - a. Mostrar el punto de llamada del botón Surf en el editor de MATLAB, al dar clic derecho en el botón en el área de bosquejo y seleccionando: View Callbacks > Callback.
 - b. En el editor, el cursor se mueve al punto de llamada del botón Surf en el archivo de código que contiene el código:

```
% --- Executes on button press in pushbutton1.
function pushbutton1_Callback(hObject, eventdata, handles)
% hObject    handle to pushbutton1 (see GCBO)
% eventdata  reserved - to be defined in a future version of MATLAB
% handles    structure with handles and user data (see GUIDATA)
```

- b. Después del comentario que empieza con % handles..., agregar el siguiente código al punto de llamada.

```
% Display surf plot of the currently selected data.  
surf(handles.current_data);
```

- c. Repetir ambos incisos para los puntos de llamada de los botones Mesh y Contour.

Agregar el siguiente código para el punto de llamada del botón Mesh, `pushbutton2_Callback`:

```
% Display mesh plot of the currently selected data.  
mesh(handles.current_data);
```

Agregar el siguiente código para el punto de llamada del botón Contour, `pushbutton3_Callback`:

```
% Display contour plot of the currently selected data.  
contour(handles.current_data);
```

- d. Guardar el código seleccionando: File > Save.
 - e. Correr el programa seleccionando en el editor del bosquejo Tools > Run.
12. En el menú desplegable se puede hacer la siguiente prueba: seleccionar Membrane, y dar clic en el botón Mesh. Esto desplegará la gráfica grid del logo MathWorks L-shaped Membrane.

4. RESULTADOS

La plataforma fue programada de manera modular, esto significa que se iba atendiendo cada uno de los objetivos específicos. Esta sección de resultados se enfoca en la presentación de cada uno de los módulos.

Ciertos módulos tienen comprobación de resultados con publicaciones anteriores, como el caso de las curvas IDT del país con las curvas IDT presentadas por la SCT; comparativo con software comercial actual, el módulo análisis de frecuencias con desarrollos como Ax o Easyfit; y aplicaciones predictivas o diagnósticas a datos ya medidos, módulos lógica difusa, series de tiempo e interpolación espacial.

4.1. Módulo Curvas IDT

El módulo Curvas IDT permite calcular las intensidades de lluvias en cualquier punto de país a partir de una base de datos interpolada con información de EMAS. Un Análisis de frecuencias a los registros EMAS, calcula los coeficientes μ y σ de Chen y Bell para el trazo de la curva IDT.

Como ya se señaló anteriormente, la lluvia se clasifica mediante condiciones medibles, intensidad (I) y duración (D) y se asigna una probabilidad de recurrencia o período de retorno (T). Al gráfico que relaciona de manera funcional a la Intensidad con las variables independientes T y D, se le conoce como Curva IDT. Para la construcción de este gráfico, es necesario el análisis numérico de hietogramas (tabla relacional de precipitación y tiempo).

Es importante señalar que esa tesis utiliza datos creados y publicados anteriormente por el CIAQ. Los desarrolladores de la base de datos utilizada en el módulo IDT de esta tesis, calcularon mediante el método de Chen y Bell (descritos anteriormente) a partir de datos de EMAS obtenidas del ERIC (Extractor rápido de información climatológica) y ajuste de frecuencias a datos de los 32 estados. La base de datos utilizada tiene registros de precipitación máxima en 24 horas desde

1930 y hasta 2010. Se trazaron curvas para periodos de retorno de 10, 25 y 50 años; y duraciones de 10, 30, 60, 120 y 240 minutos para las 189 EMAS analizadas. Se presenta un ejemplo de la base de datos consultada. La Tabla 4-1 presenta los registros de la estación 1003 Calvillo.

Tabla 4-1 Registros Precipitación máxima en 24 horas en Calvillo

Año	Pmax 24 horas
1932	27
1933	46,3
1934	41,2
1935	49,5
1936	59
1937	40,7
1938	39,6
1939	39,2
1940	60
1941	53
1942	45,6
1943	44,6
1944	28,6
1945	49,6
1946	39,7
1947	51,5
1948	35,4
1949	38,2
1954	47,4
1955	43,5
1956	56,2
1957	39,2
1958	50,6
1959	52,5
1960	32,7

1961	42,6
1962	64,5
1963	47,5
1964	38,5
1965	38
1966	51
1967	62
1968	66,5
1969	31
1970	43
1971	35,5
1972	52,2
1973	74,7
1974	33
1975	59,5
1976	39
1977	47
1978	40
1979	41
1980	55
1981	31
1982	53,5
1983	35,5
1984	40
1985	46
1986	100
1987	55,5
1988	60,5

Tras la aplicación de la metodología de Chen y Bell, se obtienen los coeficientes $\mu=0.5497$ y $\sigma=1.1653$. Con los coeficientes se determinan los valores para la curva IDT (Tabla 4-2) y se traza la curva (Figura 4-1):

Tabla 4-2 Registros curva IDT Calvillo

ESTACION	CALVILLO, CALVILLO
LONGITUD	-102,717
LATITUD	21.883
T10D10	82.0102195
T10D30	45.3634206
T10D60	29.7826326
T10D120	19.11355
T10D240	12.0673301
T25D10	98.8061458
T25D30	54.6539783
T25D60	35.8822005
T25D120	23.0280594
T25D240	14.5387537
T50D10	111.511774
T50D30	61.6820138
T50D60	40.4963455
T50D120	25.9892715
T50D240	16.4083134

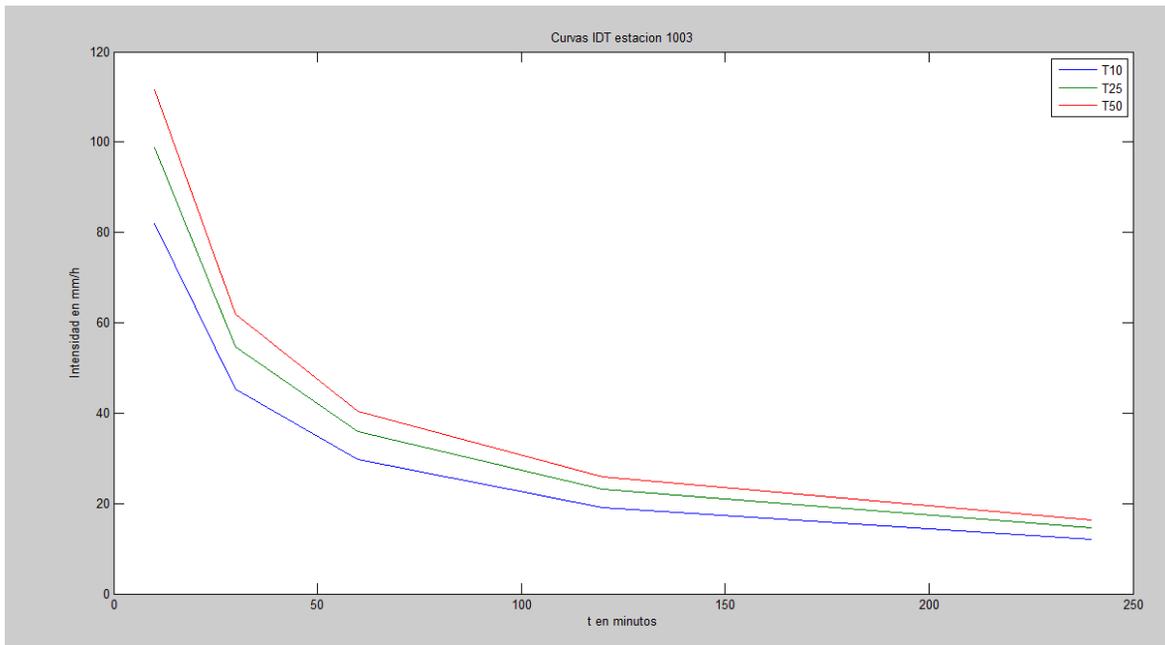


Figura 4-1 Curvas IDT en Calvillo

Una vez que se tienen los valores de las curvas IDT en cada estación, se crea un ráster con resolución de 100m (Tabla 4-3) mediante el algoritmo de Hutchinson para cada una de las duraciones e intensidades calculadas con Bell y Chen.

Tabla 4-3 Ráster con valores de Intensidad de lluvia

Longitud	Latitud	Intensidad T10d10
-102.824346957280	21.622143290000	84.184672622179
-102.822328427570	21.622143290000	84.153191877378
-102.820309897860	21.622143290000	84.121302011050
-102.818291368160	21.622143290000	84.089004193619
-102.816272838450	21.622143290000	84.056299965735
-102.814254308740	21.622143290000	84.023191243728
-102.812235779030	21.622143290000	83.989680324244
-102.810217249320	21.622143290000	83.955769888011
-102.808198719610	21.622143290000	83.921463002722
-102.806180189900	21.622143290000	83.886763124978
-102.804161660190	21.622143290000	83.851674101284
-102.802143130490	21.622143290000	83.816200168051
-102.800124600780	21.622143290000	83.780345950598

-102.798106071070	21.622143290000	83.744116461135
-102.796087541360	21.622143290000	83.707517095695
-102.794069011650	21.622143290000	83.670553630026
-102.792050481940	21.622143290000	83.633232214429
-102.790031952230	21.622143290000	83.595559367537
-102.788013422520	21.622143290000	83.557541969037
-102.785994892820	21.622143290000	83.519187251357
-102.783976363110	21.622143290000	83.480502790307
-102.781957833400	21.622143290000	83.441496494710
-102.779939303690	21.622143290000	83.402176595034
-102.777920773980	21.622143290000	83.362551631052
-102.775902244270	21.622143290000	83.322630438567
-102.773883714560	21.622143290000	83.282422135222

Con cada una de las duraciones y los estados, se construye una base de datos a la que la interfaz gráfica se conecta. Dicha base de datos está disponible en: <http://redciaq.uaq.mx/system/idt/CIAQ/IDTpais/Auxiliares/puntos/>

Con los puntos cargados como un matriz de Matlab. Se utiliza la triangulación de Delauney para crear un polinomio interpolador que permita estimar el valor de intensidad de lluvia para cualquier punto en cada uno de los estados. Esta información se manipula por el usuario mediante una interfaz compuesta por 3 botones de acción: cargar mapa, nuevo cálculo y continuar; un menú selector, un panel de botones de radio y 3 elementos de texto editable (Figura 4-2).

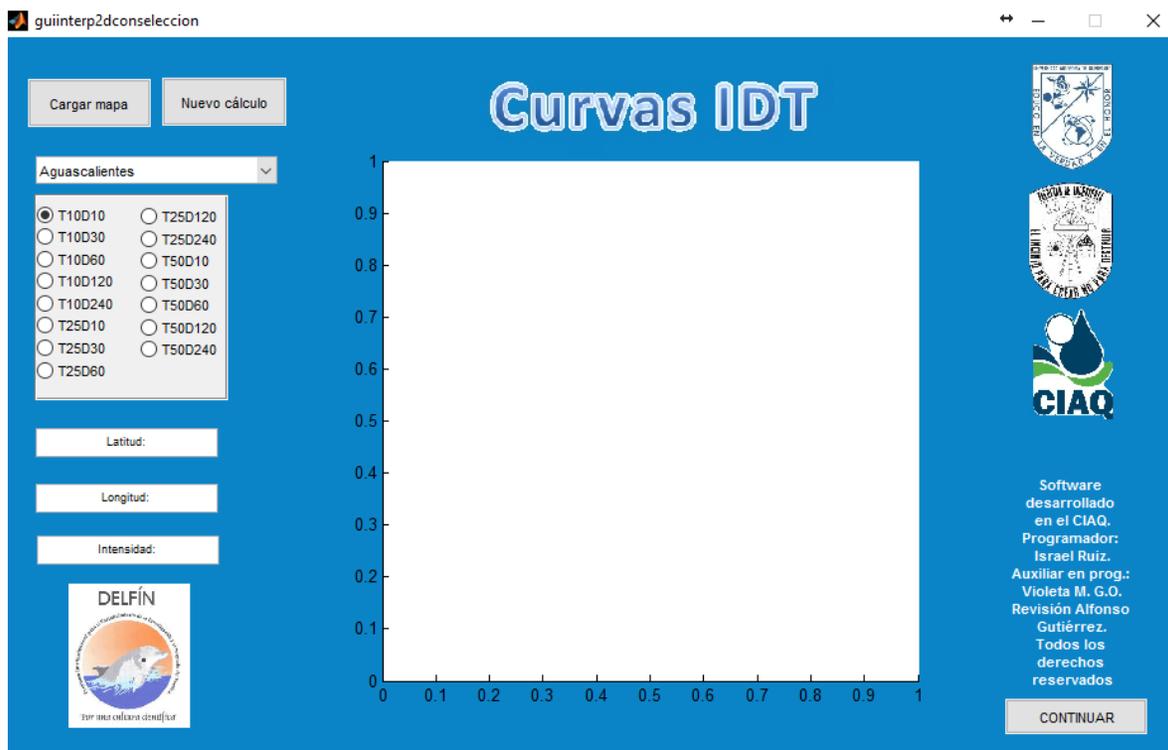


Figura 4-2 Interfaz de bienvenida Módulo IDT

Para utilizar el programa es necesario seleccionar un estado (menú selector) y una combinación período de retorno (T) y de duración (D). En el ejemplo se utiliza Querétaro con una tormenta de 25 años y 30 minutos de duración (Figura 4-3).

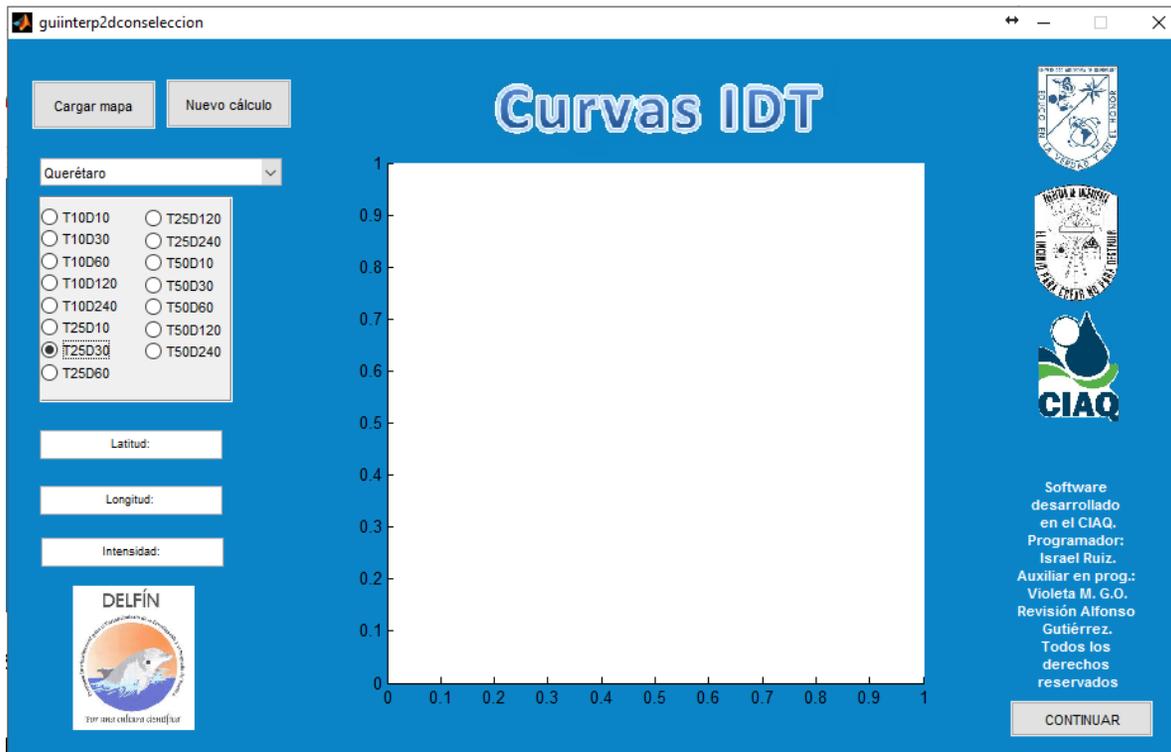


Figura 4-3 Panel de Control con botones de radio

Para utilizar el programa es necesario seleccionar un estado (menú selector) y una combinación período de retorno (T) y de duración (D). En el ejemplo se utiliza Querétaro con una tormenta de 25 años y 30 minutos de duración (Figura 4-4).

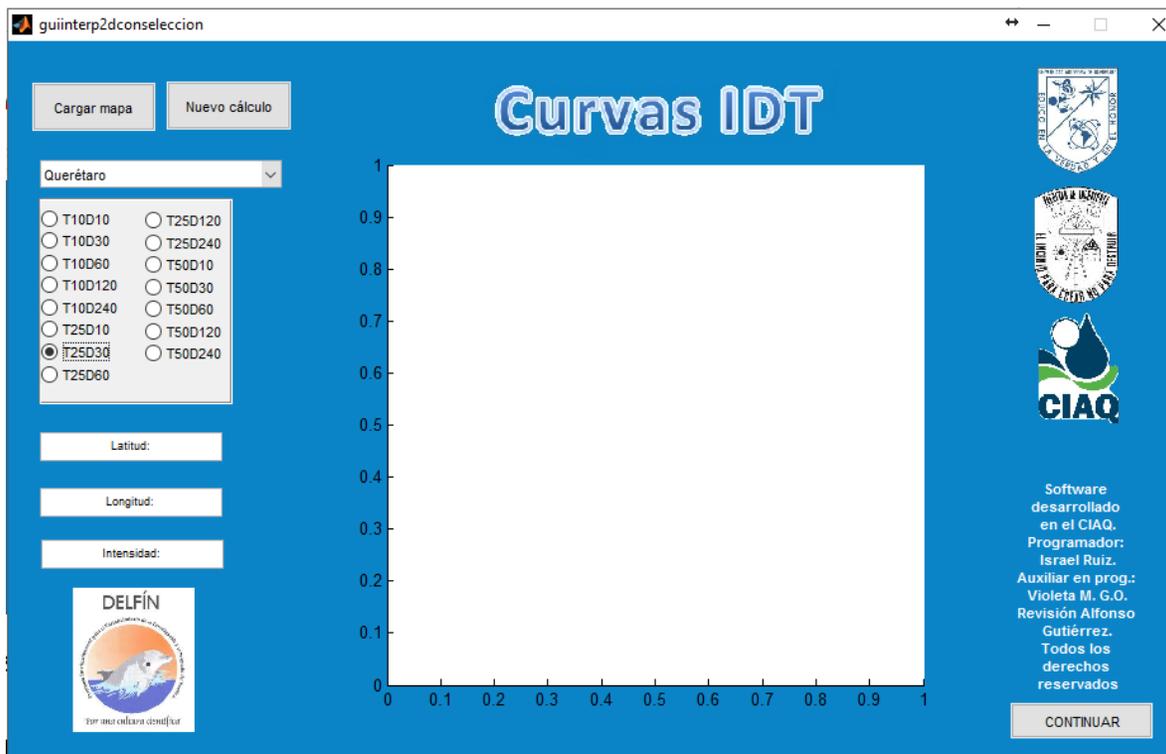


Figura 4-4 Selección de Estado y condiciones de tormenta

En este momento el programa puede tomar dos direcciones distintas, la primera cuando no hay escrita una dirección en latitud y longitud y la segunda cuando uno definió anteriormente la coordenada del punto a analizar.

- 1) No hay datos de longitud y latitud: Se presiona el botón Cargar mapa para generar un mapa con el estado seleccionado, además el programa no detecta un punto de análisis por lo que el cursor cambia a uno de selección espacial en donde mediante un clic, el usuario define la coordenada a analizar. Finalmente al presionar Nuevo cálculo (Figura 4-5), el programa encuentra la intensidad en el punto seleccionado mediante la triangulación de Delauney.

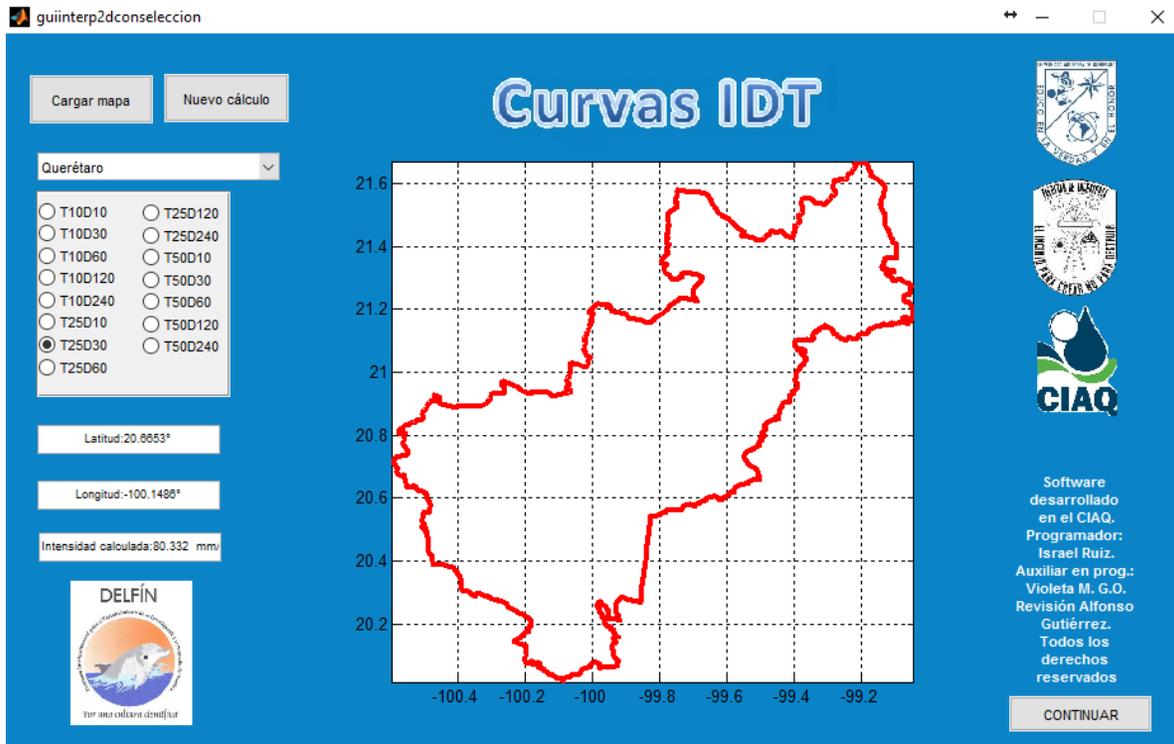


Figura 4-5 Selección de punto de análisis por geografía

- 2) Cuando hay un registro de latitud y longitud, el programa considera éste como el punto de análisis y al presionar Nuevo cálculo, estima la intensidad solicitada (Figura 4-6).

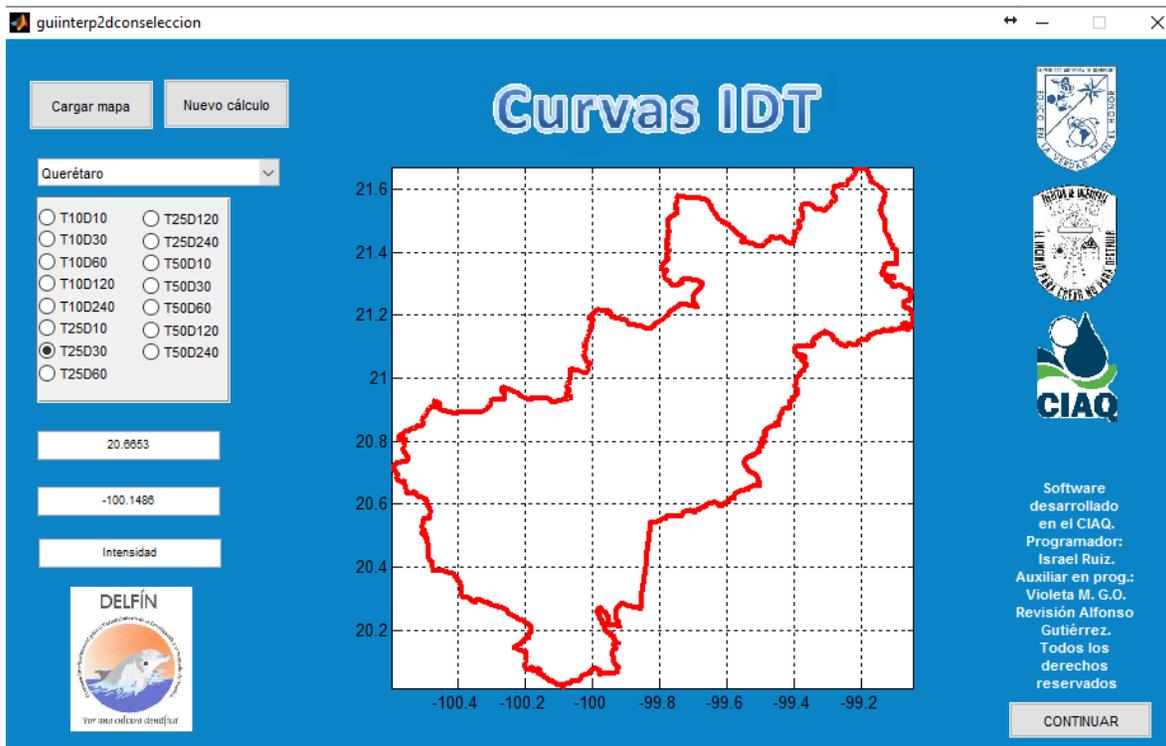


Figura 4-6 Latitud y Longitud previamente definidas

Una vez que se tiene un punto de análisis, el botón de Continuar vincula la interfaz anterior con la interfaz IDT. Ésta interfaz (Figura 4-7 Ventana para análisis puntual de una curva IDT) toma como dato de entrada el punto de análisis seleccionado en la interfaz anterior y en ese punto, dibuja las curvas IDT de 10, 25 y 50 años al presionar el botón de Graficar (Figura 4-8 Mensaje de espera debido al tiempo de ejecución del módulo IDT y Figura 4-9 Gráfica con las curvas IDT 10, 25 y 50 años).

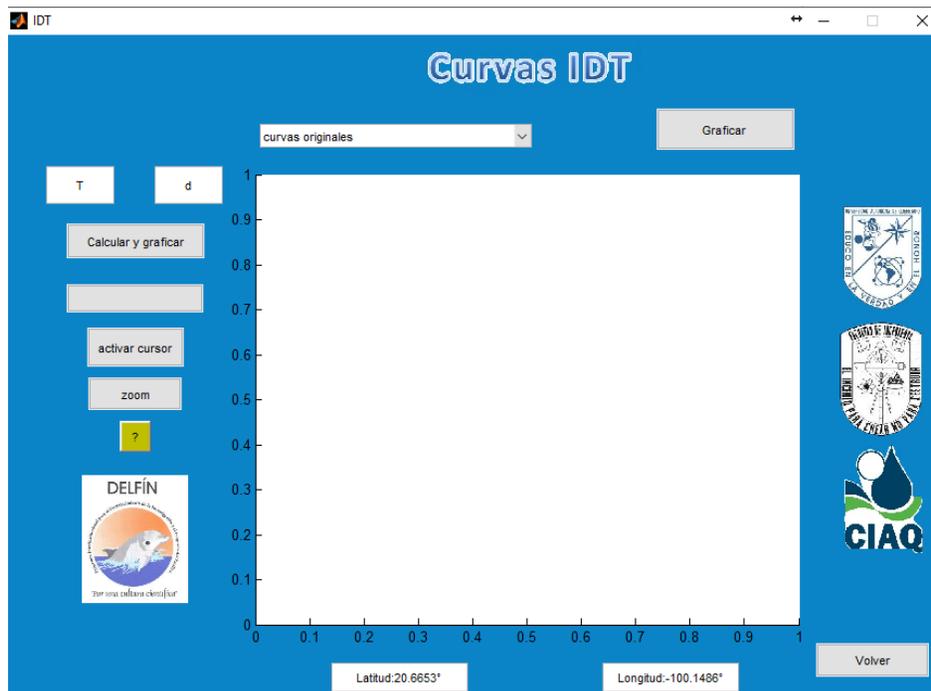


Figura 4-7 Ventana para análisis puntual de una curva IDT

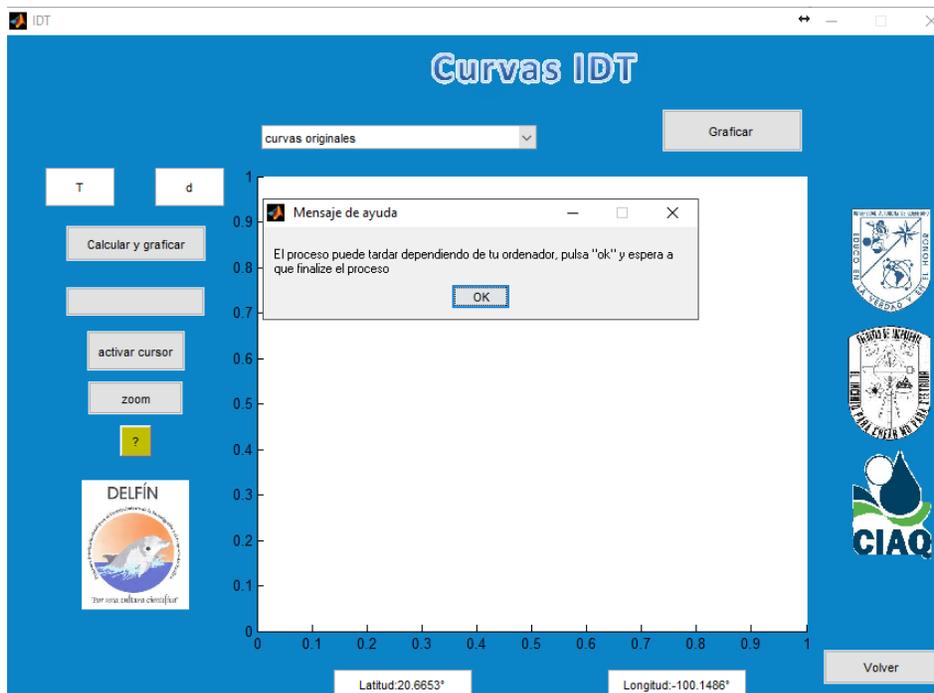


Figura 4-8 Mensaje de espera debido al tiempo de ejecución del módulo IDT

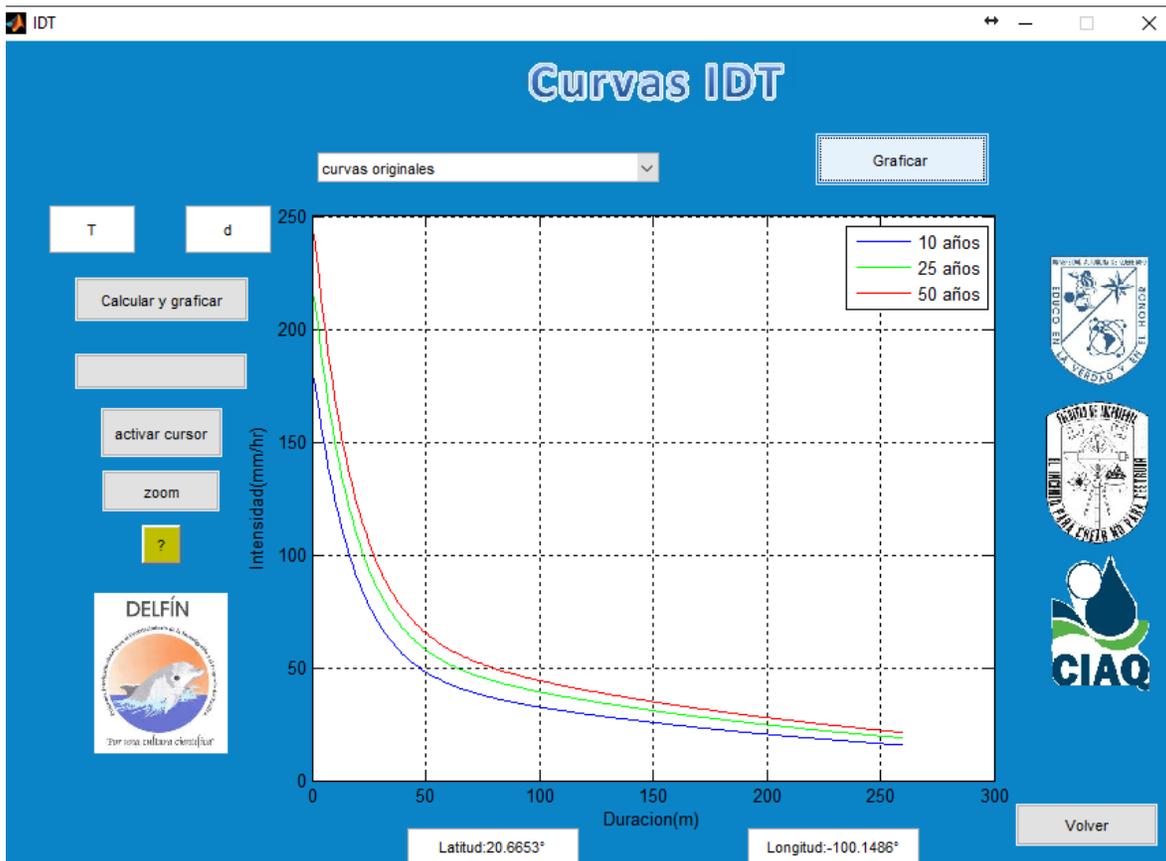


Figura 4-9 Gráfica con las curvas IDT 10, 25 y 50 años

Una vez graficadas las curvas 10, 25 y 50, el usuario puede calcular una lluvia de diseño distinta a las existentes en la base de datos. Para este cálculo es necesario llenar los campos disponibles para T y D. En este ejercicio se decide graficar la curva de 32 años (interpolación) y la de 75 años (extrapolación, Figura 4-11) con duraciones de 90 minutos. Por cuestiones de visualización, se hace un zoom a la gráfica para ver la ubicación de la línea estimada, ver Figura 4-10 Herramientas gráficas en la interfaz.

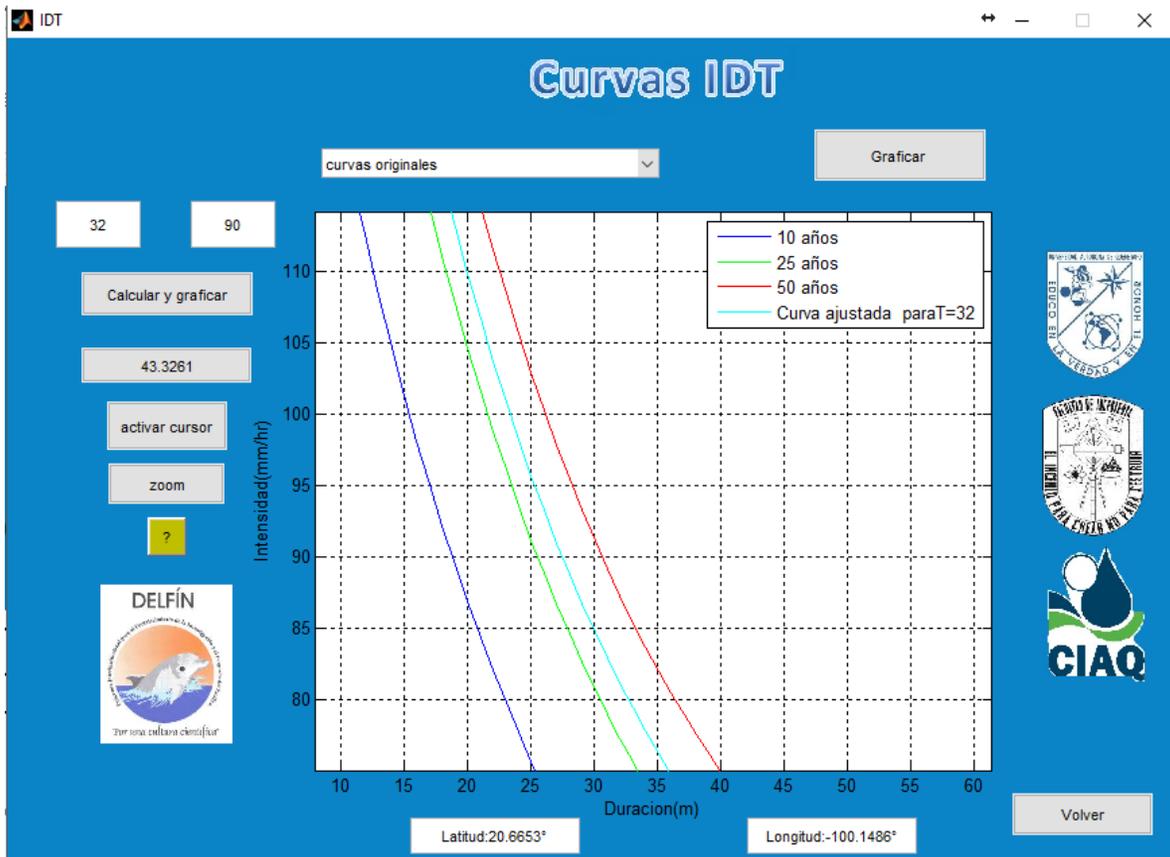


Figura 4-10 Herramientas gráficas en la interfaz

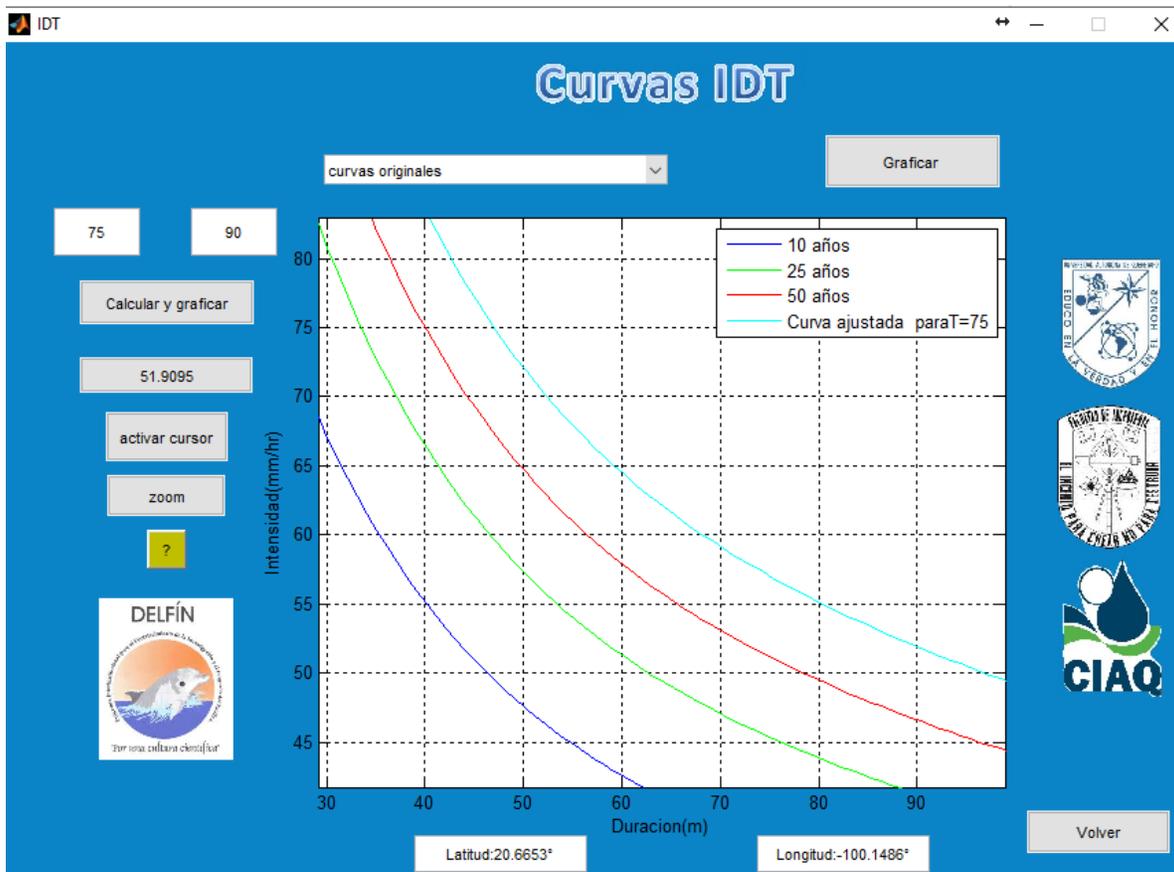


Figura 4-11 Curva extrapolada

El cálculo de las nuevas curvas y sus intensidades puntuales (botón activar cursor Figura 4-12) es posible mediante el ajuste de curvas. A diferencia de los modelos de Chen y Bell, se asume que la curva IDT tiene un comportamiento exponencial decreciente, se sumen ecuaciones del tipo doble exponencial y se calculan los parámetros para encontrar el mejor ajuste lineal (Best Linear Unbiased Estimator) mediante el teorema de Gauss-Markov.

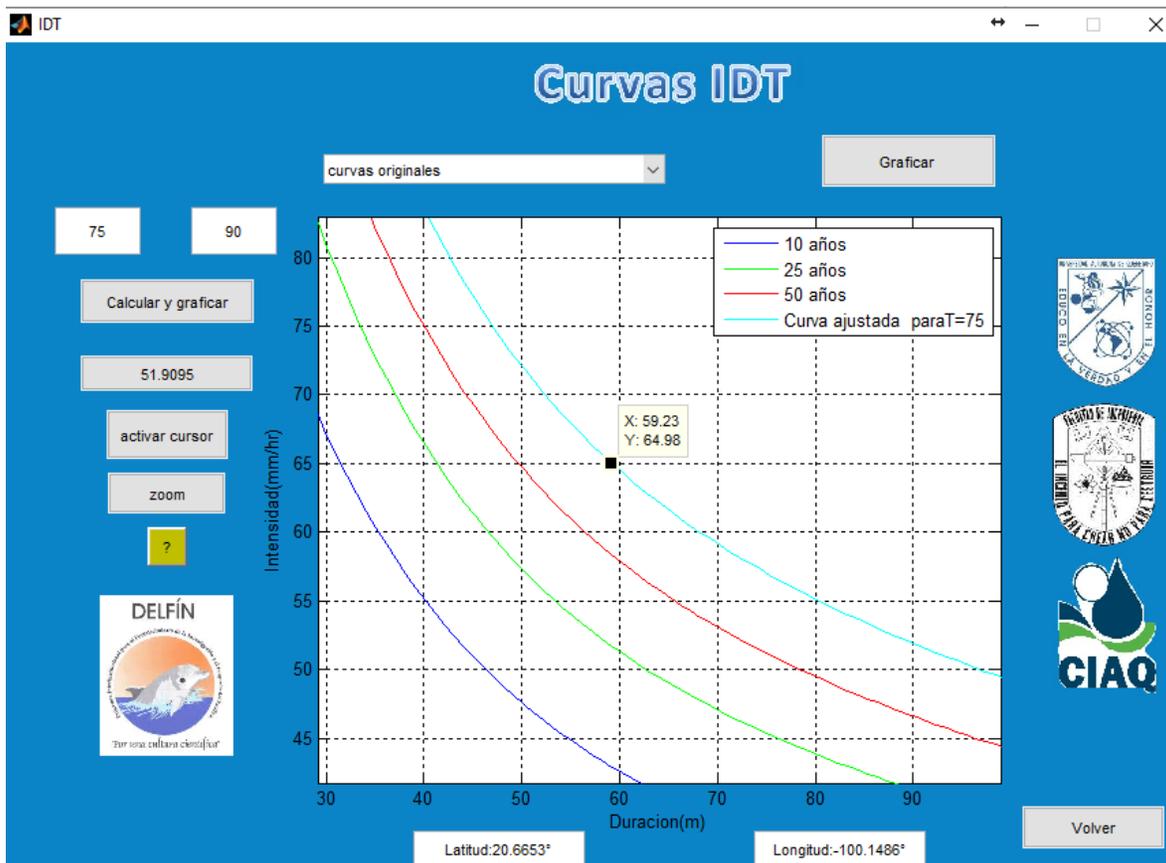


Figura 4-12 Cursos de datos

Finalmente, el botón de Volver, permite ir a la interfaz de entrada de locación por si se quiere analizar un punto adicional.

Para la validación de este programa, se realizó un comparativo con las curvas IDT presentadas por la SCT para todo el país. Si bien se tiene el registro de 189 EMAS en todos los estados en los tres periodos de retorno presentes en la base de datos (10, 25 y 50 años), se presenta el comparativo del programa con las curvas de Querétaro para los tres periodos de retorno en las tres estaciones consideradas; así como las curvas para T=10 en Baja California.

Tabla 4-4 Comparativo módulo IDT vs Curvas oficiales SCT

FID_	Nombre Estación	Coord.		CURVAS	ACTUAL	CURVAS	ACTUAL	CURVAS	ACTUAL	CURVAS	ACTUAL	CURVAS
		X	Y	IDT	D10	IDT	D30	IDT	D60	IDT	D120	IDT
15	San Vicente	-116.26723	31.147171	43		25	23.7137	18	15.5689	14	11.9096	9
	Santo Domingo	-115.86466	30.763269	47	51.7862	33	28.6452	24	18.8065	15	12.0694	9
	Santa Cruz	-115.63064	30.916436	62	59.0478	47	32.6619	32	21.4437	24	13.7619	16
	Santo Tomas	-116.36481	31.544168	66	50.2625	30	27.8023	19	18.2532	13	11.7143	9
	San Carlos	-116.46324	31.779192	61	48.8927	37	27.0447	28	17.7558	18	11.3951	10
	Ensenada	-116.63499	31.877621	48	47.3766	26	25.9954	16	17.0668	13	10.953	8
	Mexicali	-115.46572	32.6305	70	32.1033	35	17.7578	32	11.6586	16	7.4821	9
	El Centinela	-115.84788	32.5273	73	40.0886	41	22.1747	33	14.5585	28	9.3432	10
	Cañada Verde	-116.4136	32.494909	95	54.7044	52	30.2594	33	19.8663	18	12.7496	10
	Presa Rodriguez	-116.90072	32.446204	55	42.7532	31	23.6487	25	15.5262	17	9.9642	9
	Tijuana	-117.02953	32.516259	98	45.376	60	25.0994	40	16.4786	23	10.5755	12
	Valle Palmas	-116.61375	32.360518	52	43.6062	21	24.1205	25	15.8359	7	10.163	3
	Olivares	-116.6145	32.044139	74	53.25	34	29.4549	24	19.3382	19	12.4106	14
	Boquilla de Sta Rosa	-116.74708	32.044139	64	54.0472	35	29.8958	28	19.6276	23	12.5964	20
Agua Caliente	-116.44878	32.111934	52	50.2519	23	27.7965	16	18.2494	13	11.7119	8	

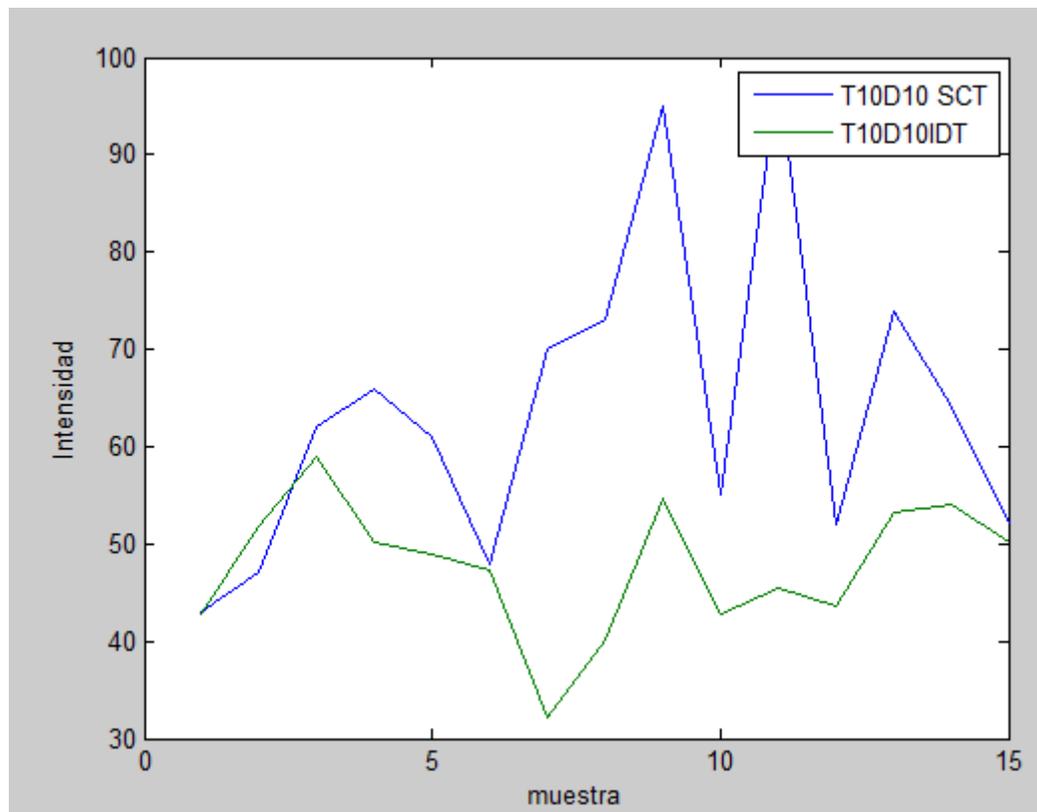


Figura 4-13 Comparativo módulo IDT vs Curvas SCT

Tabla 4-5 Estaciones de Querétaro T=10 años

QUERETARO	-100.3824	20.599955	99	133.632	61	73.9176	40	48.5294	23	31.1435	13	19.6624
EL PUEBLITO	100.44963	20.533799	105	134.539	68	74.4194	46	48.8589	26	31.3561	14	19.7967
SAN JUAN DEL RIO	-99.999	20.366612	125	141.555	79	78.3002	60	51.4067	32	32.9912	16	20.829

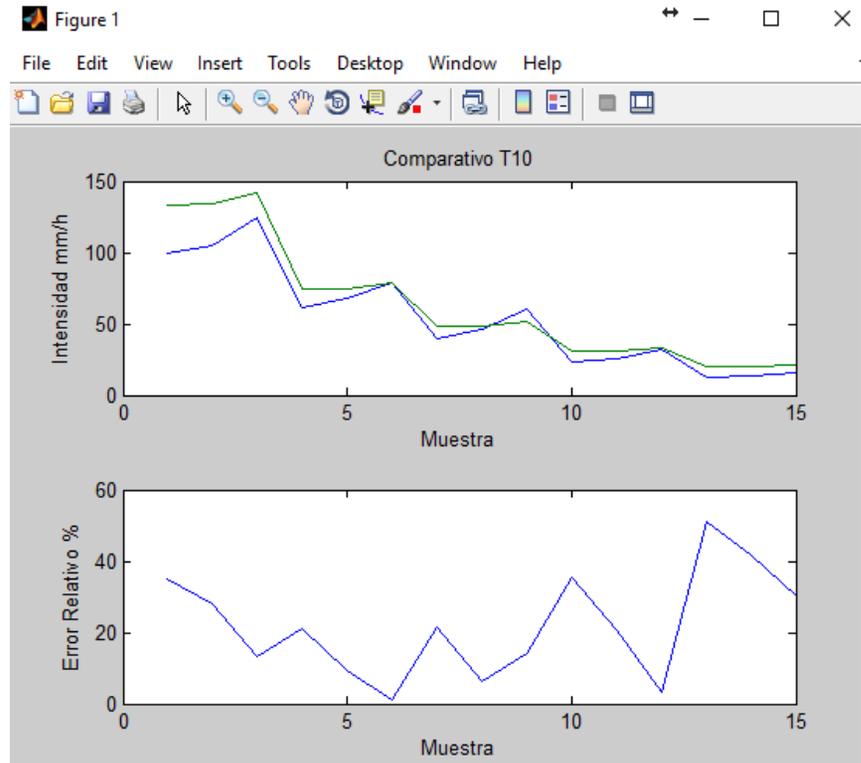


Figura 4-14 Comparativo para T=10 años

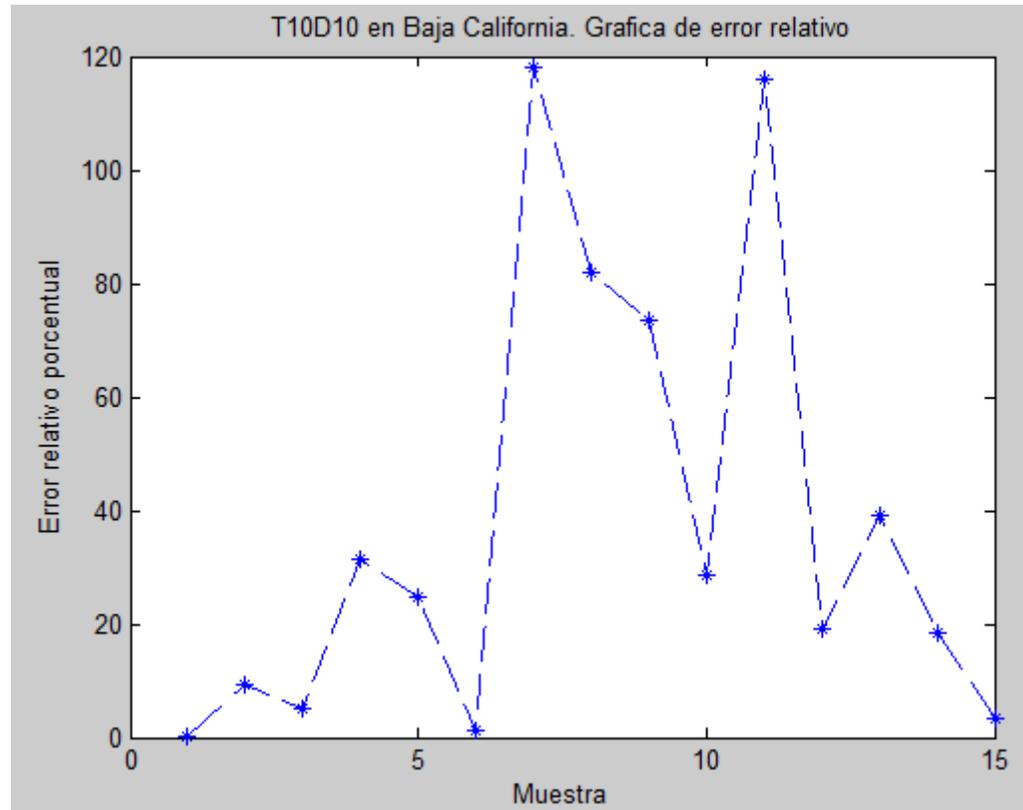


Figura 4-15 Gráfica Error relativo en Baja California

Tabla 4-6 Estaciones de Querétaro T=25 años

QUERETARO	-100.3824	20.599955	109	160.994	68	89.0529	45	58.4663	25	37.5218	14	23.6894
EL PUEBLITO	-100.44963	20.533799	126	162.093	80	89.6608	57	58.8553	32	37.7779	18	23.8511
SAN JUAN DEL RIO	-99.999	20.366612	151	170.546	96	94.3362	74	61.935	39	39.7479	20	25.0948

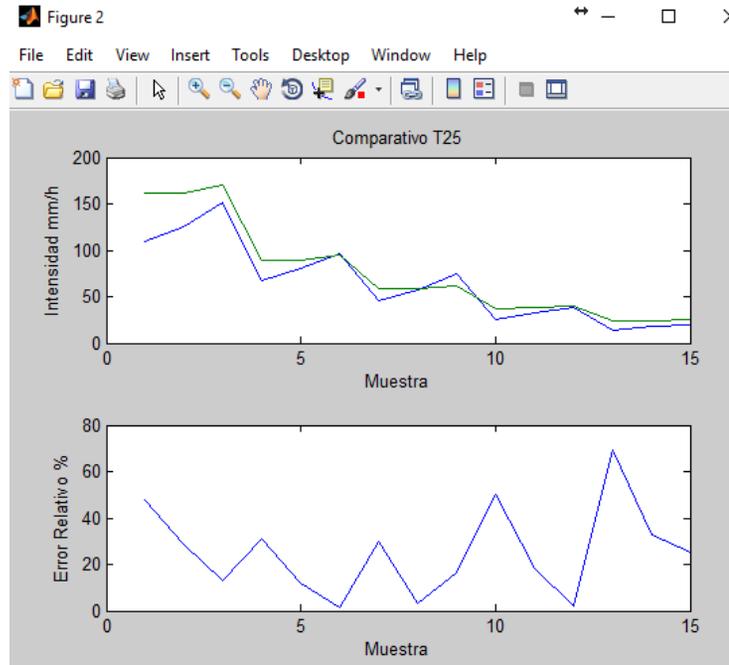


Figura 4-16 Comparativo para T= 25 años

Tabla 4-7 Estaciones de Querétaro T=50 años

QUERETARO	-100.3824	20.599955	178	181.6967	95	100.5043	63	65.9845	41	45.4552	24	26.7356
EL PUEBLITO	-100.44963	20.533799	155	182.9369	98	101.1904	68	66.4349	44	42.6358	25	26.9181
SAN JUAN DEL RIO	-99.999	20.366612	194	192.4764	103	106.467	74	69.8993	47	44.8591	26	28.3218

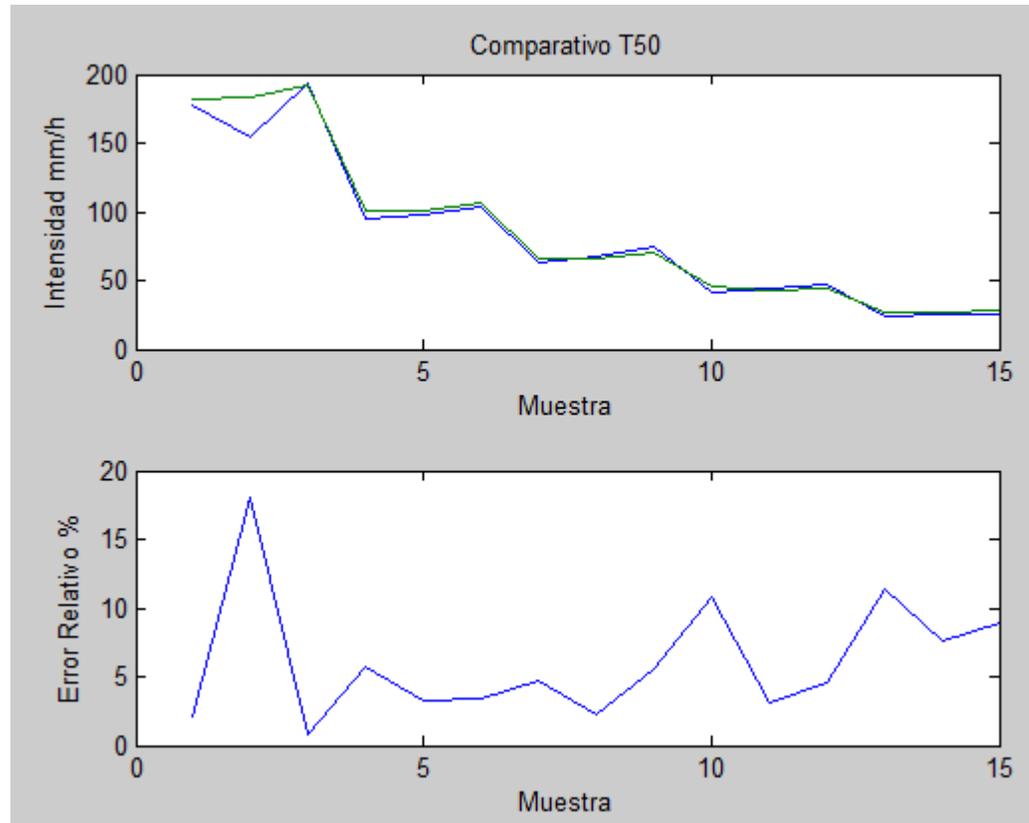


Figura 4-17 Comparativo para T=50 años

4.2. Módulo de Análisis de Frecuencias

El módulo Análisis de Frecuencias, ver Figura 4-18, se utiliza para ajustar funciones de distribución de probabilidad a una serie de datos. Encuentra los coeficientes de la distribución mediante el método de momentos de probabilidad.

Las distribuciones consideradas en este módulo son: Gumbel, exponencial de dos parámetros, Gamma, y la combinación de éstas.

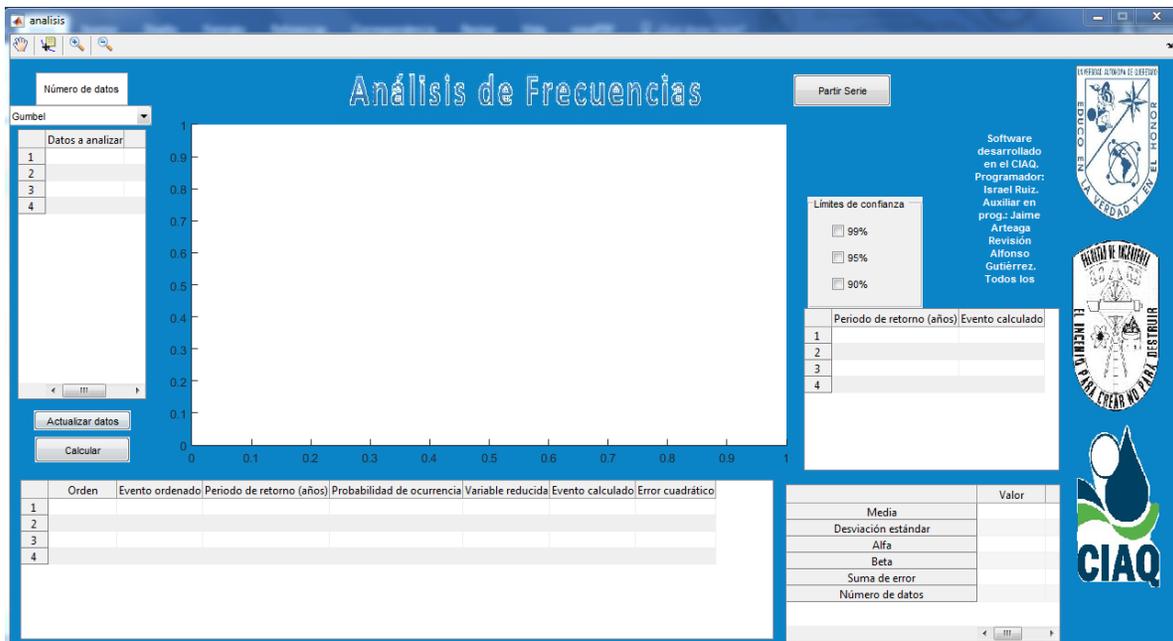


Figura 4-18 Interfaz gráfica Análisis de Frecuencias

La interfaz trabaja a partir del ingreso del número de datos en el campo indicado. Una vez puesto el número de muestras, con el botón de actualizar se crea una tabla dinámica que permite la inserción de datos para ajuste. En el ejemplo se realiza el ajuste a 15 datos (Figura 4-19).

$$x[n] = [15 \ 14 \ 17 \ 18 \ 95 \ 62 \ 35 \ 48 \ 95 \ 47 \ 12 \ 63 \ 65 \ 53 \ 56] \quad (4.1)$$

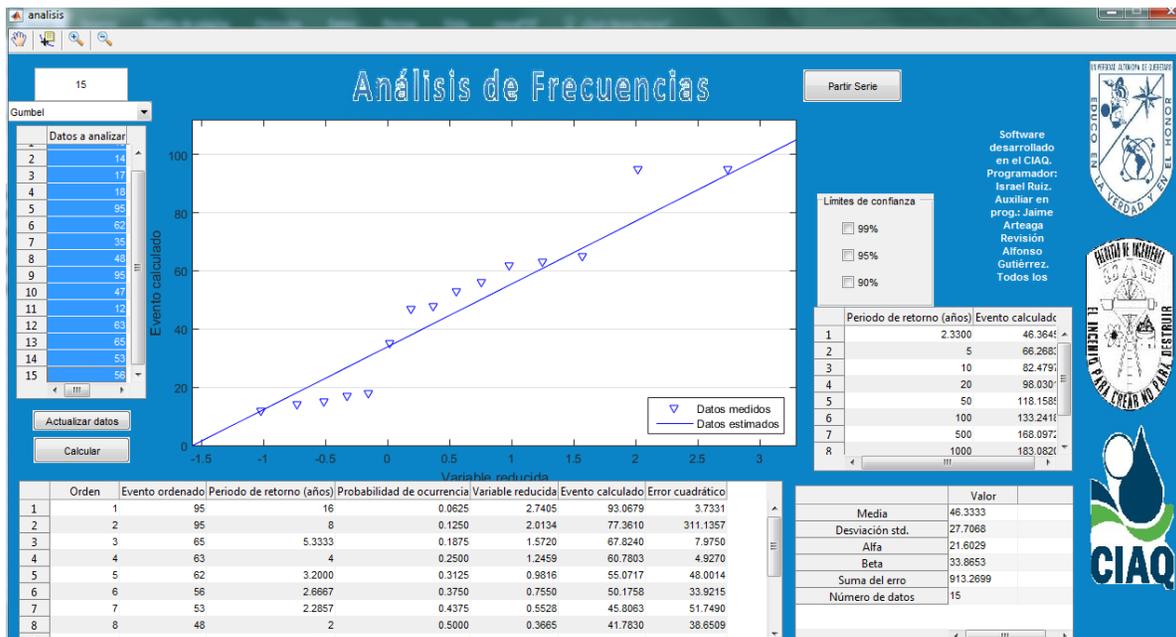


Figura 4-19 Ajuste de Frecuencias por Gumbel a 15 registros

Al presionar el botón **Calcular** se realiza el ajuste seleccionado en el menú. Las tablas muestran los distintos parámetros de ajuste de la distribución y sus resultados de errores, variable reducida y probabilidades asociadas mediante la ecuación de Weibull.

El programa permite además la estimación de funciones con margen de error. Al habilitar los botones de 90 (Figura 4-20), 95 y 99% (Figura 4-21) se gráfica las distribuciones que consideran este error porcentual negativo y positivo.

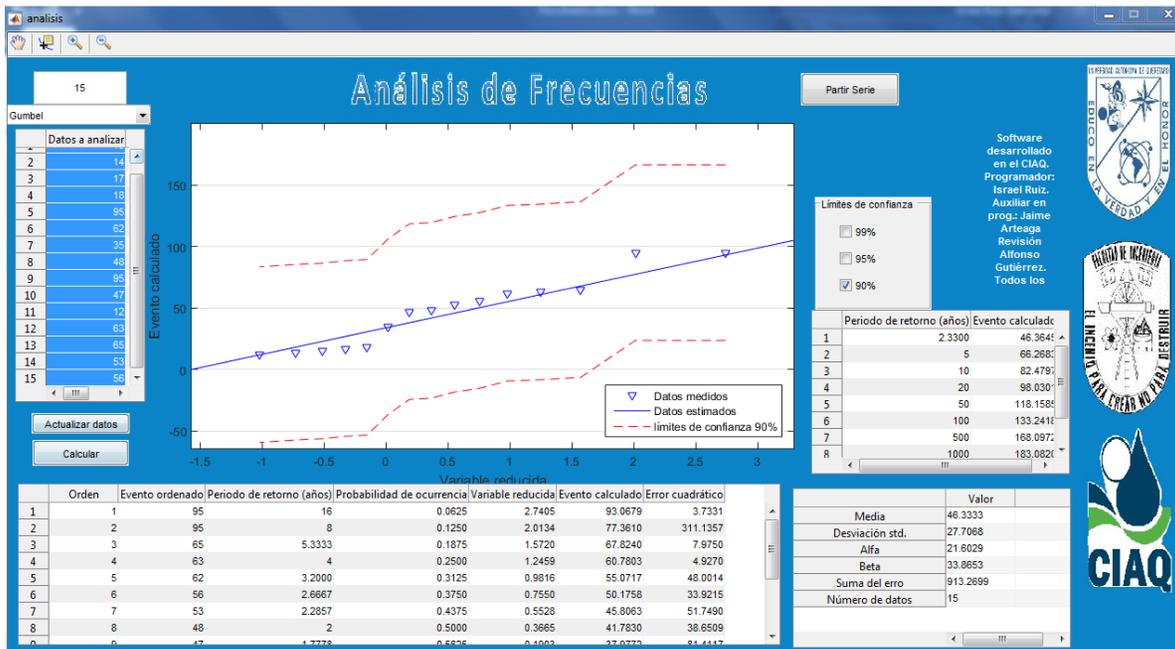


Figura 4-20 Límites de confianza en la plataforma.

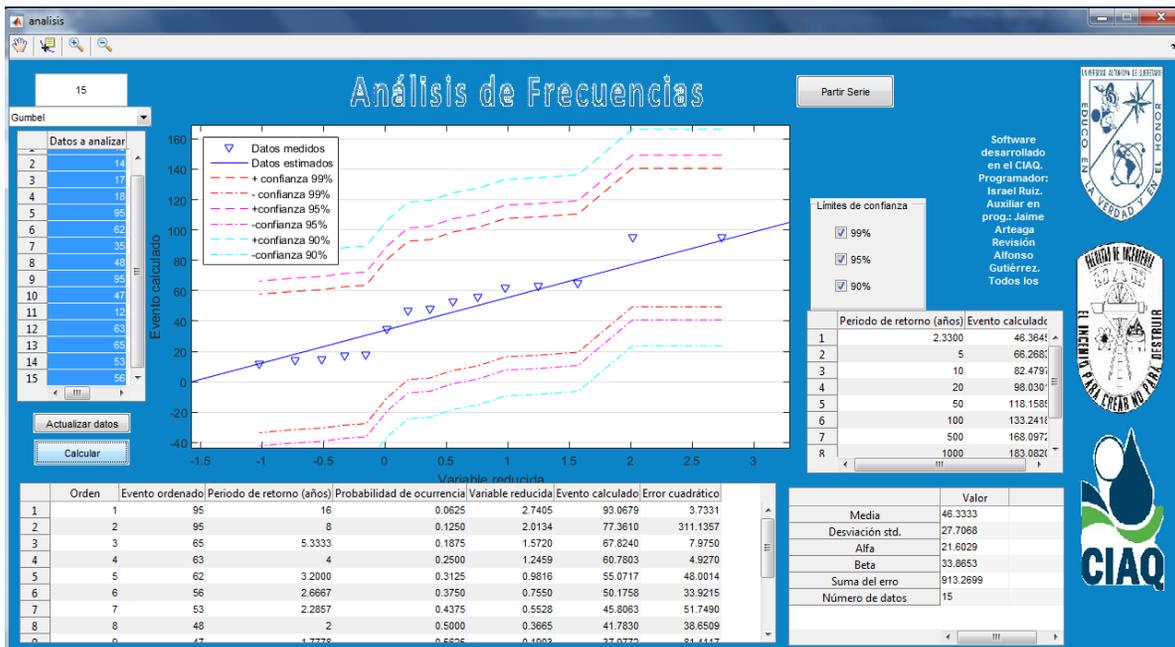


Figura 4-21 Límites de confianza a la par en la ventana gráfica.

Finalmente, si se quiere trabajar con distribuciones mezcladas, se procede a la Interfaz Partir serie. Esta interfaz (Figura 4-22 Rompimiento gráfico de serie) mediante un cursor de selección espacial gráfica los datos y permite al usuario mediante un click seleccionar el punto de quiebre de la serie.

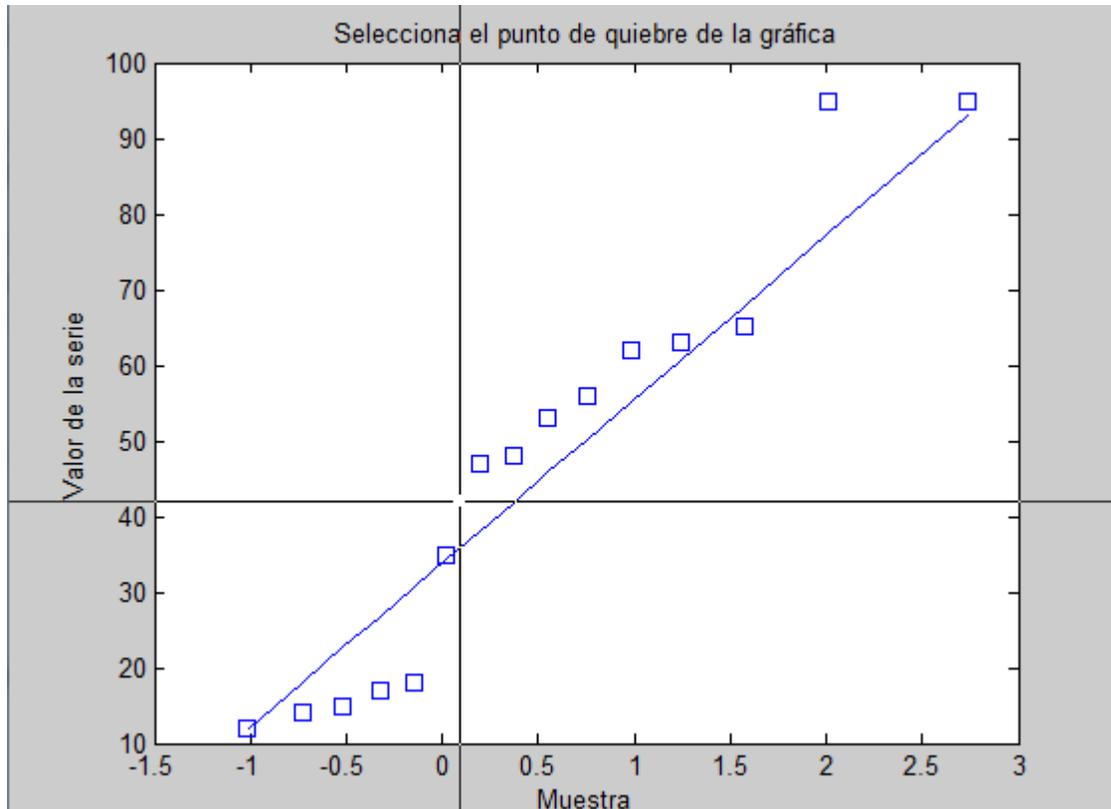


Figura 4-22 Rompimiento gráfico de serie.

Una vez que la serie se partió en 2, se realiza un ajuste de frecuencias a cada parte de la muestra y se obtienen os distintos parámetros de ajuste. Con estos nuevos parámetros de ajuste, se grafican las 2 nuevas distribuciones de probabilidad generadas (Figura 4-23 Serie combinada Gumbel-Gumbel).

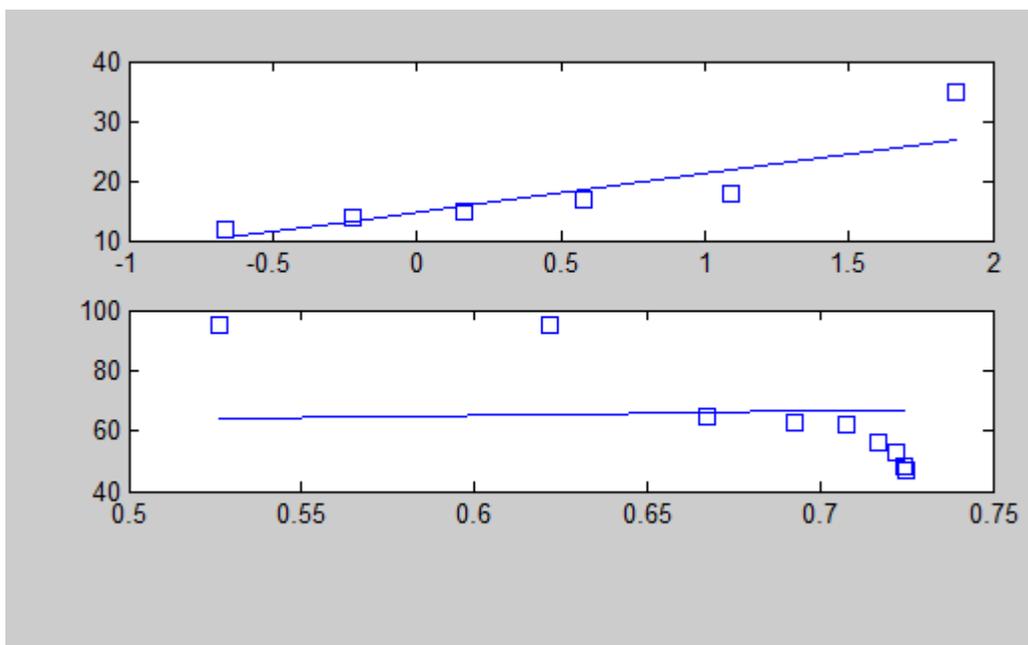


Figura 4-23 Serie combinada Gumbel-Gumbel

4.3. Módulo Interpolación Espacial

Este módulo analiza datos en tres dimensiones (2 variables independientes y una dependiente). Se realiza una triangulación de Delauney con los registros ingresados en la matriz de datos y se calculan las isolíneas para la variable dependiente.

En el ejemplo se utilizan los valores de la Tabla 4-8 Ejemplo Interpolación espacial.

Tabla 4-8 Ejemplo Interpolación espacial

X	Y	Z
5	4	11
3	7	16
1	4	9
4	9	12
2	8	14

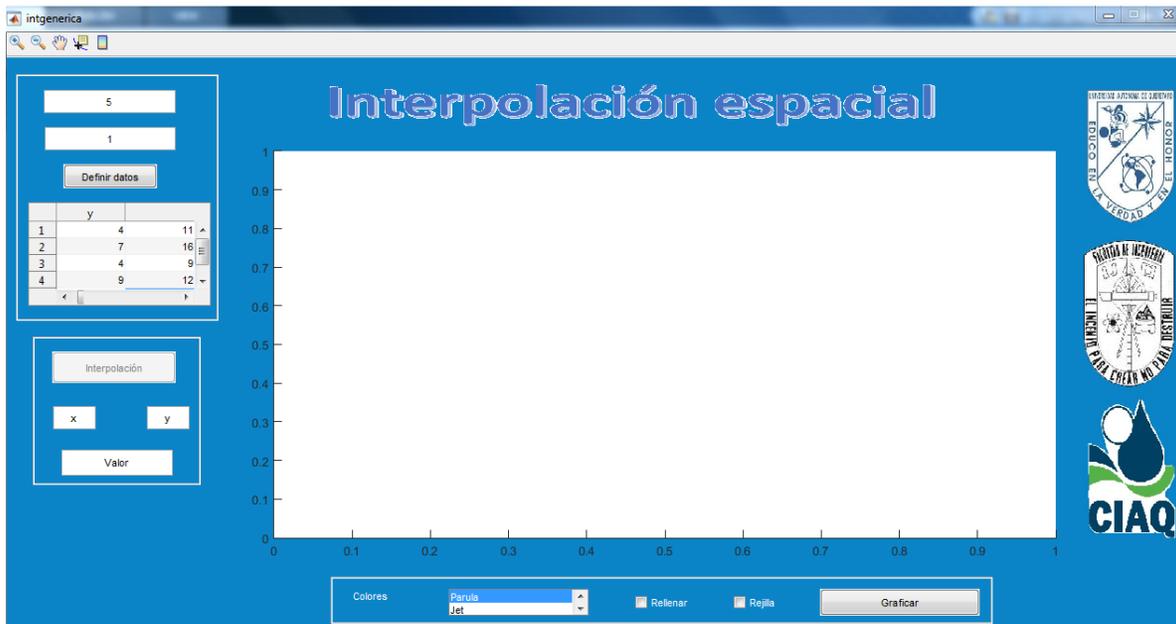


Figura 4-24 Interfaz gráfica Interpolación Espacial

Una vez capturados los datos, el programa dibuja el mapa de isolíneas (Figura 4-25 Mapa de isolíneas) con los datos ingresados al oprimir el botón Graficar.

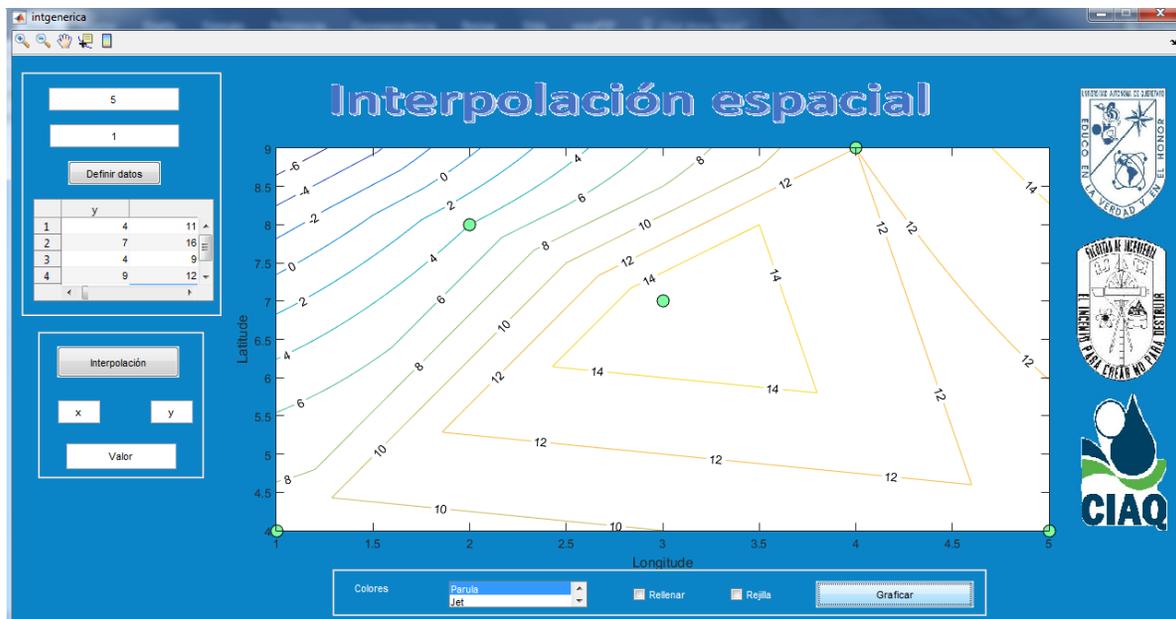


Figura 4-25 Mapa de isolíneas

Cabe señalar que la gráfica se modifica dependiendo de las características seleccionadas en la interfaz. Las características a editar es el relleno, rejilla y rampa de colores (Figura 4-26 Mapa de isolíneas con relleno sólido). La rampa de colores sigue un patrón determinado por Matlab y clasificado por nombres seleccionados por Mathworks. La rejilla consiste en líneas paralelas a los ejes que forman una cuadrícula y es apoyo visual para interpretar el gráfico.

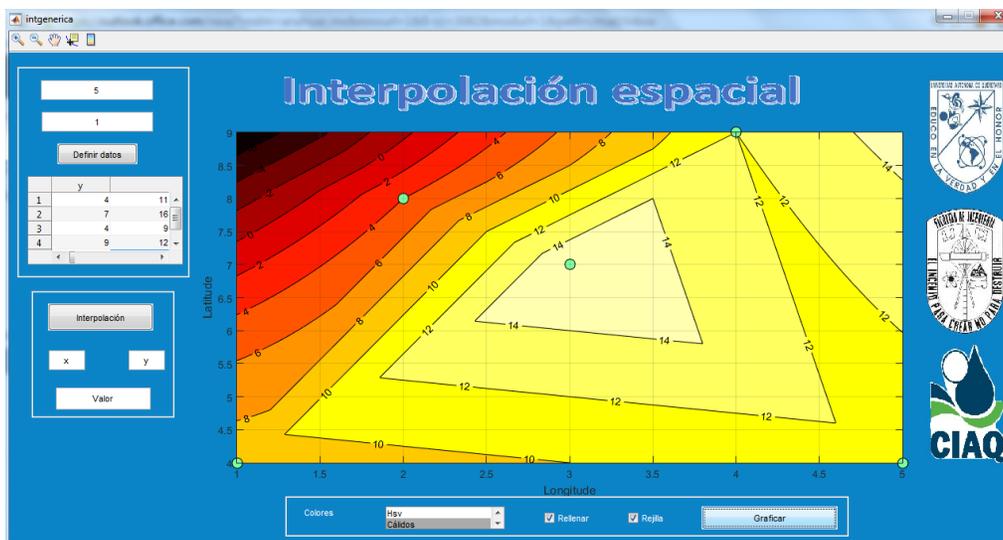


Figura 4-26 Mapa de isolíneas con relleno sólido

En la barra de herramientas de la interfaz hay 5 botones para iniciar herramientas gráficas, en orden de izquierda a derecha (Figura 4-27 Herramientas gráficas interpolación Especial) están: Acercamiento (Zoom in), Alejamiento (Zoom out), Desplazamiento gráfico (Pan), Cursor de datos (data cursor) y Rampa de colores (Colormap). Estas herramientas gráficas sirven para un análisis puntual en un área, es recomendable hacer zoom a la zona de interés y después mediante el cursor de datos, conocer los valores x,y,z del punto de interés.

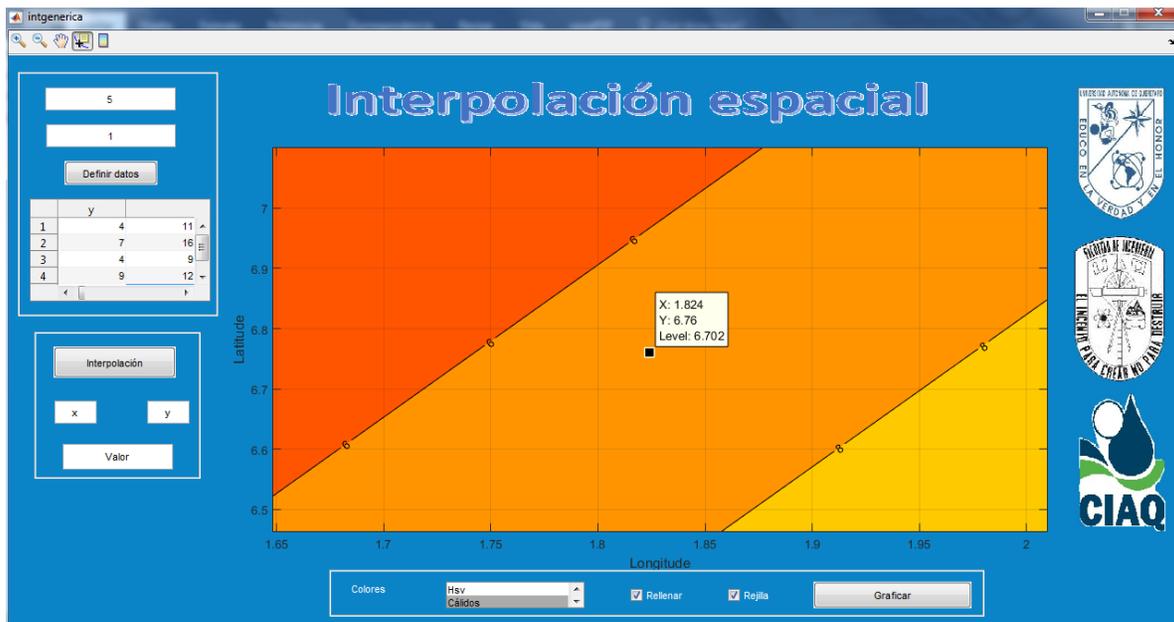


Figura 4-27 Herramientas gráficas interpolación Especial

De otra manera, el programa es capaz de interpolar y extrapolar espacialmente para conocer el valor de la función en un punto dado.

Para el ejemplo, se busca el valor de la función para cuando $x=2$ y $y=6$ utilizando los campos de ingreso de texto y además con el cursor de datos, ver Figura 4-28 Interpolación puntual espacial.

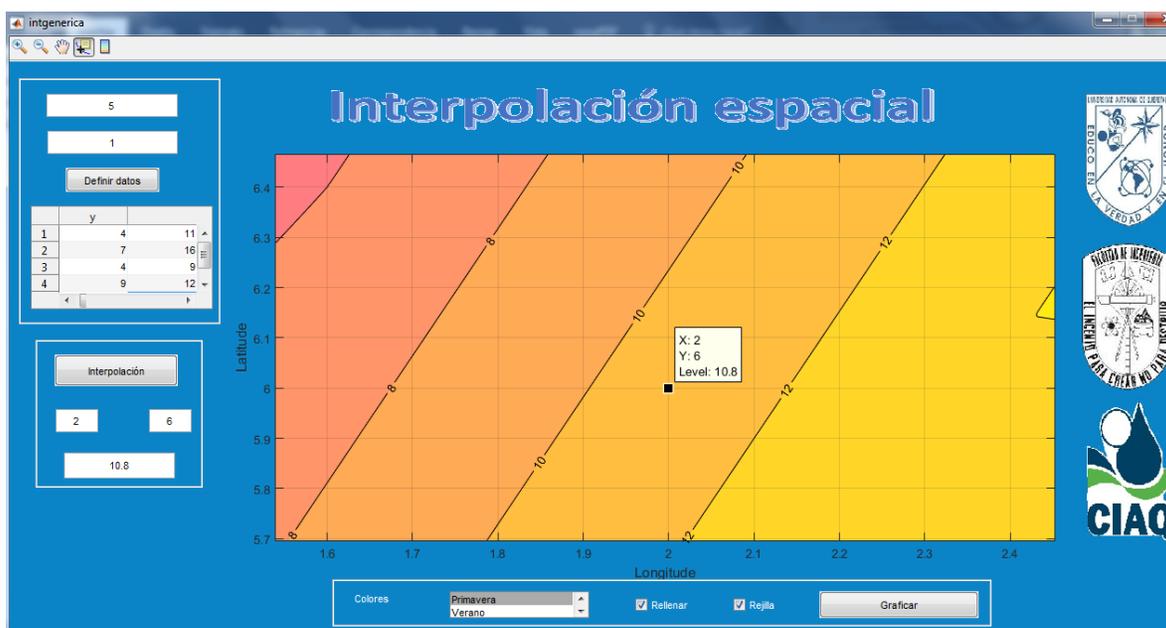


Figura 4-28 Interpolación puntual espacial

Para la validación de este programa, se realiza un mapa de Isolíneas sobre datos de Red CIAQ. A este mapa se le quita una estación y se busca ese dato mediante interpolación espacial Figura 4-29.

Los datos para este experimento son los presentados en la Tabla 4-9.

Tabla 4-9 Interpolación espacial en red CIAQ

Estacion	x	y	Temperatura	Radiación Solar
Chulavista	346746.47	2282052.99	25.9	957
Belén	353295.4	2284403.72	26.6	951
Real del Parque	353727.51	2279775.36	27.3	923
Candiles	353923.32	2272580.35	26.7	916
Cimatario	356546.01	2273989.37	25.4	842
Centro Histórico	355299.15	2277822.96	26.1	896
Milenio	359743.17	2278038.71	25.4	954
El Refugio	359094.12	2284372.02	24.7	940
Cerro de las Campanas	352844.61	2277483.98	26.2	NaN
San José el Alto	355528.08	2283897.8	24.8	849

:

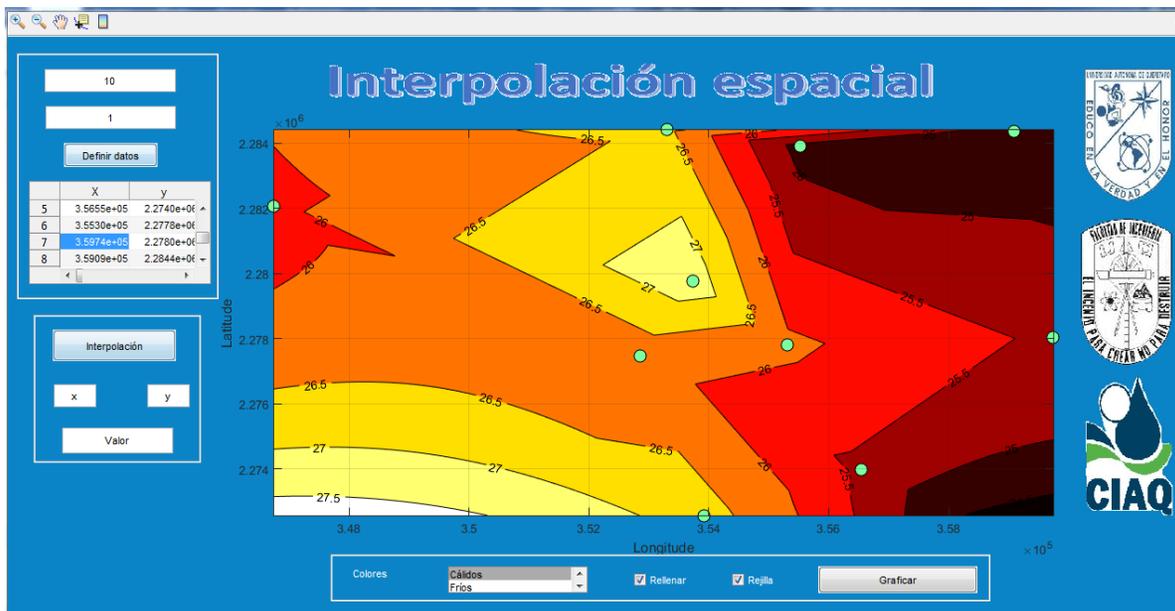


Figura 4-29 Interpolación espacial en red CIAQ

Los Resultados de la interpolación espacial se presentan en la Tabla 4-10 resultados interpolación espacial Red CIAQ y en la Figura 4-30.

Tabla 4-10 resultados interpolación espacial Red CIAQ

Estación	Temperatura	Radiación Solar	Temperatura Int	Radiación Int	Error relativo T	Error Rel. Rad.
Chulavista	25.9	957	25.75	956.23	0.57915058	0.08046
Belén	26.6	951	26.35	948.17	0.93984962	0.29758
Real del Parque	27.3	923	27.42	911.63	0.43956044	1.23185
Candiles	26.7	916	26.85	927.54	0.56179775	1.25983
Cimatario	25.4	842	25.32	895.68	0.31496063	6.3753
Centro Histórico	26.1	896	25.94	884.14	0.61302682	1.32366
Milenio	25.4	954	25.36	937.21	0.15748031	1.75996
El Refugio	24.7	940	24.61	947.14	0.36437247	0.75957
Cerro de las Campanas	26.2	NaN	26.26	925.07	0.22900763	##### #
San José el Alto	24.8	849	25.9583	823.47	4.67056452	3.00707

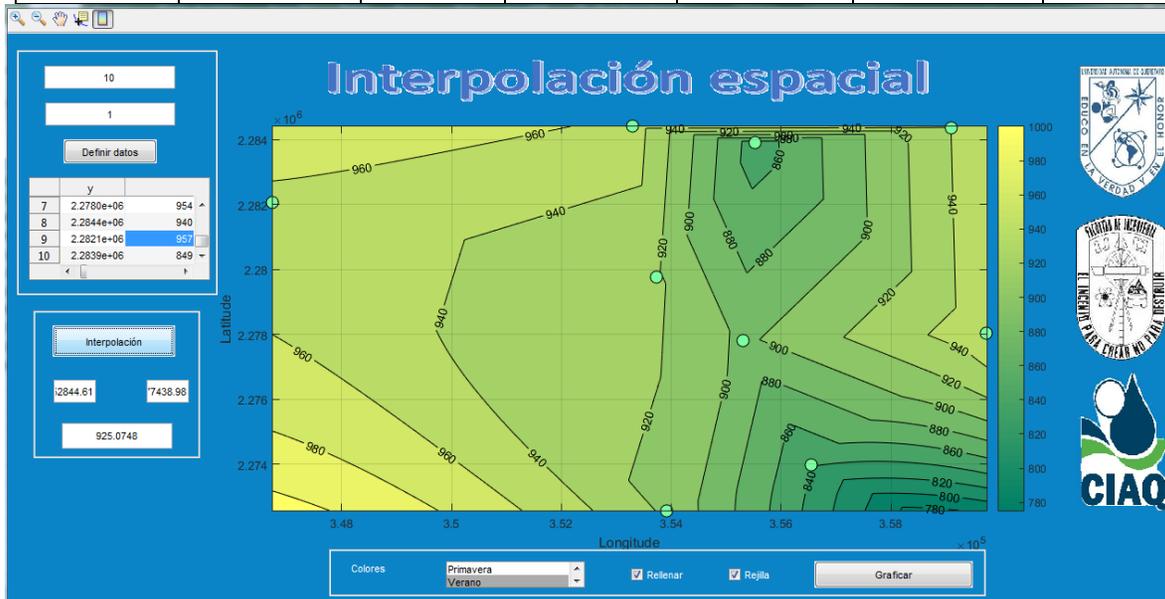


Figura 4-30 Mapa Isolíneas interpoladas en red CIAQ

4.4. Módulo de Lógica Difusa

El módulo de lógica difusa permite estimar valores mediante la defusificación por mínimo de máximos. El módulo es funcional para relaciones de dos variables independientes y una dependiente. La salida es un coeficiente entre 2 variables. En este módulo se pueden editar los límites de las funciones de pertenencia para cada subconjunto tanto de las variables independientes como de las variables dependientes. También es posible editar la base del conocimiento para establecer las relaciones entre los subconjuntos borrosos para cada variable independiente (Figura 4-31).

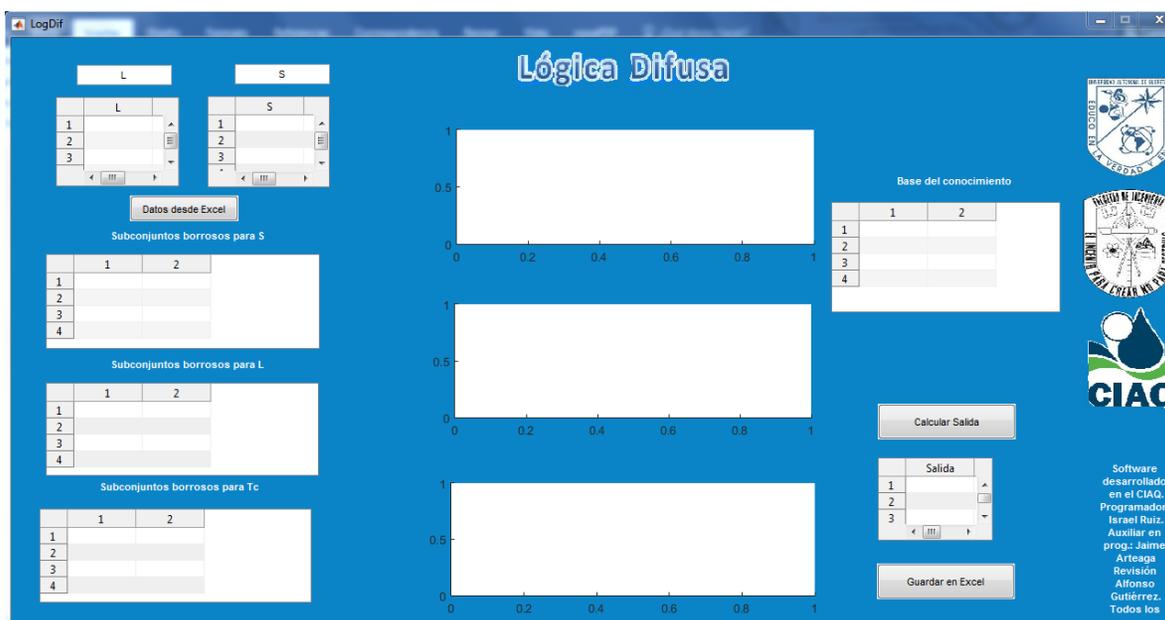


Figura 4-31 Módulo de Lógica Difusa

El modulo puede trabajar tanto con pares de números en las tablas L y S además de una hoja de Excel con múltiples parámetros para estimar el valor de la función. Si se selecciona datos desde Excel (Figura 4-32), se habilita una interfaz para lectura de archivo.

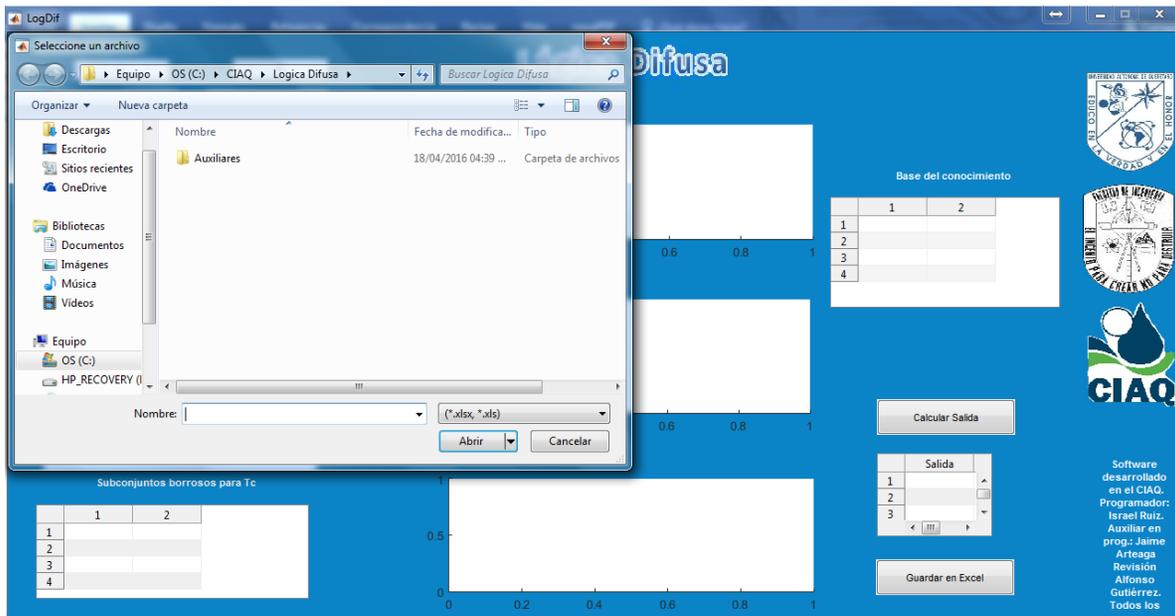


Figura 4-32 Cuadro de diálogo para lectura de archivo

Una vez seleccionado el archivo con el cual alimentar la plataforma, se cargan las 2 primeras columnas de la hoja de datos en las variables L y S. Después al oprimir Calcular salida, se ejecuta la rutina de Lógica Difusa para calcular la salida de la función mediante la asignación de una probabilidad basada en el área de las funciones de membresía determinadas en los subconjuntos borrosos.

La plataforma está calibrada para una relación de cociente entre 2 polinomios univariados, una función de L debe ser el numerador y mientras que para S hay una relación funcional en el denominador. Un ejemplo en hidrología de esta relación, es la ecuación de Kirpich para calcular el tiempo de concentración. De acuerdo a la ecuación empírica de Kirpich, el tiempo de concentración se calcula como:

$$T_c = 0.666 \left(\frac{L}{\sqrt{S}} \right)^{0.77} \quad (4.2)$$

Como se observa, hay una relación funcional como cociente de 2 variables. Si L crece, Tc crece y por el contrario si S es muy grande, el Tc va a ser pequeño. A partir de estas condiciones (base del conocimiento), el módulo de lógica difusa

puede replicar esta ecuación y calcular un Tiempo de concentración para distintas L y S (Figura 4-33).

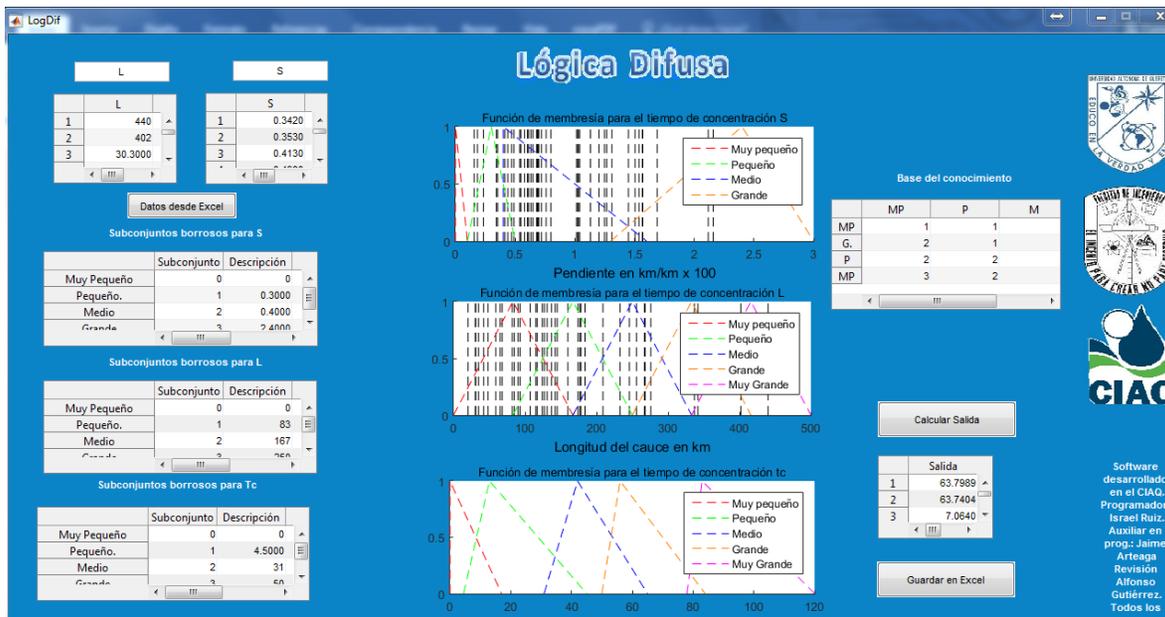


Figura 4-33 Corrida del módulo Lógica Difusa

En la tabla de salida, se despliega el resultado del cálculo para cualquier condición de L y S, está tabla es exportable, mediante el botón de Guardar datos en Excel a través de una interfaz gráfica de usuario.

Para la validación del modelo, se tomaron datos de L y S para 43 cuencas de México y se comparan los registros obtenidos por fórmula con los obtenidos por el modelo.

Tabla 4-11 Resultados tiempo de concentración Kirpich

Pendiente	Longitud	Kirpich	Modelo	Error Relativo
0.342	440	64.0828578	63.7989418	0.44304516
0.353	402	59.0532984	63.7404348	7.93712877
0.413	30.3	7.59369964	7.06397084	6.97589872
0.49	20.2	5.20329471	7.5134304	44.3975562
0.538	342	44.3335273	46.0474439	3.86596045
0.548	337	43.5239066	46.0541642	5.81348917
0.234	82	20.3409823	21.4747243	5.5736836

0.685	267	33.3853676	21.190383	36.5279326
1.028	184	21.4375789	22.2416764	3.75087834
1.449	131	14.4600878	7.93825879	45.1022781
0.72	246	30.7489565	21.2730994	30.8168412
1.041	209	23.5330095	22.2902263	5.28102139
1.565	137	14.5301526	7.52287605	48.2257603
1.537	127	13.8020243	7.62116936	44.7822348
1.249	145	16.5561817	23.129687	39.7042354
0.184	124	30.6805539	22.5011392	26.6599316
0.156	93	26.1976455	23.1942636	11.4643201
1.014	177	20.9169406	22.189974	6.08613552
1.13	131	15.9128166	22.6357441	42.2485073
1.504	90	10.6760273	7.73867007	27.5135794
2.159	35	4.48878173	6.86743922	52.9911597
0.76	107	15.8633767	6.05897436	61.8052671
0.69	68	11.6133716	5.93320283	48.9105918
0.698	65	11.1671489	5.94660807	46.7490932
0.613	31	6.63834138	7.03415142	5.96248395
0.603	276	35.9717187	21.4186471	40.4569817
0.632	268	34.5359238	21.1360889	38.7997002
0.649	256	32.9999989	21.1126492	36.022273
0.68	233	30.1461245	21.1791291	29.7451017
2.118	59	6.76017466	6.03419285	10.7390985
1.202	178	19.6763049	22.9305825	16.5390689
1.569	117	12.854987	7.50894292	41.587316
0.47	117	20.4469624	6.3275491	69.053843
0.438	85	16.4273837	5.67217465	65.4712232
0.438	85	16.4273837	5.67217465	65.4712232
1.024	48	7.62910567	6.67473684	12.5095768
0.799	31	5.99447779	7.03415142	17.3438566
1.267	160	17.7617908	23.2072369	30.6582044
1.131	115	14.3892313	6.9734359	51.5371201
1.31	174	18.7049621	23.3952541	25.0751213
1.689	142	14.504674	7.10465623	51.0181598
0.579	42	8.57341936	6.59220754	23.1087707

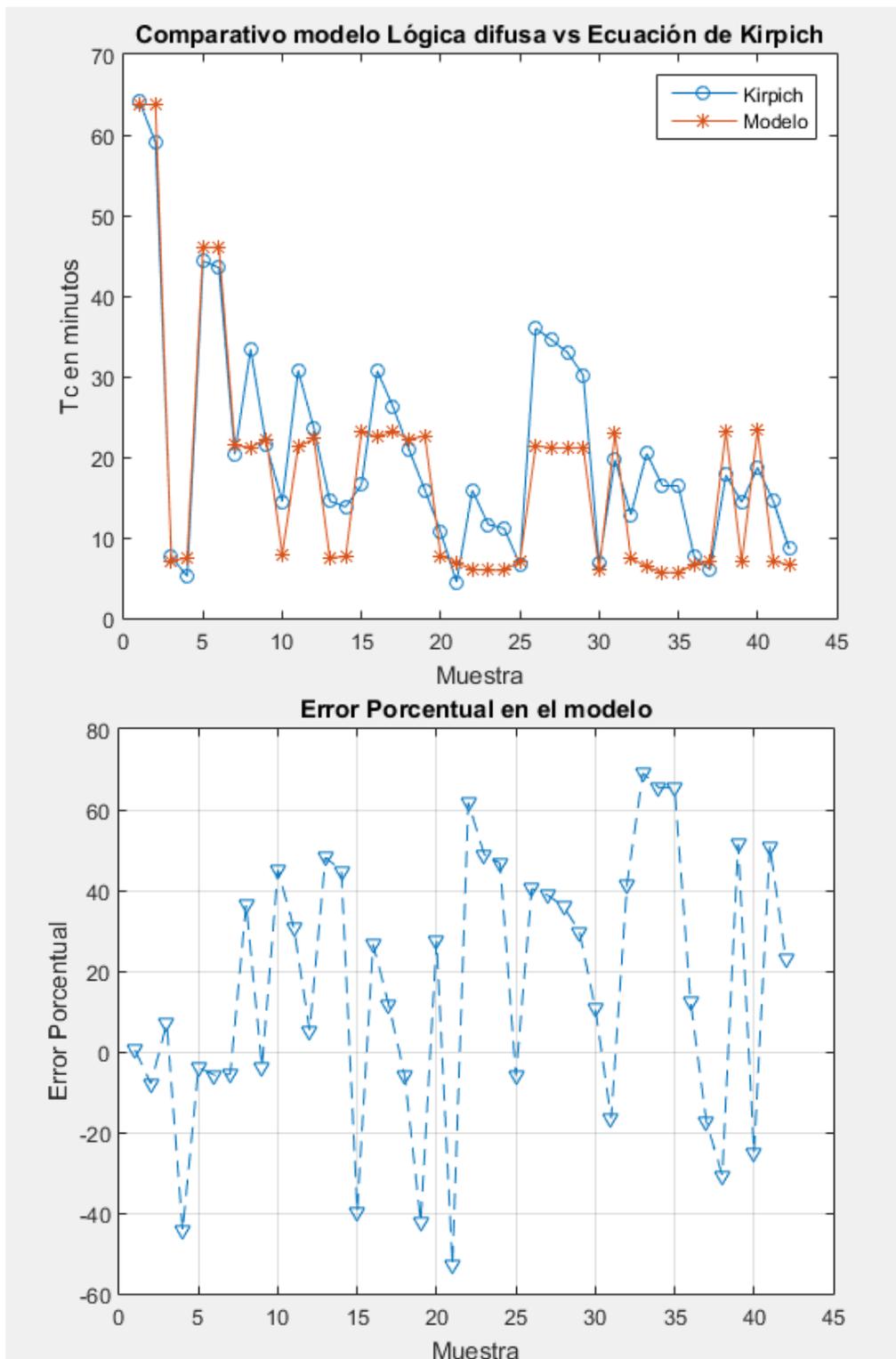


Figura 4-34 Gráfica comparativa del modelo de Lógica Difusa vs cálculo de concentración de Kirpich y el error porcentual del modelo.

4.5. Módulo Manejador de Red CIAQ

Como ya se señaló, Red CIAQ genera información climatológica a cada minuto, esta información es presentada a tiempo real en el portal redciaq.uaq.mx; la plataforma es capaz de presentar tanto la información en tiempo real como la histórica. El manejador tiene la flexibilidad para analizar datos, descargar y graficar la información de Red CIAQ.

	1	2
Temperatura °C		
Temperatura mínima °C		
Temperatura máxima °C		
Sensación térmica °C		
Índice de calor °C		
Precipitación 1 min mm		
Precipitación 24 h mm		
Precipitación 12 h mm		
Precipitación 6 h mm		
Intensidad mm/h		
Dirección viento °		
Velocidad viento m/s		
Radiación solar		
Punto de rocío °C		
Húmedad relativa %		
Presión mb		

Fecha:

Estaciones

- 1 Chulavista
- 2 Belén
- 3 Real del Parque
- 4 Candiles
- 5 Cimatarío
- 6 Centro Histórico
- 7 Milenio
- 9 El Refugio
- 11 Cerro de las C...
- 12 Geociencias
- 31 Viñedos
- 33 San José el Alto
- 34 Santa Rosa
- 37 PIQ
- Otra
- Todas

Generar Tabla

Figura 4-35 Módulo Manejador de Red CIAQ

El programa es capaz de presentar la información de una sola estación o hasta de 15 al mismo tiempo, al habilitar el checkbox de cada estación, se presentan los últimos registros adquiridos en cada estación. Es importante señalar que si alguna estación pierda la comunicación con el servidor Red CIAQ, la rutina Json generadora de datos, se queda con el último registro recibido y es éste el que se envía al manejador CIAQ.

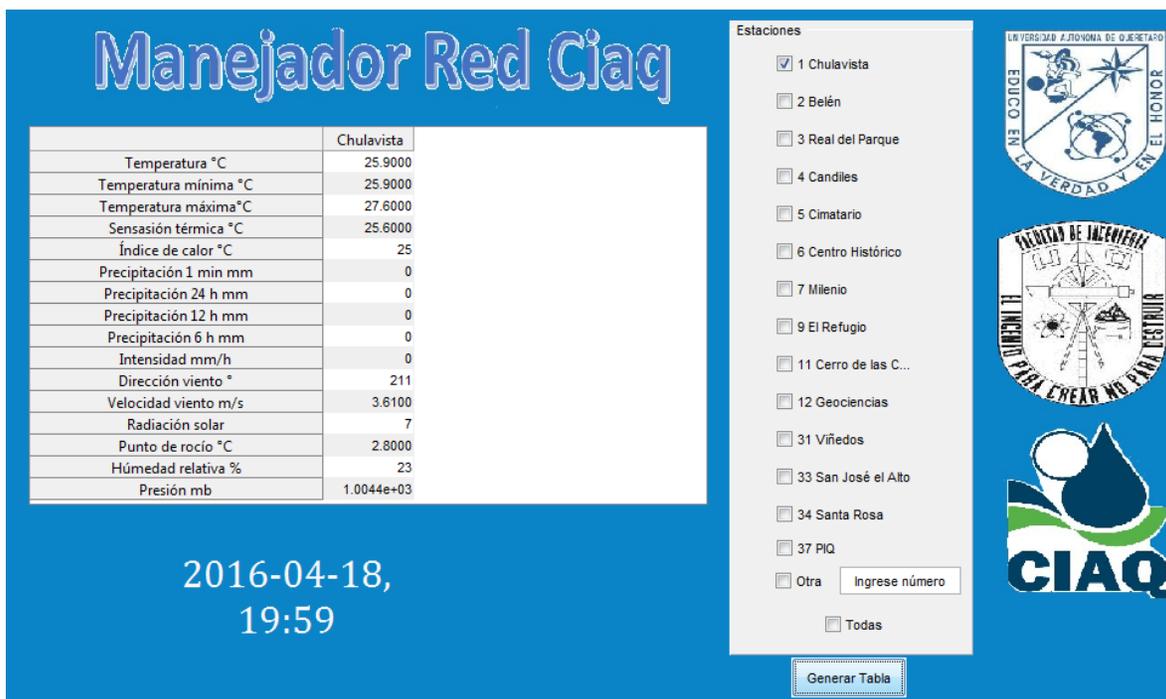


Figura 4-36 Presentación de datos de la Estación Chulavista el día 18 de abril de 2016 a las 19:59 h en el Manejador Red CIAQ.

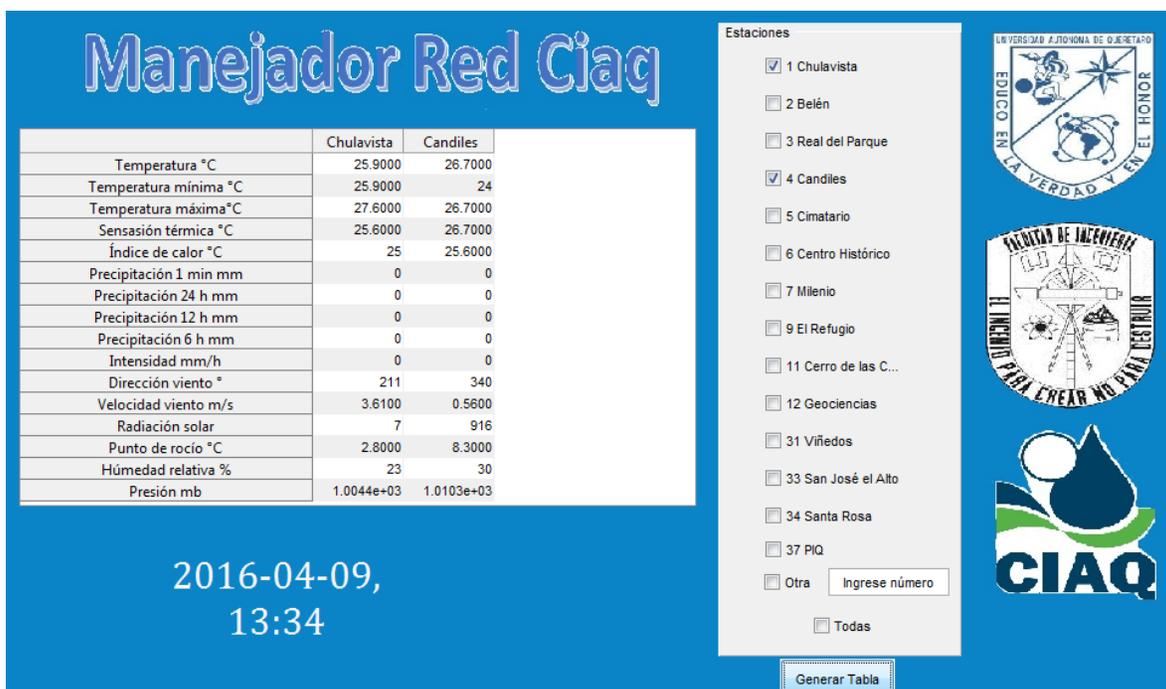


Figura 4-37 Presentación de datos de múltiples estaciones el día 9 de abril de 2016 a las 13:34 h en el Manejador Red CIAQ.

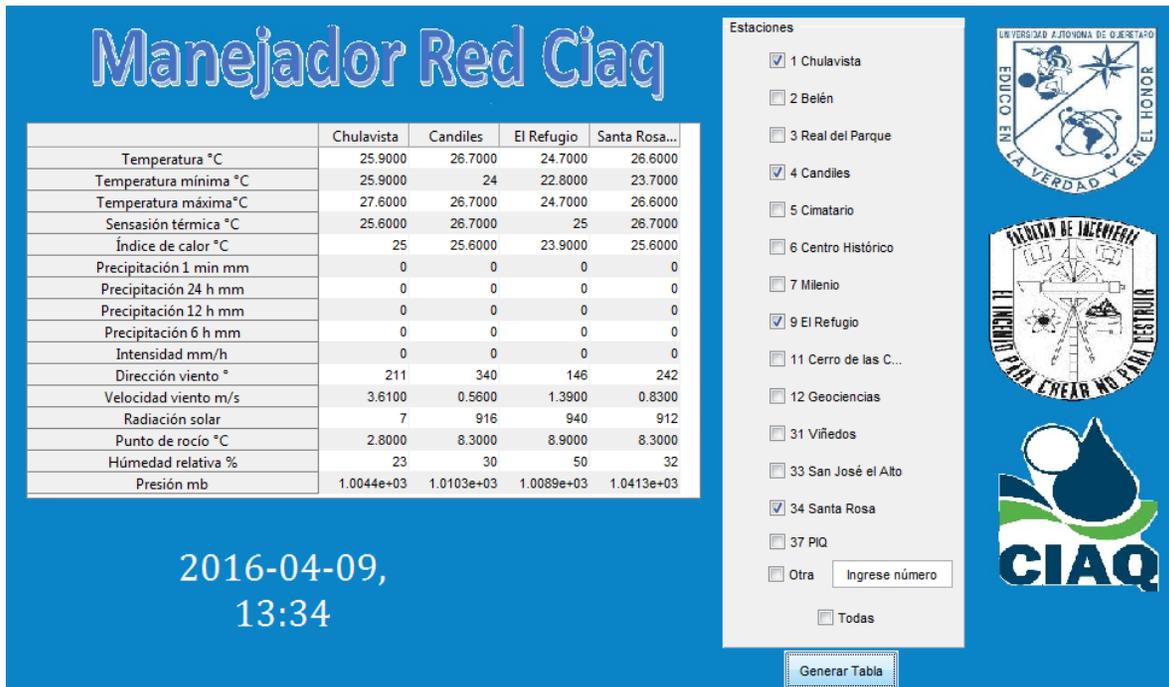


Figura 4-38 Estaciones en tiempo real en simultáneo

La interfaz manejo de datos permite la descarga histórica de registros desde Red CIAQ (Figura 4-39). Como la base de datos en línea está clasificada por mes y por estación, en cada corrida se puede descargar un mes completo de alguna estación.

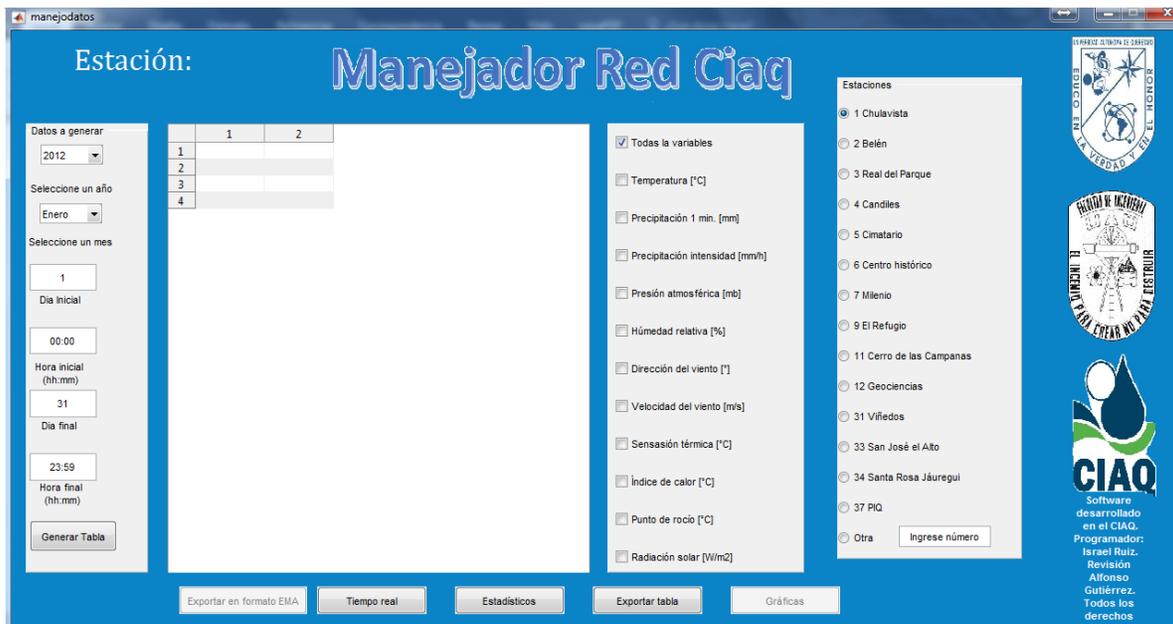


Figura 4-39 Interfaz para manejo gráfico de datos históricos

Esta interfaz permite una vez generados los datos exportarlos en dos formatos, uno es tal cual se presentan en pantalla y el llamado formato EMA coincide con el estándar utilizado por CONAGUA en sus EMAS. Además, desde esta pantalla es posible abrir 2 ventanas emergentes, una con los estadísticos principales de la tabla de datos y otra con los datos en tiempo real que presenta Red CIAQ. Es importante señalar que el panel de botones en esta interfaz es con botones de radio, esto quiere decir que sólo puede ser seleccionada una estación en cada corrida.

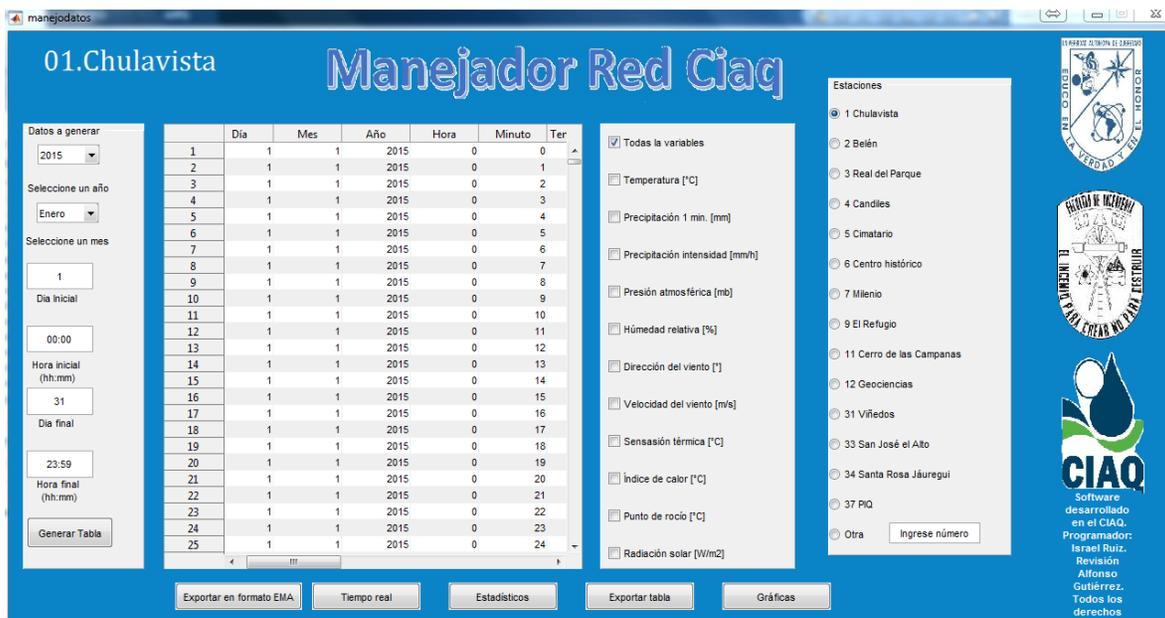


Figura 4-40 Datos históricos descargados desde Internet

Una vez generada la tabla de datos, se habilitan los botones de Exportar datos y Gráficas (Figura 4-40). Los dos botones de exportar, abren cuadros de diálogo para crear archivos .xlsx que contienen los registros de la tabla.

El botón Tiempo Real liga con la interfaz descrita como Manejador CIAQ en tiempo Real y el botón Estadísticos, muestra el resumen de los datos generados en la tabla.

Dentro del botón de Gráficas, se recuperan los datos generados y en un entorno gráfico, se visualizan las variables seleccionadas en diferentes colores (Figura 4-41).

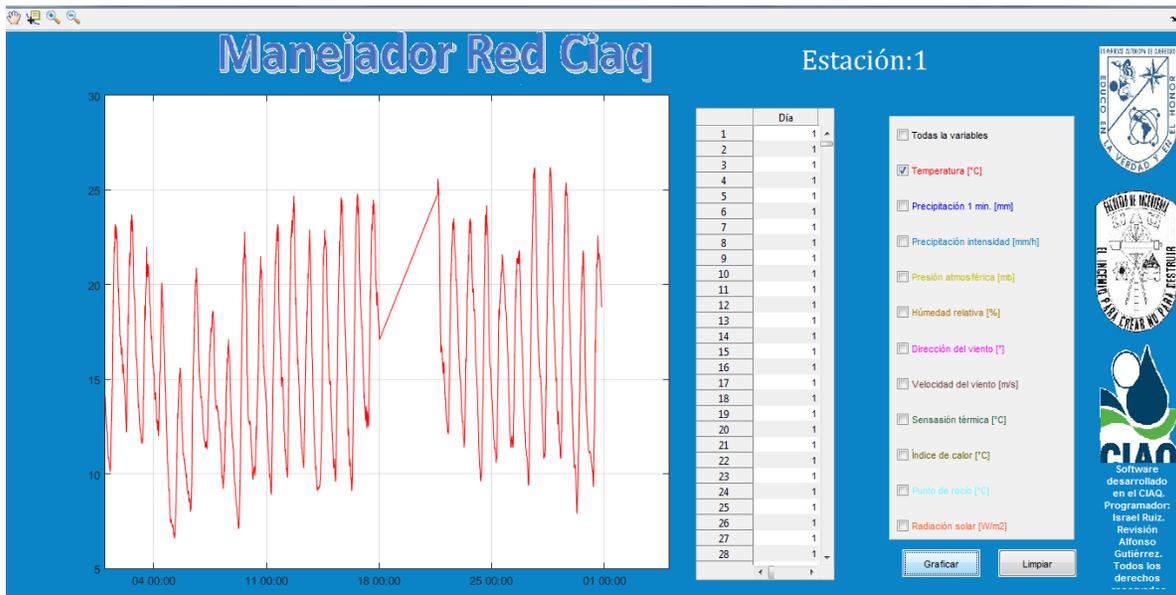


Figura 4-41 Gráficos del mes en Chulavista

Las gráficas pueden presentar desde una hasta todas las variables seleccionadas en los checkbox. Además, la barra de herramientas permite la manipulación de la visualización del gráfico (ver Figura 4-42) y la extracción de datos mediante el cursor de datos.

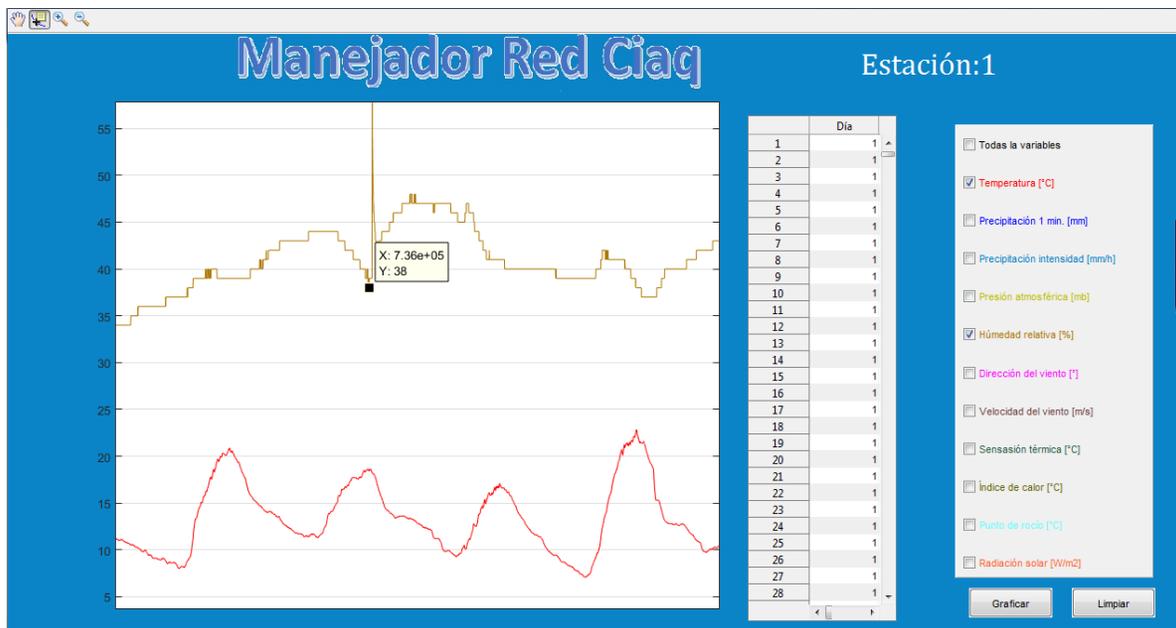


Figura 4-42 Datos gráficos en simultáneo para la estación Chulavista

4.6. Módulo Series de Tiempo

El módulo Series de tiempo calcula y evalúa modelos autorregresivos a una serie de datos. Los modelos que calcula este módulo son AR de orden 1,2 y 3. El módulo determina el mejor ajuste a partir del criterio de Akaike.

Para utilizar este módulo, se abre la interfaz Entrada (Figura 4-43). En esta interfaz se utiliza el cuadro de dialogo de Cargar archivo para enviar datos desde Excel a la plataforma. Los datos con los que trabaja el módulo deben de ser una columna en la posición Columna B de Excel. No hay un límite en la longitud de éste vector columna; sin embargo, el tiempo de ejecución depende completamente del tamaño de vector.



Figura 4-43 Interfaz de bienvenida para módulo Series de Tiempo

Una vez cargado el archivo a analizar, Hay que utilizar alguno de los 4 botones de entrada: Normalización, Estandarización, Generación de Muestras y Pronóstico.

Cada uno de éstos botones abre interfaces gráficas para el desarrollo de las metodologías para encontrar los modelos AR que se ajustan a los datos. Dentro del botón Normalización, al cargar los datos (Previamente obtenidos) y oprimir el botón Corrida de datos, Se calculan los estadísticos sesgados e insesgados de la muestra. A partir de estos estadísticos se calculan también los correlogramas (Figura 4-44). En esta interfaz (Figura 4-45), además, al presionar el botón de normalización, se puede redefinir a la muestra en base a los criterios de logaritmos, Box y Cox (Figura 4-46) y doble potencia (Figura 4-47).

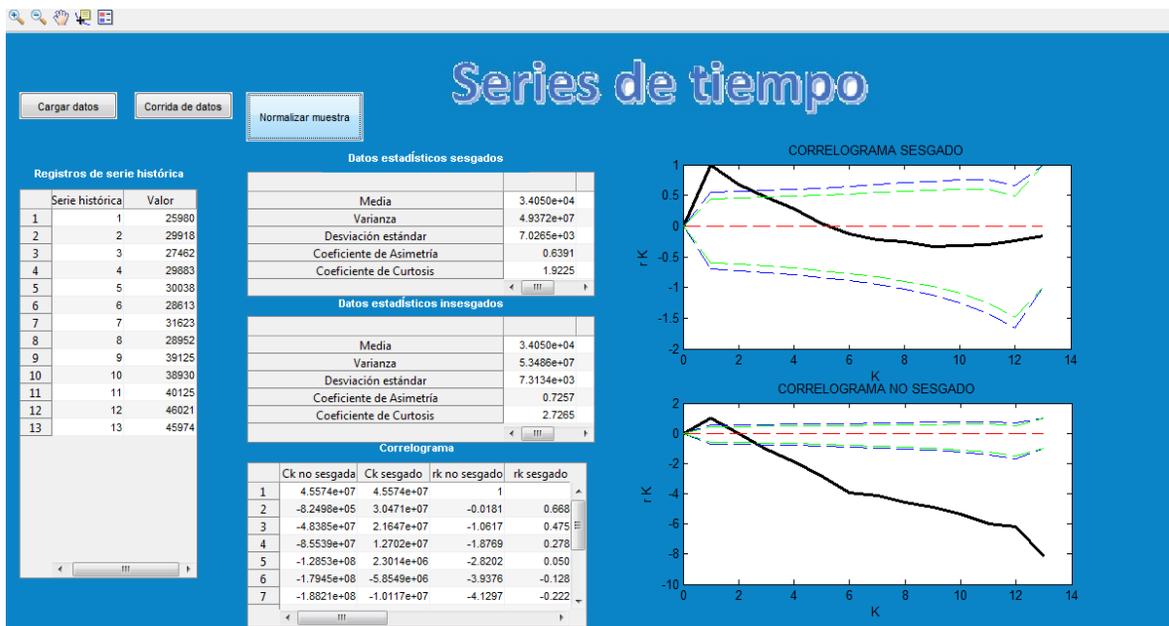


Figura 4-44 Correlogramas sesgados e insesgados

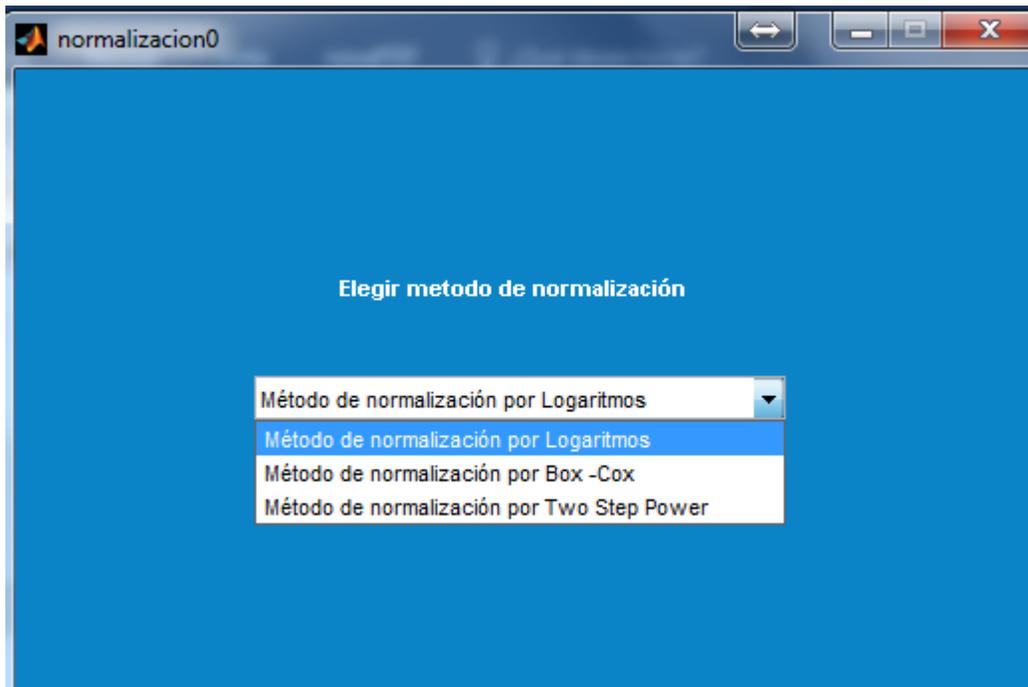


Figura 4-45 Métodos de normalización de muestras



Figura 4-46 Normalización por el estadístico de Box y Cox



Figura 4-47 Normalización por doble potencia

El segundo botón (estandarización) calcula los parámetros ϕ de los modelos AR de orden 1 a 3 de acuerdo a la selección en el menú indicado y el coeficiente de error medido con el criterio de Akaike. Los estadísticos (Figura 4-48, Figura 4-49, Figura 4-50, Figura 4-51) son retomados de la interfaz anterior y son los correspondientes a la muestra sin normalizar.

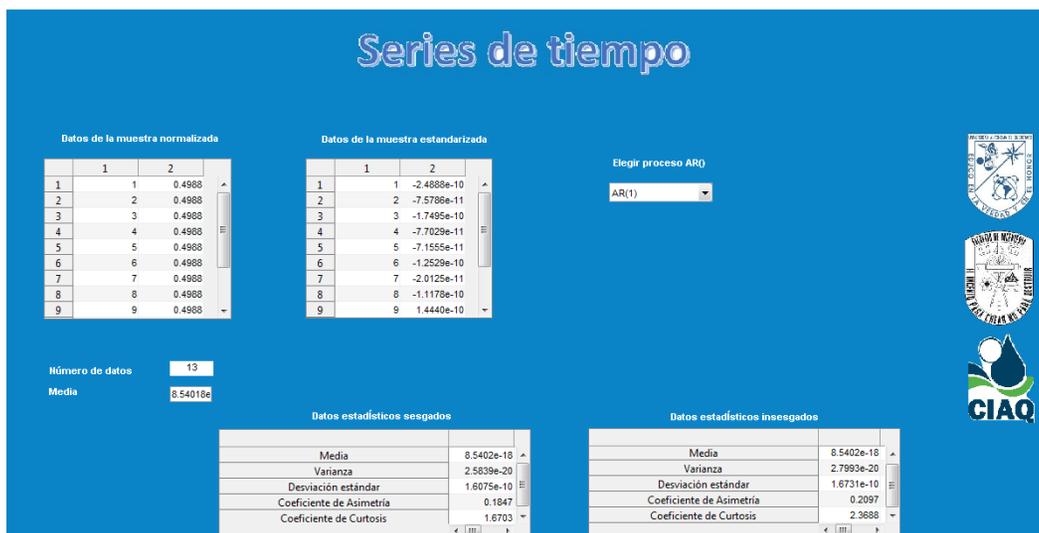


Figura 4-48 Estadísticos de la muestra.



Figura 4-49 Tablas de coeficientes phi y valor de criterio de Akaike Modelo AR(1)



Figura 4-50 Coeficientes phi modelo AR(2)



Figura 4-51 Coeficientes modelo AR(3)

El botón Generación de muestras evalúa los modelos AR (previamente seleccionado el indicado) y estima muestras sintéticas, hay que recordar que los modelos AR tienen un componente estocástico por lo que cada muestra sintética es distinta a la anterior. Las muestras sintéticas utilizan a la muestra normalizada por alguno de los tres métodos (Logaritmos, Box-Cox y doble potencia). La Figura 4-52 presenta el comportamiento de 11 muestras sintéticas para la serie normalizada y el proceso AR(2) con el estadístico de Box y Cox.

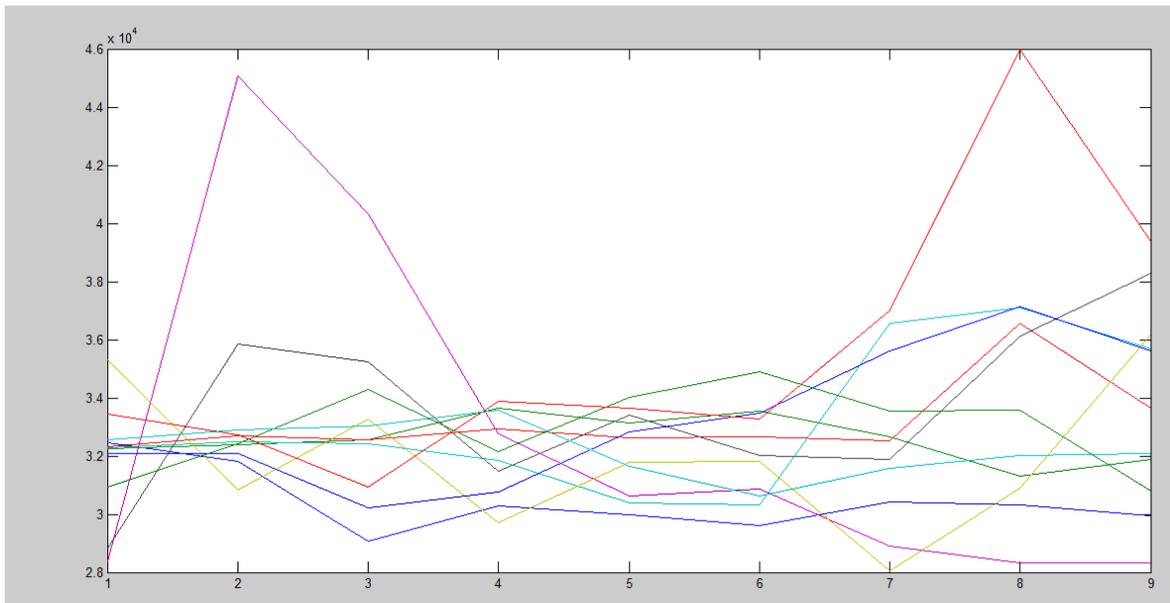


Figura 4-52 Muestras sintéticas por procesos AR

Series de tiempo

Número de muestras a generar

11

Generar

Muestras generadas

	1	2	3	4	5	6
3.1797e+04	3.2387e+04	3.2681e+04	3.2491e+04	4.5061e+04	3.0840e+04	3.3272e+04
2.9078e+04	3.2543e+04	3.2563e+04	3.2438e+04	4.0313e+04	3.3272e+04	3.3272e+04
3.0305e+04	3.3634e+04	3.2932e+04	3.1836e+04	3.2760e+04	2.9706e+04	3.3272e+04
2.9981e+04	3.3126e+04	3.2623e+04	3.0377e+04	3.0643e+04	3.1777e+04	3.3272e+04
2.9598e+04	3.3532e+04	3.2667e+04	3.0314e+04	3.0860e+04	3.1797e+04	3.3272e+04
3.0416e+04	3.2646e+04	3.2529e+04	3.6562e+04	2.8883e+04	2.8037e+04	3.3272e+04
3.0326e+04	3.1291e+04	3.6546e+04	3.7118e+04	2.8315e+04	3.0882e+04	3.3272e+04
2.9964e+04	3.1889e+04	3.3651e+04	3.5676e+04	2.8322e+04	3.6176e+04	3.3272e+04

Datos estadísticos sesgados

	1	2	3	4	5	6
Media	3.0192e+04	3.2675e+04	3.3003e+04	3.3059e+04	3.1625e+04	3.1517e+04
Varianza	1.4274e+06	5.2784e+05	1.8349e+06	4.6056e+06	2.7090e+07	9.8677e+06
Desv. Estandar	1.1948e+03	726.5245	1.3546e+03	2.1461e+03	5.2048e+03	3.1413e+03
Coef. de asimetría	0.2528	-0.2274	1.2681	0.7191	1.7658	-0.2274

Datos estadísticos inesgados

	1	2	3	4	5	6
Media	3.0192e+04	3.2675e+04	3.3003e+04	3.3059e+04	3.1625e+04	3.1517e+04
Varianza	1.5464e+06	5.7182e+05	1.9878e+06	4.9894e+06	2.9347e+07	1.0690e+07
Desv. Estandar	1.2435e+03	756.1907	1.4099e+03	2.2337e+03	5.4173e+03	3.2696e+03
Coef. de asimetría	0.2871	-0.2582	1.4399	0.8165	2.0050	-0.2582

Figura 4-53 Interfaz de generación de muestras

Finalmente, los modelos Autorregresivos permiten el pronóstico de datos a futuro, en el Botón pronóstico, se utiliza alguno de los procesos AR previamente calculados y se estima el valor de la muestra en N pasos subsecuentes.

Series de tiempo

Número de años a pronosticar: 11

Número de datos de la serie histórica: 13

Pronosticar

	AR(1)	AR(2)	AR(3)
1	3.6479e+04	3.6423e+04	3.6717e+04
2	3.5845e+04	3.3375e+04	3.3441e+04
3	3.4914e+04	3.2759e+04	3.3556e+04
4	3.5372e+04	3.4498e+04	3.3908e+04
5	3.6117e+04	3.5603e+04	3.5486e+04
6	3.3701e+04	3.2854e+04	3.3519e+04
7	3.4817e+04	3.4601e+04	3.4658e+04
8	3.4599e+04	3.4393e+04	3.4182e+04
9	3.6980e+04	3.6967e+04	3.7357e+04
10	3.4302e+04	3.4111e+04	3.4420e+04
11	3.2996e+04	3.2604e+04	3.3896e+04
12	3.3464e+04	3.3215e+04	3.3505e+04
13	3.5626e+04	3.5525e+04	3.5824e+04
14	3.6009e+04	3.5878e+04	3.6236e+04

Notese que se incluye la generación de una muestra del mismo tamaño que la muestra real

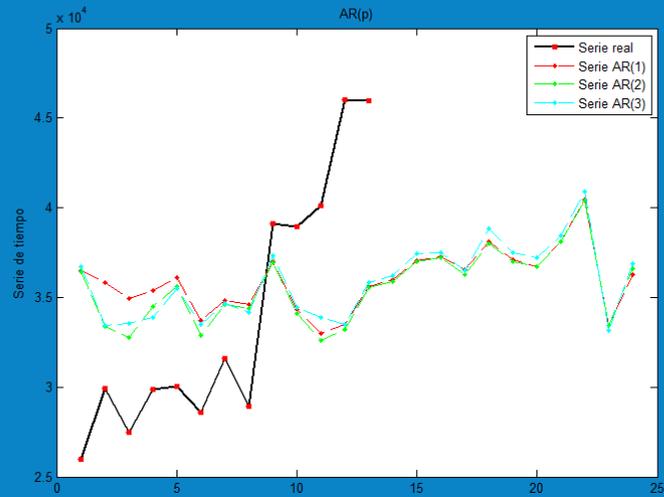


Figura 4-54 Caminata hacia el futuro con un modelo AR

5. CONCLUSIONES Y RECOMENDACIONES

Para obtener un grado de doctorado se necesita dominar un tema y desarrollar conocimiento en esta disciplina. Podría definir a la especialidad de este trabajo como hidroinformática. De esta disciplina se sabe que la automatización, informática y el análisis numérico, fortalecen y permiten entender mejor a los procesos hidrológicos. El estudio de la hidrología permite entender el mundo físico a través de números y ecuaciones; además la informática permite la solución de problemas numéricos complejos en segundos. Es así que la hidroinformática facilita la comprensión del mundo físico y reduce los tiempos de cálculo para la modelación hidrológica.

El análisis numérico busca resolver mediante operaciones aritméticas y booleanas, metodologías matemáticas más avanzadas. Es así que software como Matlab, se vuelve imprescindible en la ingeniería. Al iniciar con esta tesis, ya existía un expertiz importante en la clase de métodos numéricos en Matlab, por lo que realizar un doctorado con estas herramientas fue un paso natural en el desarrollo del conocimiento en este software. El tema de programación orientada a objetos era limitativo a estudiantes de informática o carreras basadas en la programación; sin embargo, con los becarios participantes en esta tesis (la mayoría ingenieros civiles), entendí que las herramientas matemáticas (básicas en todo ingeniero y universales) evolucionan a herramientas de análisis numérico y teniendo un objetivo claro y accesible, alumnos no especializados en software, se vuelven expertos programadores.

Desde el punto de vista de automatización, esta tesis obligó el trabajo en equipo para la instalación y puesta en marcha de estaciones meteorológicas. En esta etapa el equipo de Red CIAQ debió capacitarse en los componentes de una estación climatológica, la instalación, mantenimiento e interpretación de registros. Hubo una evolución tangible en base al trabajo del equipo. Como inicio de la tesis, se pensaron en 6 estaciones en la ciudad de Querétaro, éstas se instalaron y

controlaron a tiempo real, lo que permitió que en base a la publicidad y al prestigio adquirido, se sumarán instancias gubernamentales como CONAGUA o Protección Civil a Red CIAQ. En los 4 años de duración del doctorado, Red CIAQ pasó de 6 estaciones a 45 a lo largo de todo el estado. La automatización de estaciones ha permitido generar informa y reportes muy acertados sobre la actualidad del estado; sin embargo, no se ha encontrado una completa autosuficiencia. Red CIAQ es sensible al interés del responsable de la estación, a un adecuado internet y a la toma de energía eléctrica. Desafortunadamente los presupuestos no dan para tener personas de planta monitoreando el status de la red, es necesario establecer responsables de monitoreo, manejo de datos y mantenimiento a software desde el inicio. Una red meteorológica es un escalón para tener investigaciones en áreas hidrológicas y climáticas por lo cual es una excelente forma de vinculación y publicidad a proyectos de investigación por lo que una correcta instalación y manejo de datos, es una muy buena tarjeta de presentación.

La plataforma hidroinformática recopila y genera conocimiento en diversas disciplinas; en cuanto a curvas IDT, el software muy rápido carga mapas y presenta las lluvias de diseño para 3 períodos de retorno previamente calculados. Como se recurre a una base de datos, el programa es instantáneo, no así la interpolación; en el proceso de crear la triangulación de Delauney, y evaluar este modelo pasan casi 40 segundos en una computadora con altas prestaciones (Procesador XEO de 4 núcleos a 3.7 Ghz y 8 Gb de RAM). Si el programa se ejecuta en computadores con menores recursos, se observaron corridas de hasta 5 minutos. Desafortunadamente el programa no cuenta con una barra de estado o avisos del progreso de la interpolación y para un usuario nuevo esto facilita la creencia errónea de un malfuncionamiento en el programa. De las gráficas comparativas de errores y registros del módulo con las curvas SCT, se observa variación hasta del 70%; sin embargo, no hay una validación de los resultados de SCT, las isoyetas presentadas por ellos pueden o no ser verídicas.

El módulo de Análisis de frecuencias puede ser ampliado con nuevas distribuciones de probabilidad y la estimación de parámetros puede realizarse además del método de momentos por máxima verosimilitud (Matlab ya tiene herramientas para estos analisis). En cuanto a la competencia del modelo, se encuentra una ventaja crucial sobre AX (software apoyado por CONAGUA), el programa oficial solo corre en entornos de 32 bits y no ha sido actualizado de manera satisfactoria; mientras que en esta tesis se cuenta con versiones para sistemas de 32 y 64 bits. Además, al correr bajo entorno Matlab, puede ser manipulado tanto en Windows como en Linux y Mac. El programa EasyFit ofrece una demo funcional con las capacidades de ajustar funciones univariadas solamente. Además, al ser programado como software a la medida, se pueden incorporar botones, gráficos y estimadores según necesidades del cliente.

El módulo de Interpolación espacial, nace de una tesis de maestría que evaluaba diversos algoritmos de interpolación espacial. En la tesis se mostró que la triangulación de Delauney entregaba los menores errores en la estimación de datos faltantes en estaciones distribuidas heterogéneamente en el espacio. El módulo de esta tesis, corre instantáneamente si bien nunca se probó con más de 15 estaciones meteorológicas; tiene la ventaja de la particularización en escalas gráficas y en rampas de colores. A futuro se propone el crecimiento de este módulo con la incorporación de imágenes satelitales y porque no la capacidad de generar mapas y capas para sistemas de información geográfica.

El módulo de lógica difusa entregó resultados casi idénticos a la ecuación analizada; sin embargo, es limitada la aplicación; es necesaria una validación con ecuaciones con relaciones distintas a la de Kirpich. También con lógica difusa es posible analizar funciones de más de 2 variables y es algo que no se evaluó en este trabajo.

El manejador de Red CIAQ funciona a la perfección, inclusive hay 2 versiones para distintas transmisiones de datos. Red CIAQ presenta servicios web

que envían información tanto como texto csv como con cadenas Json. El módulo contiene interfaces de visualización para ambos casos y en el manejo de datos históricos, descarga instantáneamente la información de las bases de datos. En cuanto a manejo de gráficos, las herramientas de Matlab son las más avanzadas y populares en el mercado, por lo que al estar presentes en el módulo compite el visualizador con las mejores interfaces del mercado.

Finalmente, el módulo de series de tiempo tiene a su análogo en el programa MOSETI, se intentó validar los datos de ambos modelos pero esto es imposible, al tener los modelos AR un componente estocástico, cada ejecución va a entregar distintos resultados.

6. REFERENCIAS

- Abbot, M.B. 1991. Hydroinformatics: information technology and aquatic environment. Avebury, Brookfield.
- Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr. AC-19 (6): 716-723.
- Aparicio Mijares, Lafragua Contreras, Gutiérrez López, Mejía Sermeño y Aguilar Garduño. 2006. Evaluación de los recursos hídricos. Elaboración de balance hídrico integrado por cuencas hidrográficas. Documentos Técnicos del PHI-LAC, N° 4. Publicado por el Programa Hidrológico Internacional (PHI) de la Oficina Regional de Ciencia para América Latina y el Caribe de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, Montevideo, Uruguay.
- Banerjee, S., B. P. Carlin y A. E. Gelfand. 2004. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall/CRC Press, Taylor and Francis Group, USA.
- Bedient, P. B., y W. C. Huber. 2002. Hydrology and Floodplain Analysis (3a Ed.). Prentice Hall, Upper Saddle River, New Jersey.
- Benítez, A., 2002. Comunicación personal al equipo investigador. Facultad de Ciencias Forestales, Universidad de Talca, Santiago, Chile.
- Bernard, M. M. 1932. Formulas for rainfall intensities of long durations. Trans. ASCE. 96:592-624
- Box, G. E. P. y G. M. Jenkins. 1970. Time Series Analysis, Forecasting and Control. Ed. Holden-Day, San Francisco.
- Campos Aranda, D. F. 1999. Hacia el enfoque global en el análisis de frecuencia de crecientes. Ingeniería Hidráulica en México. XIV(1):23-42.
- Chow, V.T. 1994. Handbook of Applied Hydrology. McGraw-Hill, New York.
- Close, C. 1926. The Early Years of the Ordnance Survey, 1926. Reeditado por David y Charles. 1969. ISBN 0-7153-4477-3, 141-144 p.
- Cormen T., C. Leiserson, R. Rivest, C. Stein. 2009. Introduction to algorithms. The MIT Press, Cambridge, Massachusetts. ISBN 978-0-262-53305-8.
- Daniell, T. M., K. A. Daniell. 2006. Human impacts, complexity, variability and non-homogeneity: four dilemmas for the water resources modeller. Climate Variability and Change- Hydrological impacts, Proceedings of the fifth FRIEND Conference.

- Dershowitz, N. y Y. Gurevich. 2008. A natural axiomatization of computability and proof of Church's Thesis. *Bulletin of Symbolic Logic* 14. ISSN 10798986, 299-350.
- Garduño Velazco, H. 1990. La hidrología como ciencia de la tierra, *Revista interdisciplinaria de ciencia y tecnología del agua, Ingeniería Hidráulica en México*. Número especial: Octubre.
- Giraldo Gómez, N. D. 2006. NOTAS DE CLASE: Series de Tiempo con R. Edit. Centro de Documentación Rafael Botero, UN-Medellín. 164 p.
- Gutiérrez López, Lebel Thierry y Mejía Zermeño. 2005. Estudio espacio-temporal del régimen pluviométrico en la zona meridional de la República Mexicana. *Revista interdisciplinaria de ciencia y tecnología del agua, Ingeniería Hidráulica en México*. XX(1): Enero/Marzo.
- Hazen, A. 1930. Water supply. *American Civil Engineers Handbook*. John Wiley & Sons, New York, NY, 1444–1518 p.
- Heo, J. H., D. C. Boes y J. D. Salas. 2001. Regional Flood Frequency Analysis Based On A Weibull Model: Part 1. Estimation And Asymptotic Variances. *J. Hydrol.* 242(3-4): 157-170.
- Izenman, A. J. 2008. *Modern multivariate Statistical Techniques: Regression, Classification and Manifold learning*. Springer, New York. ISBN 978-0-387-78188-4
- Kinnison. 1985. *Applied Extreme Value Statistics*. Macmillan Publishing Company, Battelle Memorial Institute, USA.
- Koutsoyiannis, D., D. Kozonis, y A. Manetas. 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. *J. Hydrol.* 206: 118-135
- Maidment D. R. 1992. *Handbook of Hydrology*. McGraw-Hill, Estados Unidos.
- Mansell, R. S., S.A. Bloom y G. Sun. 2000. A model for wetland hydrology: description and validation. *Soil Science*. 165:384–397.
- Matheron, G. 1963. Principles of geostatistics. *Econ. Geol.* 58: 1246–1266.
- Matheron G. 1971. *The Theory of Regionalized Variables and its Applications*. Fontainebleau. Centre de Morphologie Mathématique de Fontainebleau.
- National Weather Service
2012:<http://w1.weather.gov/glossary/index.php?word=storm>

- Palacios Galindo Eugenio S. 2004. Modelo para manejo de la incertidumbre hidrológica en la planificación de la operación del SIC. Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas, Santiago de Chile.
- Price, R. 1997. Hydorinformatics. Society and the Market, Inaugural address at IHE Delft.
- Rajyalakshmi N.V. y S. Dutta. 2006. La regionalización de los procesos de precipitación y escurrimiento en la agricultura del arroz dominado las cuencas hidrográficas. *Water Sci. Tech.* 53(10):145-158.
- Ramírez, A. y A. Gutiérrez. 2005. Predicción Hidrológica mediante el método de la avenida índice para dos poblaciones. *Revista interdisciplinaria de ciencia y tecnología del agua, Ingeniería Hidráulica en México.* XX (2): Abril/Junio.
- Raudkivi A, J. 1979. *Hydrology an Advanced Introduction to Hydrological Processes and Modelling.* Pergamon Press, Great Britain, 479 p.
- SEMARNAT. 2002. Norma Oficial Mexicana NOM-011-CNA-2000, Conservación del Recurso Agua, Que Establece las Especificaciones y el Método para Determinar la Disponibilidad Media anual de las Aguas Nacionales. *Diario Oficial, México.*
- Sherman, C.W. 1931. Frequency and intensity of excessive rainfall at Boston. Mass. *Trans. Am. S.C.E.* 95: 951-960.
- Sherman, M. W. y J. P. Chesley. 1941. *Hydraulics of Steady Flow in Open Channels,* John Wiley & Sons, Inc., New York, 154-155 p.
- Skelton, R. A. 1958. *Cartografía, Historia de la Tecnología Oxford.* 6: 612-614.
- Smith, J. A. 1993. Precipitation. In: *Handbook of Hydrology.* McGraw-Hill, USA, 1-47 p.
- Stein, A., F. Van der Meer y B. Gorte (Ed.). 2002. *Spatial Statistics for remote sensing.* Springer, USA. ISBN 0-7923-5978-X
- Suárez Buenrostro, N. 2011. Caracterización de eventos extremos de precipitación, empleando distribuciones mezcladas. Tesis de Maestría. Universidad Autónoma de Querétaro, Querétaro.
- Témez, J. 1978. *Cálculo Hidrometeorológico de caudales máximos en pequeñas cuencas naturales.* Dirección General de Carreteras, Madrid, España. 111p.
- UNESCO. 2006. *Water a shared responsibility. The United Nations World Water Development Report 2.* UN, USA.

- Urquiola, A., S. Infante y F. Cedeño. 2010. Modelo Dinámico Espacio Temporal para Predecir Niveles de Precipitaciones en Venezuela. En: Chacón, R. et al. (Ed.) Modelos Computacionales en Ingeniería: Desarrollos Novedosos y Aplicaciones. ISBN 978-980-7161-02-2.
- Varas, E., y S. Lara. 1998. Métodos Regionales Para Estimar La Probabilidad De Frecuencia De Crecida. Ingeniería Del Agua. 5 (3): 51-57.
- Viessman, W. y G. L. Lewis. 2002. Introduction to Hydrology (5a Ed.). Pearson, USA.
- Willems, P. 2000. Compound intensity/duration/frequency-relationships of extreme precipitation for two season and two storm types. J. Hydrol. 233: 189-205.
- Woodcock C.E., A. H. Strahler, y D. L. B. Jupp. 1988. The use of variograms in remote sensing: Scene models and simulated images. Remote Sens. Environ. 25: 323–348.
- Yue, S. 2000. The Gumbel Mixed Model Applied To Storm Frequency Analysis, Water Resour. Manag. 14: 377-389.
- Zimmerman, D. A., G. de Marsily, C. A. Gotway, M. G. Marietta, C. L. Axness, R. L. Beauheim, R. L. Bras, J. Carrera, G. Dagan, P. B. Davies, D. P. Gallegos, A. Galli, J. Gómez-Hernández, P. Grindrod, A. L. Gutjahr, P. K. Kitanidis, A. M. Lavenue, D. McLaughlin, S. P. Neuman, B. S. RamaRao, C. Ravenne, y Y. Rubin. 1998. A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modelling advective transport by groundwater flow. Water Resour. Res. 34 (6): 1273-1413.