



**UNIVERSIDAD MICHOACANA  
DE  
SAN NICOLAS DE HIDALGO.**

**Facultad de Ingeniería Civil.**

**Cuerpo Académico Gestión Integral del Agua.**

**MODULO DE GESTION DE BASES DE DATOS DE ESTACIONES  
CLIMATOLOGICAS.**

**TESIS QUE PARA OBTENER EL TITULO DE INGERIERO CIVIL**

**PRESENTA:**

**P.I.C. Manuel Servin Arreygue.**

Director: Dra. Sonia Tatiana Sánchez Quispe.

**Morelia Mich., Agosto de 2010**



## *RESUMEN*

Actualmente existe un problema muy grave, en la gestión de los datos climatológicos, esto debido principalmente al tipo de gestores utilizados y la falta de actualización en los procedimientos de toma de datos y recopilación de los mismos. Uno de los principales problemas más frecuentes que tenemos con este tipo de gestores es el diseño de la interface que inter actúa con el usuario, requiere una especialización exhaustiva en los gestores para poder tener acceso a los datos climatológicos, y aun más complicado el actualizar los mismos. Al mismo tiempo existe un problema en los métodos de recolección de datos, en ellos pueden existir ciertos errores que pueden ser fácilmente detectables por métodos estadísticos.

Debido a la existencia de estos problemas se pretende realizar un modulo que gestione los datos climatológicos, obtenidos en las estaciones con registros diarios, adicionalmente realizar una revisión en la consistencia de los datos y llenado de los datos faltantes con algunos criterios, basado en la correlación de los datos en las estaciones.

Previendo estas necesidades se planteo realizar un gestor de los datos, utilizando como base primaria Autocad (en su lenguaje nativo, Autolisp), facilitando así el introducir las áreas de estudio y el manejo de algunos datos.

Para la revisión de la consistencia de datos, se evaluarán cambios en la media, cambios en la tendencia y pruebas de independendencia.

En base al nivel de correlación existente entre estaciones se propondrán varios métodos de llenado de datos faltantes.



## DEDICATORIAS



A mi madre por haberme dado la vida y por todos sus esfuerzos para que yo estudiara, y tuviera una formación profesional.

A mis hermanas que siempre estuvieron presentes en todos mis logros y metas propuestas.

A mis tíos Marcelo, Mario, Gerardo por su apoyo moral durante toda mi vida y la carrera y por sus sabios consejos que hicieron en mi un cambio de perspectiva de las cosas.



## AGRADECIMIENTOS

A la Dra. Sonia Tatiana Sánchez Quispe y el por su asesoría, comentarios y dirección mi más sincera gratitud.

A la UMSNH por albergarme como mi segunda casa 5 años para poder adquirir todos los conocimientos necesarios para comenzar mi vida profesional.

A los maestros de la facultad por transmitir sus conocimientos y vivencias.

A mis compañeros de facultad que siempre me apoyaron en toda la carrera.

A mis amigos(as) y compañeros que sin duda alguna son parte importante en este trabajo anímicamente hablando.



## INDICE

<b>indice</b> .....	<b>1</b>
<b>indice de tablas</b> .....	<b>3</b>
<b>indice de figuras</b> .....	<b>4</b>
<b>1.1 Introducción</b> .....	<b>1</b>
1.1.1 Breve historia del SMN.....	1
1.1.2 Descripción de la propuesta.....	2
1.1.3 PLANTEAMIENTO DE LA SOLUCION:.....	2
1.1.3.1 Objetivo General.....	2
1.1.3.2 Como soluciones particulares se plantearon procesos individuales:.....	2
1.1.4. METODOLOGIA UTILIZADA.....	3
<b>2.1 Revisión bibliográfica</b> .....	<b>8</b>
2.1.1 Antecedentes:.....	8
2.1.2 Marco teórico para la revisión de la calidad de los datos.....	10
Metodologías para la revisión de la consistencia en los datos.....	10
Pruebas básicas.....	10
Análisis con promedios móviles. (D.F Campos Aranda 2007).....	10
Prueba de las secuencias. (D.F Campos Aranda 2007).....	11
Test de Helmert. (D.F Campos Aranda 2007).....	11
Pruebas específicas.....	12
Evaluación de errores por cambios en la media.....	12
Evaluación de errores en la tendencia.....	14
Evaluación de errores en la independencia.....	17
Metodologías utilizadas en el llenado de datos y herramientas auxiliares.....	18
Correlación (Ronald E. Walpole Et.Al. 2000).....	18
Llenado de datos.....	19
METODO DEL U.S. WEATHER SERVICE. (D.F Campos Aranda 1998).....	19
Método de la relación normalizada. (D.F Campos Aranda 1998).....	20
2.1.3 Marco teórico para la programación del modulo.....	21
Autolisp.....	21
DCL (lenguaje de control de dialogos).....	22
Visual Basic.....	24
<b>3.1 MODULO DE GESTION</b> .....	<b>27</b>
3.1.1 El modulo de gestión tiene como requerimientos mínimos:.....	27
3.1.2 Como ejecutar el programa.....	27
3.1.3 Descripción del proceso.....	27



---

Selección adecuada de la zona horaria UTM .....	27
Primer interface (preselección y analisis superficial de estaciones) .....	28
Herramienta de gráficos .....	32
Opciones al graficar. ....	33
Segunda interface (revisión de los datos).....	36
<b>4.1 ANALISIS DE RESULTADOS Y REPORTE.....</b>	<b>43</b>
<b>Conclusiones.....</b>	<b>48</b>
<b>Recomendaciones.....</b>	<b>48</b>
<b>Referencias.....</b>	<b>49</b>



## INDICE DE TABLAS

TABLA 1, ELEMENTOS CLIMATOLÓGICOS CONTENIDOS POR EN EL CLICOM .....	8
TABLA 2, DATOS CLIMATOLÓGICOS CONTENIDOS EN ERIC.....	9
TABLA 3, RANGOS DE TOLERANCIA PARA EL NUMERO DE SECUENCIAS PERMISIBLE... 11	
TABLA 4, VALORES DE LA TOLERANCIA .....	13
TABLA 5, VALORES MÍNIMOS DE N .....	15
TABLA 6: DESCRIPCIÓN DETALLADA DEL ANÁLISIS DE LOS RESULTADOS Y LOS DATOS MOSTRADOS EN LOS REPORTES. ....	39
TABLA 7: RESULTADO DE LAS PRUEBAS ESTADÍSTICAS.....	44
TABLA 8: RESULTADO DE LAS PRUEBAS ESTADÍSTICAS ESTACIONES USADAS. ....	46
TABLA 9: RESULTADO DE LAS LA PERSISTENCIA DE ANDERSON.....	46
TABLA 10: MATRIZ DE CORRELACIONES.....	47



## INDICE DE FIGURAS

FIGURA 1: INTERFACE DE PROGRAMACIÓN VLISP .....	22
FIGURA 2, PRE VISUALIZACIÓN DEL DIALOGO EN DISEÑO.....	23
FIGURA 3, EDICIÓN DEL DIALOGO .....	24
FIGURA 4: ENTORNO DE PROGRAMACIÓN DE VISUAL BASIC.....	25
FIGURA 5. CUADRO DE DIALOGO INICIAL DE SELECCIÓN ZONA. ....	27
FIGURA 6, CUADRO DE DIALOGO DE AJUSTE DE DISTANCIA. ....	28
FIGURA 7 CUADRO DE DIALOGO INICIAL SOBRE POSICIÓN DE LAS ESTACIONES.....	29
FIGURA 8, ZONIFICACIÓN DE ÁREA DE DIBUJO .....	30
FIGURA 9, SIMBOLOGÍA FIGURA 4 .....	30
FIGURA 10, SELECCIÓN DE FILTROS DE BÚSQUEDA.....	31
FIGURA 11, DESPLIEGUE DE DATOS Y OPCIÓN DE SELECCIÓN. ....	31
FIGURA 12, RESUMEN DE DATOS Y BOTÓN DE ACCIÓN. ....	31
FIGURA 13, HERRAMIENTA DE GRÁFICOS.....	32
FIGURA 14: CUADRO DE DIALOGO REVISIÓN DE DATOS. ....	36
FIGURA 15: HERRAMIENTAS DE CONSISTENCIA. ....	37
FIGURA 16: HERRAMIENTAS DE LLENADO DE DATOS.....	37
FIGURA 17: HERRAMIENTAS DE GENERACIÓN DE REPORTE.....	38
FIGURA 18: GENERACIÓN DE REPORTE (PRUEBA DE SECUENCIAS).....	38
FIGURA 19: GENERACIÓN DE UNA TABLA (DATOS DIARIOS) .....	39
FIGURA 20: MAPA DE LOCALIZACIÓN DE LAS ESTACIONES. ....	43
FIGURA 21: MAPA DE LOCALIZACIÓN DE LAS ESTACIONES GENERADO CON EL PROGRAMA.....	44



# **CAPITULO I**

# **INTRODUCCION**

CIBERMITANIOS.COM



## 1.1 INTRODUCCIÓN

La anotación manuscrita de fenómenos meteorológicos se ha practicado por siglos, estas prácticas fueron fomentadas principalmente por los cambios climáticos continuos de la tierra, y la preocupación del ser humano por comprender los fenómenos climatológicos. Los primeros registros de fenómenos meteorológicos son de fechas de congelación y deshielo de ríos lagos y mares. (N. Plummer, Et.Al. 2007)

### 1.1.1 Breve historia del SMN

En México comenzaron los registros del comportamiento del clima en febrero de 1877, a cargo del observatorio meteorológico astronómico creado por decreto presidencial debido a la iniciativa del Secretario de Fomento, Vicente Riva Palacio, dependiente de la comisión de exploración geográfica del territorio mexicano. Durante el gobierno de Porfirio Díaz, y se instalo en la azotea de palacio nacional.

En 1878 el observatorio se traslado al castillo de Chapultepec, en 1880 el observatorio se independizo de la comisión geográfica y se dictaminó una partida especial en el Presupuesto de Egresos Nacionales.

En 1901 se forma el Servicio Meteorológico Nacional, el cual ya contaba con 31 secciones meteorológicas estatales, 18 observatorios y estaciones independientes, las cuales transmitían información al Observatorio Meteorológico de Tacubaya por vía telegráfica.

Durante el gobierno de Manuel Ávila Camacho (1940-1946), se creó la Secretaría de Recursos Hidráulicos, la cual incorporó al Servicio Meteorológico Nacional y lo denominó como Dirección de Geografía y Meteorología. En 1947, México firmó el Convenio de la Organización Meteorológica Mundial (OMM), Organismo especializado de las Naciones Unidas, encargado de la vigilancia del tiempo y del clima mundial. El Ingeniero Federico Peña, fungió como Director de Geografía y Meteorología de 1947 a 1960. Durante esos años, esa Dirección también se encargó de la cartografía del territorio nacional.

Para 1980, el Servicio Meteorológico Nacional contaba con una red de 72 observatorios, con 9 estaciones de radio sondeo, con más de 3000 estaciones climatológicas, 5 estaciones de radar meteorológico y un centro de Previsión del Golfo.

En 1989, al crearse la CNA, el Servicio Meteorológico se integró como una Subgerencia dependiente de la Subdirección General de Administración del Agua y, en 1990, se transformó en la actual Gerencia del Servicio Meteorológico Nacional pasando a formar parte de la Subdirección General Técnica de la CNA en 1995. (smn.conagua.gob.mx, CONAGUA)

En nuestra época la CONAGUA para la organización de los datos climatológicos en formatos digitales, o base de datos adopto un programa de gestión de datos climatológicos muy usado en varios países, llamado CLICOM, este programa viene equipado con una serie de herramientas muy completo, pero debido a las limitantes existentes en el sistema operativo el



cual tiene como base el MS-DOS, resulta muy compleja la utilización del mismo y por tanto una especialización para poder usarlo.

El problema principal de los datos climatológicos no es solo el proceso para extraerlos de las bases de datos que los contienen sino también la disponibilidad de los mismos.

El IMTA realizo una recopilación de varias estaciones climatológicas y realizo un programa mucho más sencillo en el proceso de interacción con el usuario y fácilmente por medio de las coordenadas geodésicas de nuestra cuenca podemos encontrar las estaciones más cercanas a nuestra área de estudio. Este programa solo nos permite realizar consultas y no tiene su gestor de los datos, llamado ERIC (extractor rápido de información climatológica) y no nos proporciona una herramienta integral de revisión y selección de los datos. Debido a estas carencias se propuso realizar el siguiente trabajo:

### **1.1.2 Descripción de la propuesta.**

La propuesta consiste en realizar una interfaz de gestión, consistencia y llenado de datos climatológicos.

Lo primero que se requiere para la interfaz, es una base de datos eficiente. Basada en los datos contenidos en los manejadores utilizados actualmente (CLICOM, ERIC).

La interfaz será programada con la capacidad de realizar:

- Búsqueda automática de estaciones.
- Herramientas graficas y numéricas para selección primaria de estaciones de trabajo.
- Herramientas de evaluación de la consistencia de los datos.
- Herramientas de evaluación de la correlación entre las estaciones.
- Herramientas de llenado de datos.
- Herramientas generadoras de reportes.

La interfaz tendrá como base de funcionamiento Autocad, y será programada en Autolisp, todo esto con la facilidad de introducir los polígonos de las áreas de influencia.

### **1.1.3 PLANTEAMIENTO DE LA SOLUCION:**

#### **1.1.3.1 Objetivo General.**

Realizar una herramienta de gestión, consistencia y llenado de los datos climatológicos, simplificado, amigable y fácil de usar por cualquier usuario.

#### **1.1.3.2 Como soluciones particulares se plantearon procesos individuales:**

- Conseguir información climatológica actualizada.



- Realizar una depuración de los datos climatológicos y una nueva codificación para poder usarlos en la nueva interfaz.
- Encontrar una forma simplificada de organización de la interfaz.
- Organizar herramientas de análisis gráfico para la pre-selección de las estaciones.
- Organizar una segunda interface de análisis más concreto en las estaciones de interés.
  - Pruebas de consistencia de datos.
  - Análisis correlación.
  - Llenado de datos.
- Desarrollar una rutina para crear reportes de los resultados de los análisis propuestos.
- Realizar un manual de usuario del programa.
- Depurar errores más comunes.

#### 1.1.4. METODOLOGIA UTILIZADA

El lenguaje de programación utilizado como base para el desarrollo del modulo de gestión de bases de datos climatológicos fue Autolisp y DCL para el manejo de los diálogos. El lenguaje Autolisp es un lenguaje Lisp adaptado a la interface de Autocad con col fin de ampliar su versatilidad y poder personalizar su uso. El lenguaje DCL nos permite crear cuadros de dialogo personalizados para la utilización de estos por medio del lenguaje Autolisp.

Para la obtención de los datos climatológicos actualizados se hizo la petición de la información climatológica directamente en la CONAGUA.

Los datos controlados recabados en la CONAGUA y controlados por el programa son:

- Evaporación diaria en mm.
- Evento de granizo diario.
- Precipitación diaria en mm.
- Evento de neblina.
- Estado de la nubosidad.
- Evento de tormenta eléctrica.
- Temperatura observada a las 8:00 horas.
- Temperatura máxima del día.
- Temperatura mínima del día.

Se propuso utilizar bases de datos individuales para cada estación climatológica y para cada dato en particular de la misma, todo esto con el propósito de tener un índice físico de archivos y así agilizar la búsqueda de los datos, y en caso de actualizar una en particular el proceso sea casi inmediato. El formato de las bases de datos se propuso en un formato ASCII y de manera comprimida y codificada.



Para depurar y adaptar los datos climatológicos, se elaboro un programa con la capacidad de leer los reportes generados por los manejadores de bases de datos actuales (CLICOM, ERIC), depurar los datos y reescribir las bases de datos en un formato más simplificado y ya previsto anteriormente para el manejo desde el gestor propuesto.

En base a la experiencia en las necesidades y requerimientos en los estudios hidrológicos se dio la propuesta de que la interface:

- Tomara como base el polígono de la cuenca geo-referenciado en coordenadas UTM especificando la zona horaria en la que se encuentra dibujada la zona de estudio.
- En base a una poligonal cerrada dibujada en Autocad y geo-referenciada se ubicaran las estaciones más cercanas al centroide del polígono a una distancia X.

Con la finalidad de simplificar la selección de la zona horaria UTM se desarrollo un paso adicional en la recopilación de los datos con opción de escoger el estado y el municipio de la zona de estudio.

En base a los análisis realizados comúnmente en la selección de estaciones para un estudio hidrológico se propuso la utilización de 1 sola herramienta para la obtención de reportes gráficos rápidos. Esta herramienta con múltiples opciones de datos fuentes a graficar:

Los filtros de selección de datos son de 2 tipos:

- Con criterios de selección temporal.
  - Datos anuales.
  - Datos mensuales.
  - Datos medios mensuales.
  - Datos diarios.
- Con criterios de selección numérica.
  - Máximos
  - Mínimos
  - Medios
  - Totales
  - % vacios

Todos estos con la posibilidad de combinación de un elemento de cada tipo al mismo tiempo con cualquier otro de otro tipo distinto. Y adicionalmente con opciones de tipo de grafico colores, sobre posición de datos, etc.

Los criterios de selección temporal tomados son muy simples, para una selección temporal anual y mensual, se descartan los años y meses incompletos y se muestra únicamente los años o meses con datos completos, para una selección media mensual se hizo el análisis de los datos obteniendo un único valor según sea el criterio de selección numérica. Y refiriéndose a una selección diaria se muestra el dato diario almacenado en la base de datos directamente.



La metodología utilizada para los criterios de selección numérica fue muy simple, en base al criterio de selección temporal es decir el intervalo seleccionado (mensual, anual), para una selección de máximos el dato máximo del intervalo de tiempo indicado, para una selección de mínimos el dato mínimo del intervalo de tiempo indicado, para una selección de medios el dato medio del intervalo de tiempo indicado y el dato total se calculo realizando la sumatoria de todos los datos contenidos en el intervalo de tiempo.

Para una revisión de la consistencia y homogeneidad de los datos y algunas herramientas adicionales para el llenado de datos faltantes se utilizaron metodologías:

Metodologías para la evaluación de la consistencia y homogeneidad.

- Como pruebas básicas.
  - Medias móviles. (D.F Campos Aranda 2007)
  - Pruebas de secuencias. (D.F Campos Aranda 1998)
  - Test de Helmert. (D.F Campos Aranda 1998)
- Como Pruebas específicas.
  - Pruebas de errores en cambios de media
    - Test de Student. (D.F Campos Aranda 1998)
    - Test de Cramer. (D.F Campos Aranda 1998)
  - Errores en la tendencia.
    - Test de Mann Kendall ([http://www.fmi.fi/kuvat/MAKESENS\\_MANUAL.pdf](http://www.fmi.fi/kuvat/MAKESENS_MANUAL.pdf))
    - Regresión Lineal. (Ronald E. Walpole Et.Al. 2000)
  - Errores en la Independencia.
    - Persistencia de Anderson (J.D. Salas Et. Al. 1985)

Metodologías utilizadas en el llenado de los datos.

- Coeficiente de correlación.
  - Correlación lineal. (William H. Press, 1982)
- Llenado de datos
  - METODO DEL U.S. WEATHER SERVICE. (D.F Campos Aranda 1998)
  - Método de la relación normalizada. (D.F Campos Aranda 1998)
  - Promedio aritmético de las estaciones. (D.F Campos Aranda 2007)

Para la implementación de estos métodos a una interface se desarrollaron rutinas adicionales para el manejo de nuevos objetos, todo esto con el propósito de simplificar la interface.

Para la organización del interface se propuso tener algunos filtros de selección de datos y unos filtros de selección de procesos.



En base a las rutinas que se implementaron para simplificar el programa se realizaron reportes de los resultados de los test aplicados a las series de datos y demás herramientas generadas. Todos estos en base a los criterios tomados en cada metodología en particular y concluyendo si el test encontró o no problemas en la serie de datos. Como herramienta adicional se desarrollo una opción de generar reportes de los datos extraídos directamente de las bases de datos, aplicando filtros de selección temporal y numérica, de igual forma que en la herramienta grafica realizada para la preselección de las estaciones.



# **CAPITULO II**

## **REVISION**

### **BIBLIOGRAFICA**

CIBERBITANIOS COA



## 2.1 REVISIÓN BIBLIOGRÁFICA

### 2.1.1 Antecedentes:

En la actualidad en el país se presentan muchos problemas para la gestión y llenado de datos climatológicos debido principalmente a las deficiencias de nuestros manejadores de bases de datos climatológicos, los principales programas gestores de los datos usados en México son principalmente 2:

- CLICOM.
- ERIC.

A continuación se da una breve descripción de estos gestores de bases de datos climatológicos

CLICOM:

Se trata de un gestor de datos climatológicos, introducido en 1985 para la gestión de datos climatológicos en computadoras personales (N. Plummer, Et.Al. 2007), en México estos datos son colectados diariamente por el SMN (Servicio Meteorológico Nacional), gestionado y actualizado por la CNA actualmente la CONAGUA, conteniendo en el 13 parámetros meteorológicos, que se listan a continuación:

Tabla 1, Elementos climatológicos contenidos por en el CLICOM

CODIGO	DESCRIPCION DEL ELEMENTO
1	Temp. Obs. diaria 08:00 AM (°C)
2	Temp. Max. diaria (°C)
3	Temp. min. diaria (°C)
5	Precip. diaria (mm)
18	Evap. diaria (mm)
30	Días con tormenta (0 no hay, 1 sí hay)
31	Días con granizo (0 no hay, 1 sí hay)
32	Días con niebla (0 no hay, 1 sí hay)
43	Cobertura nubosa (0 despejado, 1 medio nublado, 2 nublado)
49	Capa de nieve (cm)
56	Vel. viento prom (m/s)
59	Dirección de la racha (décimas de grados)
91	Días con helada (0 no hay, 1 sí hay)

Cabe mencionar tiene la capacidad de almacenar mucho mas información pero en nuestro país solo recolectamos esta información esto debido a la falta de estaciones climatológicas más sofisticadas.

La principal desventaja de este programa es que su interfaz esta realizada en algún compilador que trabaja en base al S.O. MSDOS por lo cual es una interfaz en texto, complicando así su



manejo para usuarios no avanzados en el sistema, además de no contener una interfaz o alguna utilidad para la búsqueda de estaciones por zonas o municipios, o algo más simplificado.

La base de datos del manejador CLICOM es manejado independientemente por cada estado, es decir que cada estado tiene su propia base de datos y tenemos esta limitante al querer consultar estaciones de otro estado.

ERIC:

Se trata de un gestor de datos climatológicos colectados diariamente por el SMN (Servicio Meteorológico Nacional), gestionado y actualizado por el IMTA, conteniendo en él, 9 parámetros meteorológicos, que se listan a continuación:

Tabla 2, Datos Climatológicos contenidos en ERIC

CODIGO	DESCRIPCION DEL ELEMENTO
1	Temp. obs. diaria 08:00 AM (°C)
2	Temp. Max. diaria (°C)
3	Temp. min. diaria (°C)
4	Precip. diaria (mm)
5	Evap. diaria (mm)
6	Días con tormenta (0 no hay, 1 sí hay)
7	Días con granizo (0 no hay, 1 sí hay)
8	Días con niebla (0 no hay, 1 sí hay)
9	Cobertura nubosa (0 despejado, 1 medio nublado, 2 nublado)

En su última versión contiene algunos comandos de visualización y mejoras muy buenas en cuanto a la consulta para el usuario en general se refiere, contiene un listado de más de 6000 estaciones climatológicas de todo el país y tiene un manejo ágil de la información. Cabe mencionar que la última actualización realizada al programa fue realizada por el CSVA (Consejo del Sistema Veracruzano del Agua).

Su principal desventaja es la gran cantidad de vacíos e inconsistencias de los datos y poca cantidad en años registrada de cada estación esto debido a que la información climatológica no es muy accesible o no hay muchas fuentes de la misma, y la falta de una herramienta de depuración de datos climatológicos.

En resumen la base de datos CLICOM es muy completa pero su interfaz no es muy amigable, por otro lado ERIC resuelve algunos problemas con su interfaz más amigable pero contiene una gran cantidad de vacíos en sus bases y no contiene en su totalidad todas las estaciones climatológicas registradas, además de carecer de herramientas de revisión de consistencia de datos.



## 2.1.2 Marco teórico para la revisión de la calidad de los datos.

### Metodologías para la revisión de la consistencia en los datos.

En términos generales se dice que una serie climatológica es homogénea o consistente si es una muestra de una única población. Si la serie parece ser in-homogénea se le deben hacer algunas correcciones para lograr su homogeneidad y con ello lograr que las evaluaciones estadísticas muestrales sean estimaciones validas de los correspondientes parámetros poblacionales.

#### Pruebas básicas

Al decir pruebas básicas el texto refiere a pruebas muy generales para detectar problemas de inconsistencia estadística de los datos climatológicos, los métodos propuestos para el análisis de la homogeneidad solo se ha probado su efectividad en datos de precipitación, la mayoría de estos utilizados para procesar datos climatológicos anuales, algunos otros mensuales. Y los métodos utilizados fueron análisis con promedios móviles, pruebas de secuencias y test de Helmert. (D.F Campos Aranda 2007)

#### Análisis con promedios móviles. (D.F Campos Aranda 2007)

El análisis de los datos de precipitación anual acumulada es lógico que este valor refleje tendencias o saltos, al ser representados cronológicamente, esto significa el graficar los datos totales de precipitación anual, graficados en el tiempo. Con esto se tiene una serie cronológica original.

Con el propósito de suavizar la dispersión de la serie y poder analizar con mayor facilidad las tendencias en la media o la dispersión se calculan los promedios móviles. El cálculo de los promedios móviles se realiza con la siguiente formula.

$$\frac{X_1 + X_2 + \dots + X_n}{N}, \frac{X_2 + X_3 + \dots + X_n}{N}, \frac{X_3 + X_4 + \dots + X_n}{N}$$

En donde N es el orden de los promedios móviles, y la serie cronológica de precipitación anual es:

$$X_1, X_2, X_3, X_4, \dots \dots \dots, X_n$$

Las series cronológicas de precipitación anual generalmente se analizan con N=5, si se analiza con un N muy bajo se pierde efectividad pues no se reduce la variabilidad, en cambio si N es muy grande se corre el riesgo que se oculten algunos componentes como la cíclica.

Por medio de una línea recta se unen los datos de los periodos secos y de los periodos húmedos y los que varían mucho sobre la media, son probablemente errores en la toma o captura de los datos; analizando así de forma visual la homogeneidad de la serie.



### Prueba de las secuencias. (D.F Campos Aranda 2007)

También conocida como prueba general para detectar inconsistencias en las series climatológicas. El primer paso para aplicar el método a una serie de datos es calcular la mediana de la serie que se analiza.

El cálculo de la mediana para una serie de datos consiste en:

- Se ordenan los datos según su magnitud.
- Si el numero de datos es impar la mediana es el valor central, en caso contrario la mediana corresponde a la media de los 2 datos centrales.

Con base a la mediana calculada se marca en la serie con A si es mayor y con B si el dato es menor. Las secuencias de A y B se contabilizan y se concluye si u esta dentro del intervalo que marca la tabla siguiente de acuerdo al número de datos contenidos en la serie, y se concluye si se tiene un valor muy alto, se puede deber a un exceso de oscilación en la serie, por el contrario si el numero de u es muy bajo se concluye que el problema de homogeneidad se puede deber a una tendencia existente o un cambio en la media.

Tabla 3, Rangos de tolerancia para el numero de secuencias permisible

Numero de datos (n)	Rango de U	Numero de datos (n)	Rango de U
12	5-8	32	13-20
14	5-10	34	14-21
16	6-11	36	15-22
18	7-12	38	16-23
20	8-13	40	16-25
22	9-14	50	22-30
24	9-16	60	26-36
26	10-17	70	31-41
28	11-18	80	35-47
30	12-19	100	45-57

### Test de Helmert. (D.F Campos Aranda 2007)

También conocida como prueba general para detectar inconsistencias en las series climatológicas. Es una prueba muy simple y consiste en analizar el signo de las desviaciones de cada dato con respecto a la media de toda la serie. Si una desviación de cierto signo es seguida por uno igual tenemos una secuencia (S), por el caso contrario si tenemos una desviación de un cierto signo y esta es seguida por un signo contrario tenemos una secuencia (C)

La serie puede ser homogénea si la diferencia de secuencias S y C es igual a cero o está dentro de los límites probablemente dados por un error el cual depende de la longitud del registro. Y lo podemos calcular de la siguiente forma:



$$S - C = 0 \pm \sqrt{n - 1}$$

Por tanto al aplicar el test podemos concluir si tenemos:

- $S > C$  el problema de homogeneidad se puede deber a una tendencia existente o un cambio en la media.
- $C > S$  puede haber un exceso de oscilación en la serie.

### Pruebas específicas

Cuando los análisis gráficos y las pruebas básicas indicaron algún error en la serie de datos es necesario aplicar pruebas específicas para verificar si efectivamente existen errores estadísticos en la serie climatológica. Y se utilizaron métodos de evaluación de cambios en la media, evaluación de errores en la tendencia, y la evaluación de la independencia. (D.F Campos Aranda 2007)

#### Evaluación de errores por cambios en la media.

Al evaluar los cambios en la media se refiere a verificar en la serie cronológica en dos tramos distintos, si existe realmente un cambio en la media. Y los métodos propuestos fueron test de la t de Student y test de Cramer.

#### Test de la t de STUDENT (D.F Campos Aranda 1998)

Cuando la causa de la perdida de homogeneidad de la seria sea probablemente causada por un cambio abrupto en la media este test es muy útil para evaluar si realmente hay un error en el cambio de la media.

Se trata de un test robusto excepto cuando la longitud de los dos periodos es desigual por este motivo se recomienda aplicar este test que la media de los datos sea muy similar.

La prueba de t Student está definida por la siguiente ecuación.

$$td = \frac{\bar{X}1 - \bar{X}2}{\left[ \left( \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2} \right) * \left( \frac{1}{n_1} + \frac{1}{n_2} \right) \right]^{\frac{1}{2}}}$$

Siendo  $S_1^2$  y  $S_2^2$  las varianzas de xi por lo tanto  $n_2 S_2^2$  se puede calcular con la ecuación:

$$n_2 S_2^2 = \sum_1^{n_2} x_i^2 - \frac{1}{n_2} \left( \sum_1^{n_2} x_i \right)^2$$

De igual forma  $n_1 S_1^2$ .

El valor absoluto de la  $td$  se compara con el valor de la t de la distribución t Student de dos colas y con  $v = n_1 + n_2 - 2$  grados de libertad y con un 5% de nivel de significancia. Los valores de t se encuentran en la tabla correspondiente.



Prueba de una cola \*

Prueba de dos colas \*\*

Tabla 4, valores de la tolerancia

Grados de Libertad	Nivel Significancia		Grados de Libertad	Nivel Significancia	
	5% *	5% **		5% *	5% **
1	6.314	12.706	18	1.734	2.101
2	2.92	4.303	19	1.729	2.093
3	2.353	3.182	20	1.725	2.086
4	2.132	2.776	21	1.721	2.08
5	2.015	2.571	22	1.717	2.074
6	1.943	2.447	23	1.714	2.069
7	1.895	2.365	24	1.711	2.064
8	1.86	2.306	25	1.708	2.06
9	1.833	2.262	26	1.706	2.056
10	1.812	2.228	27	1.703	2.052
11	1.796	2.201	28	1.701	2.048
12	1.782	2.179	29	1.699	2.045
13	1.771	2.16	30	1.697	2.042
14	1.761	2.145	40	1.684	2.021
15	1.753	2.131	60	1.671	2
16	1.746	2.12	120	1.658	1.98
17	1.74	2.11	∞	1.645	1.96

Si el valor absoluto de  $t\bar{d}$  es mayor que  $t$  de la distribución se concluye que hay evidencia de falta de homogeneidad en las medias.

*Test de CRAMER (D.F Campos Aranda 1998)*

Algunas veces puede ser más conveniente comparar la media de toda la serie y la media de una cierta parte del registro, el propósito de esta prueba es ese, además de ser complementaria para el test de la  $t$  de Student siendo esta no recomendable cuando  $n_1$  y  $n_2$  no son parecidos.

En esta prueba  $\bar{X}$  y  $S$  son respectivamente la media y la desviación típica del registro total. Por otra parte  $\bar{X}_K$  es la media del sub periodo de  $n'$  valores y se puede calcular de la siguiente forma:

$$\bar{X}_K = \left( \sum_{i=k+1}^{i=k+n} x_i \right) / n'$$



$$tk = \frac{(\bar{X}_K - \bar{X})}{S}$$

$$Tk = \left\{ \frac{n'(n-2)}{n - n'[1+tk^2]} \right\}^{\frac{1}{2}} tk$$

La estadística  $tk$  tiene la distribución t Student de dos colas con  $v = n - 2$  grados de libertad y es utilizado de igual forma que la estadística  $td$  del método de la t de Studet.

#### Evaluación de errores en la tendencia.

Al evaluar los cambios en la tendencia se refiere a verificar si la serie de datos tiene alguna tendencia creciente o decreciente en los datos climatológicos y si esta es significativa o no en términos estadísticos. Y el método utilizado fue el test de Mann Kendall.

*Test de Mann Kendall ([http://www.fmi.fi/kuvat/MAKESENS\\_MANUAL.pdf](http://www.fmi.fi/kuvat/MAKESENS_MANUAL.pdf))*

El test de Mann Kendall para la detección de errores en la tendencia, es aplicable si asumiendo que la serie de tiempo es del tipo siguiente:

$$x_i = f(t) + \epsilon_i$$

Donde  $f(t)$  es monótona y continua, aumenta o disminuye en función del tiempo, y los residuos  $\epsilon_i$  se puede suponer que son de la misma distribución pero su media es igual a cero. Con esto suponemos que la varianza de distribución es igual a cero.

El número de datos mínimo que se debe utilizar son 10 datos, con menos datos el test se vuelve ineficiente, y se pueden tener datos faltantes.

El test estadístico de Mann Kendall se realiza con la siguiente fórmula:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(x_j - x_k)$$

Donde  $x_j$  y  $x_k$  son los datos anuales en los años  $j$  y  $k$ ,  $j > k$  respectivamente y:

$$\text{sgn}(x_j - x_k) = \begin{cases} 1 & \text{si } x_j - x_k > 0 \\ 0 & \text{si } x_j - x_k = 0 \\ -1 & \text{si } x_j - x_k < 0 \end{cases}$$

Si  $n$  es 9 o menos, el valor absoluto de  $S$  se compara directamente con la distribución teórica de  $S$  obtenidos por Mann y Kendall (Gilbert, 1987).

Los valores mínimos de  $n$  con la que estos cuatro niveles de significación se pueden alcanzar, son derivados de la tabla de probabilidad siguiente de  $S$ :



Tabla 5, Valores mínimos de n

Nivel de significancia	Requiere por lo menos n
0.1	≥ 4
0.05	≥ 5
0.01	≥ 6
0.001	≥ 7

El nivel de significancia 0.1 nos dice que existe una probabilidad de 0.1% que los valores de  $x_i$  sean de una distribución aleatoria, con una probabilidad de cometer un error al rechazar  $H_0$  o no encontrar una tendencia.

**10 datos o más:** si n es al menos 10 se utiliza el test de aproximación normal, sin embargo si hay varios atados es decir varios datos iguales el test puede reducir su efectividad al acercarse a 10 datos debido a la reducción de n por datos repetidos.

Primero calculamos la varianza S con la siguiente ecuación.

$$VAR(S) = \frac{1}{18} \left[ n(n-1)(2n+5) - \sum_{p=1}^q t_p(t_p-1)(2t_p+5) \right]$$

En donde q es el número de grupos afines (número de grupos con datos repetidos), y  $t_p$  es el número total de repeticiones contenidas de cada número. En base a S y VAR(S) obtenemos el estadístico Z de la siguiente forma:

$$Z = \begin{cases} \frac{S-1}{\sqrt{VAR(S)}} & \text{si } S > 0 \\ 0 & \text{si } S = 0 \\ \frac{S+1}{\sqrt{VAR(S)}} & \text{si } S < 0 \end{cases}$$

Para encontrar la presencia de una tendencia estadísticamente significativa utilizamos el valor de Z, un valor positivo indica una tendencia hacia arriba y un valor negativo indica una tendencia hacia abajo. El valor estadístico Z se encuentra en una distribución normal. Para probar la presencia de una tendencia monótona hacia arriba o hacia abajo, en el nivel de significación  $\alpha$   $H_0$  se rechaza si el valor absoluto de Z es mayor que  $Z_{1-\alpha/2}$  donde  $Z_{1-\alpha/2}$  se obtiene de la tabla de distribución normal acumulada.

### MÉTODO DE SEN

Este método se utiliza para estimar la verdadera pendiente de una tendencia. Este método se puede utilizar cuando se asume que la tendencia es lineal. Esto nos dice que f (t) es igual a:

$$f(t) = Qt + B$$



Donde Q es la pendiente y B es una constante.

Para estimar la pendiente Q primero calculamos las pendientes de todos los datos pares.

$$Q_i = \frac{x_j - x_k}{j - k}$$

Si tenemos n valores de  $x_j$  en la serie de tiempo nosotros tendremos  $N = n(n-1)/2$  numero de pendientes  $Q_i$  calculadas. Para estimar la pendiente de Sen, calculamos la mediana de la serie

$$Q = Q_{\left[\frac{n+1}{2}\right]} \text{ Si } N \text{ es numero impar}$$

$$Q = \frac{1}{2} \left( Q_{\left[\frac{n}{2}\right]} + Q_{\left[\frac{n+2}{2}\right]} \right) \text{ Si } N \text{ es numero par}$$

Un 100 (1- $\alpha$ ) % intervalo de confianza de dos lados sobre el estimador de la pendiente se obtiene por la técnica no para-métrica basada en la distribución normal. El método es válido para n tan pequeños como 10 a menos que existan muchos lazos.

Al principio calculamos:

$$C_\alpha = Z_{\frac{1-\alpha}{2}} \sqrt{\text{VAR}(S)}$$

Donde VAR (S) ya se ha definió anteriormente, y  $Z_{1-\alpha} / 2$  se obtiene de la tabla de distribución normal acumulada.

Después calculamos los niveles de confianza superior e inferior  $Q_{min}$  y  $Q_{max}$   $M1 = (N - C_\alpha)/2$  y  $M2 = (N + C_\alpha)/2$ .

Para obtener B se calcula la mediana de la serie  $x_i - Q_i$ .

#### *Regression lineal Simple (Ronald E. Walpole Et.Al. 2000)*

El método de regresión lineal simple es un método muy sencillo consiste en el ajuste de una ecuación de una línea recta a n datos. Para su aplicación requiere que existan dos variables únicamente, una variable de regresión independiente x y una sola aleatoria dependiente y los datos pueden presentarse por los pares de observaciones  $\{(x_i, y_i); i = 1, 2, \dots, n\}$

#### **MÉTODO MÍNIMOS CUADRADOS**

Se encontrarán a y b, y las estimaciones de  $\alpha$  y  $\beta$ , de tal forma que la suma de los cuadrados de los residuos sea mínima. La suma de los cuadrados de los residuos también recibe el nombre de suma de cuadrados de los errores alrededor de la línea de regresión SSE. La formulación general del método es:



$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

$$\frac{\partial(SSE)}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i)$$

$$\frac{\partial(SSE)}{\partial b} = -2 \sum_{i=1}^n (y_i - a - bx_i) x_i$$

Las cuales se resuelven simultáneamente para dar las fórmulas de cálculo de a y b.

$$b = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$a = \frac{(\sum_{i=1}^n y_i) - b \sum_{i=1}^n x_i}{n}$$

#### Evaluación de errores en la independencia.

Al evaluar los errores en la independencia se refiere a la evaluación de la independencia de una misma serie temporal, es decir evaluar que tan independiente es un valor del siguiente y así sucesivamente. El método utilizado para la evaluación de la independencia de los datos fue la persistencia de Anderson.

#### *Persistencia de Anderson (J.D. Salas Et. Al. 1985)*

La característica de la dependencia de las series de tiempo anual, puede ser analizada utilizando 2 cálculos estadísticos clásicos.

- El correlograma con una representación del dominio en el tiempo.
- El espectro con una representación de la frecuencia en el tiempo.

Para una serie independiente la población del correlograma es igual a cero para  $k \neq 0$ . Sin embargo series temporales independientes, presentan variabilidad de la muestra, conteniendo fluctuaciones de  $r_k$  alrededor de cero, pero estos no son necesariamente cero. En este caso lo que hacemos para determinar los límites de probabilidad de series independientes, utilizamos los límites de Anderson (1941). Los cuales se calculan:

$$r_k(95\%) = \frac{-1 \pm 1.96\sqrt{N - k - 1}}{N - k}$$

$$r_k(99\%) = \frac{-1 \pm 2.326\sqrt{N - k - 1}}{N - k}$$

Para el 95% y 99% de porcentaje de nivel de probabilidad respectivamente, siendo N el tamaño de la muestra.



El correlograma consiste en obtener múltiples correlaciones desfasando la misma serie de precipitaciones, comenzando del orden 1 hasta 25, si los datos calculados en los 25 casos están dentro de los límites de Anderson se dice que la serie de datos es independiente de lo contrario se dice que tiene problemas de independencia.

### Metodologías utilizadas en el llenado de datos y herramientas auxiliares.

#### Correlación (Ronald E. Walpole Et.Al. 2000)

La correlación es análisis sencillo de la relación entre las dos variables o dos series de datos para este caso aplicado a las precipitaciones se refieren a la relación que existe entre las alturas del mismo en diferentes sitios o estaciones climatológicas con respecto de otras. El análisis de correlación intenta medir la fuerza de las relaciones entre dos variables por medio de un simple número llamado coeficiente de correlación.

En teoría se supone con frecuencia que la distribución condicional  $f(y/x)$  de Y para valores fijos de X, es normal con media  $\mu_{y/x} = \alpha + \beta x$  y varianza  $\sigma_{y/x}^2$ , y que de la misma manera, X tiene distribución normal con media  $\mu_x$  y varianza  $\sigma_x^2$ . La densidad conjunta de X e Y es entonces:

$$f(x, y) = n(y|x; \alpha + \beta x, \sigma) n(x; \mu_x, \sigma_x)$$

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma} \exp\left\{-\frac{1}{2}\left[\left(\frac{y - (\alpha + \beta x)}{\sigma}\right)^2 + \left(\frac{x - \mu_x}{\sigma_x}\right)^2\right]\right\}$$

Para  $-\infty < x < \infty$  y  $-\infty < y < \infty$

Escríbase la variable aleatoria Y en la forma:

$$Y = \alpha + \beta X + E$$

Donde X es ahora la variable aleatoria independiente del error aleatorio E dado que la media del error aleatorio E es =, se sigue que,

$$\mu_y = \alpha + \beta\mu_x$$

$$\sigma_y^2 = \sigma^2 + \beta^2\sigma_x^2$$

Al sustituir  $\alpha$  y  $\sigma^2$  en la expresión anterior para  $f(x, y)$ , se obtiene la distribución normal bi-variada.

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-p^2}}$$

$$\exp\left\{-\frac{1}{2(1-p^2)}\left[\left(\frac{x - \mu_x}{\sigma_x}\right)^2 - 2p\left(\frac{x - \mu_x}{\sigma_x}\right)\left(\frac{y - \mu_y}{\sigma_y}\right) + \left(\frac{y - \mu_y}{\sigma_y}\right)^2\right]\right\}$$



Para  $-\infty < x < \infty$  y  $-\infty < y < \infty$  donde:

$$p^2 = 1 - \frac{\sigma^2}{\sigma_y^2} = \beta^2 \frac{\sigma_x^2}{\sigma_y^2}$$

La constante  $p$  es también llamada coeficiente de correlación poblacional.

La medida  $p$  de asociación lineal entre dos variables se estima con el coeficiente de correlación muestral  $r$ , y lo podemos calcular de la siguiente manera.

$$r = b \sqrt{\frac{S_{xx}}{S_{yy}}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

Entonces  $r^2$  o coeficiente de determinación muestral:

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{SSR}{S_{yy}}$$

### Llenado de datos

#### METODO DEL U.S. WEATHER SERVICE. (D.F Campos Aranda 1998)

Este procedimiento teórica y empíricamente por lo tanto puede ser muy confiable, y considera que el dato faltante de la estación A, puede ser estimado en base a las estaciones circundantes. Este método puede ser utilizado para estimar datos diarios, mensuales y anuales.

El método consiste en ponderar los valores observados en una cantidad  $W$ , igual al recíproco del cuadrado de la distancia  $D$  entre cada estación vecina y la estación A, La lluvia buscada  $P_x$  será igual a:

$$P_x = \frac{\sum (P_i - W_i)}{\sum W_i}$$

$P_i$  = Precipitación observada para la fecha faltante en las estaciones auxiliares circundantes (pueden ser como mínimo 2), en milímetros.

$W_i = \frac{1}{D_i^2}$ , siendo  $D_i$  la distancia entre cada estación circundante y la estación incompleta, en Km.

Se recomienda utilizar cuatro estaciones auxiliares (las más cercanas) y de manera que cada una quede localizada en uno de los cuadrantes que definen unos ejes coordenados que pasan por la estación incompleta, generalmente los ejes norte-sur y oriente-poniente.



#### Método de la relación normalizada. (D.F Campos Aranda 1998)

Este método se utiliza para realizar el llenado mensual o anual de una cierta estación con datos faltantes. Se estima a partir de los valores observados en tres estaciones cercanas, situadas uniformemente alrededor de la estación incompleta. Si la precipitación de las estaciones auxiliares esta dentro de un 10% de la estación registrada se utilizara el promedio aritmético del registro de las 3 estaciones.

Si el dato difiere en más del 10% el registro faltante será calculado de la siguiente forma:

$$P_x = \frac{1}{3} \left[ \frac{N_x}{N_A} P_A + \frac{N_x}{N_B} P_B + \frac{N_x}{N_C} P_C \right]$$

En donde:

$N_x$  Precipitación media anual en la estación incompleta en mm.

$N_A, N_B, N_C$  Precipitación media anual en las estaciones auxiliares A, B, C en mm.

$P_A, P_B, P_C$  Precipitación anual o mensual observada en la estaciones A, B, C, en mm para la misma fecha que la faltante, en milímetros.

El método no se restringe a tres estaciones si no que puede ser empleado un número mayor de estaciones, siempre y cuando tengan la misma distancia a la estación con los datos faltantes.



### 2.1.3 Marco teórico para la programación del modulo.

Para la programación del modulo el primer paso será la depuración de las bases de datos, para esto se realizo un programa que depuro y reestructuro los datos para poder usarlos desde Autocad.

En total se realizara un modulo de gestión de los datos programado en Autolisp, y un programa depurador de los datos fuente, creado en Visual Basic.

#### Autolisp

Autolisp es un lenguaje de programación diseñado para extender la funcionalidad de Autocad y tener la posibilidad de personalizarlo. Autolisp de Autocad es introducido como una interfaz de programación de aplicaciones en su versión 2.1 a mediados de 1980. Autolisp fue el primer lenguaje implementado a Autocad. Visual Lisp es un programa diseñado para acelerar el desarrollo y depuración de las aplicaciones Autolisp. Evitando así un trabajo tedioso de depuración de código. (Autocad, 2008)

Visual Lisp contiene su propio conjunto de ventanas, pero no es independiente del Autocad, los dos programas trabajan en conjunto.

Autolisp cuenta al igual que muchos lenguajes de programación con funciones y operadores para los distintos requerimientos, algunas funciones para el manejo interno del programa y algunas otras para la comunicación con otros programas, para el manejo de diálogos DCL, o para llamado de los comandos internos predefinidos por Autodesk en Autocad, y para ejecución de comandos externos.

La información detallada de los comandos y la sintaxis del lenguaje se encuentran bien documentadas en los archivos de ayuda de Autocad en la mayoría de sus versiones.

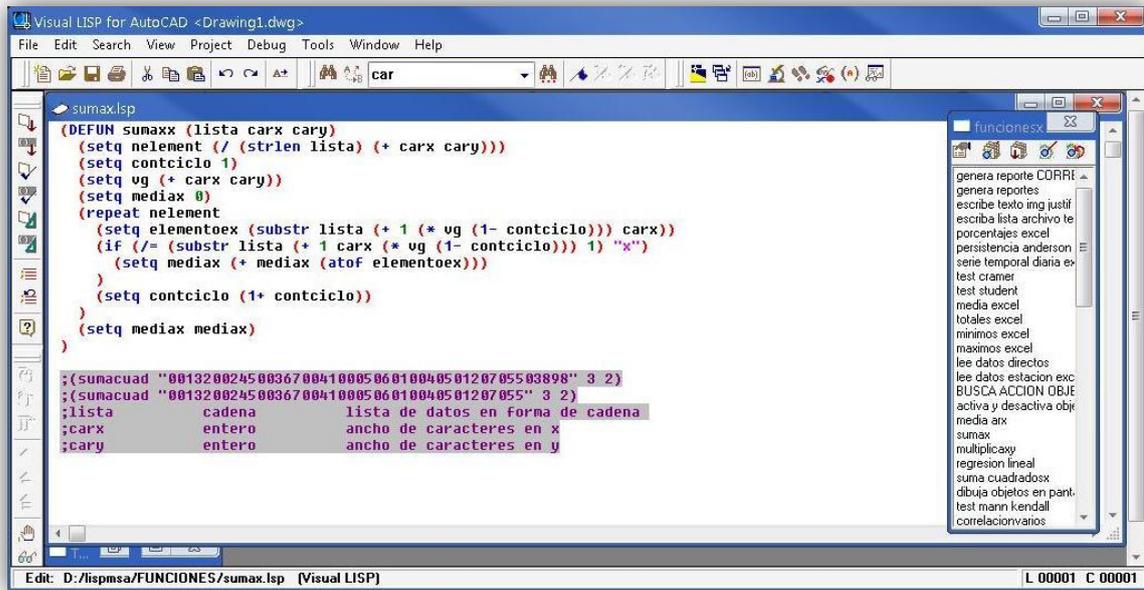


Figura 1: Interface de programación VLisp

## DCL (lenguaje de control de diálogos)

Autolisp como complemento visual tiene este lenguaje de programación que nos permite crear cuadros de diálogo con una sintaxis no tan complicada. Los cuadros de diálogo son escritos en archivos de texto en formato ASCII, en los cuales podemos definir las dimensiones y propiedades de los mismos, como la altura y el ancho de los objetos, alineación, leyendas, etc., todas estas opciones y sintaxis se encuentran bien documentadas en los archivos de ayuda de AutoCAD.

El control de las acciones de los cuadros de diálogo, en la interacción con el usuario se efectúa por medio de sentencias específicas realizadas en el lenguaje Autolisp, la única función del lenguaje DCL como tal es el de especificar las dimensiones y propiedades iniciales de los diálogos así como los identificadores de los objetos.

La edición de los archivos de texto también puede ser realizada en la interface de VLisp (figura 2), y además la interface cuenta con una herramienta de pre visualización de los diálogos para facilitar el diseño de los mismos (figura 3).

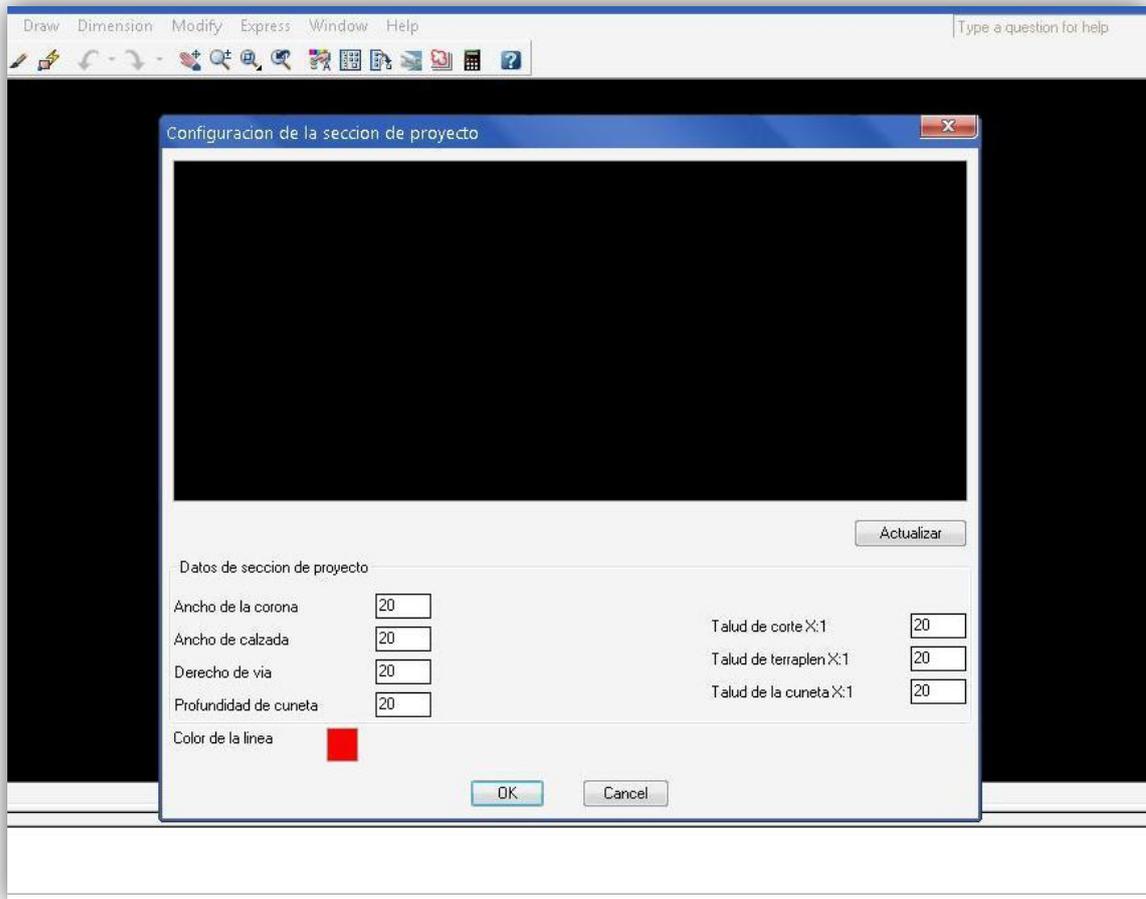


Figura 2, Pre visualización del dialogo en diseño

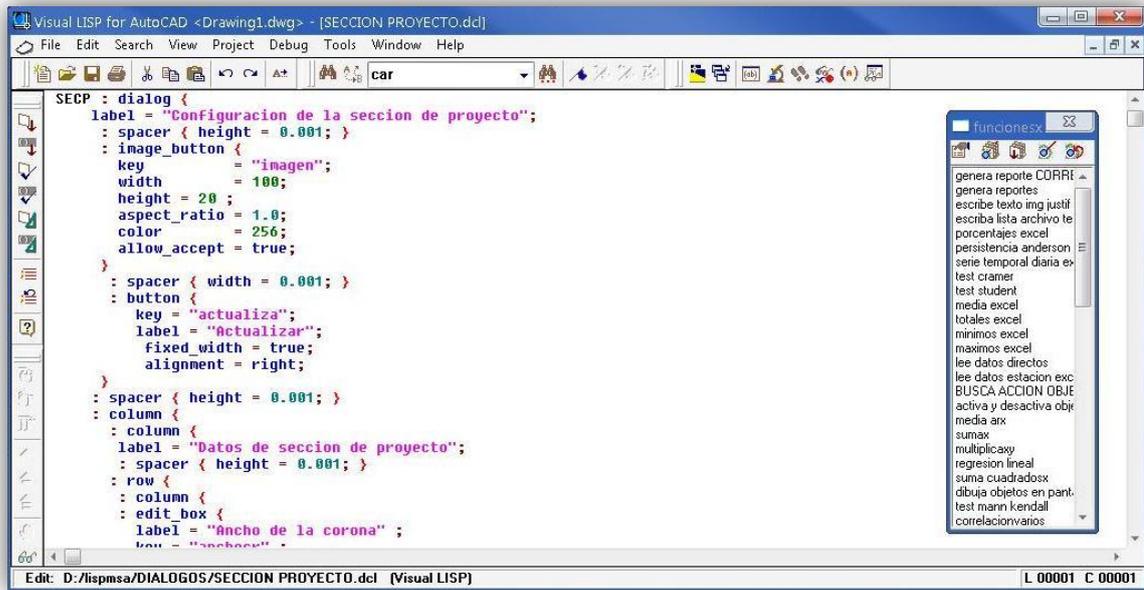


Figura 3, Edición del dialogo

### Visual Basic

El lenguaje de programación BASIC (Beginner's All-purpose Symbolic Instruction Code) es quizá el lenguaje de programación más conocido debido a su sencillez para programar, creado en un principio como ayuda para programadores principiantes, poco a poco fue evolucionando, comenzó como BASIC y luego se desarrollaron versiones como QUICKBASIC y GW-BASIC, posteriormente salió la versión QBASIC, una versión muy potente y fácil de usar, posteriormente siguieron algunas versiones más de alto nivel, y luego dio comienzo la nueva generación de lenguajes de programación visuales los cuales ya eran enfocados específicamente para la programación bajo los sistemas operativos de Microsoft.

Visual Basic ya no es más "un lenguaje para principiantes" sino que es una perfecta alternativa para los programadores de cualquier nivel que deseen desarrollar aplicaciones compatibles con Windows.

El entorno de programación de Visual Basic, es un entorno totalmente visual que nos permite arrastrar los objetos requeridos al formulario o ventanas de nuestro proyecto, y programar las acciones a realizar de cada objeto simplemente dando doble clic sobre alguno de ellos en modo diseño, además este lenguaje nos permite compilar nuestro código fuente, así como contiene un depurador de sintaxis activo en todo el proceso de programación.

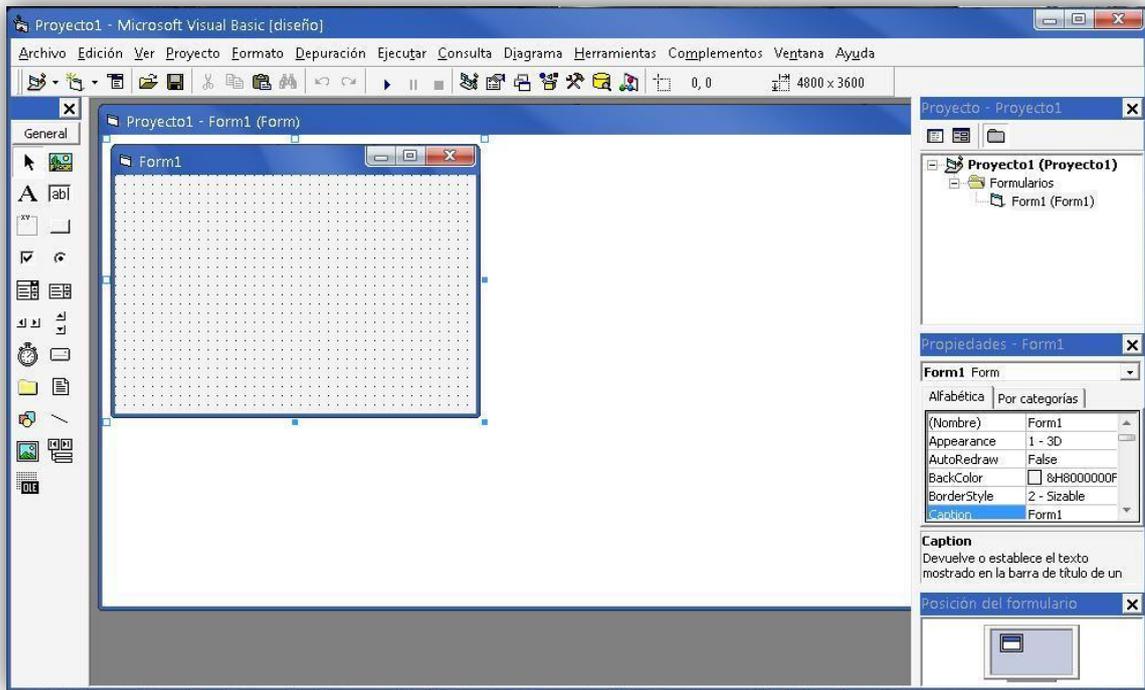


Figura 4: Entorno de programación de Visual Basic



# **CAPITULO III**

## **MODULO DE GESTION**



### 3.1 MODULO DE GESTION.

#### 3.1.1 El modulo de gestión tiene como requerimientos mínimos:

- Autocad 2008 o versiones posteriores.
- Excel de Microsoft (para generar reportes directos de los datos de las bases de datos.)
- Procesador Intel ATOM 1.3 GHZ o superior
- Memoria en gráficos 10 m.
- Memoria RAM 2Gb.
- 1.6gb de espacio libre en disco duro
- Windows XP o superior.

#### 3.1.2 Como ejecutar el programa

Para ejecutar el modulo de gestión de datos climatológicos se requiere haber cumplido los requerimientos mínimos, y tener la poligonal de la cuenca geo referenciada y trazada con una poligonal en Autocad, esta poligonal teniendo como identificador interno de la entidad en asociación (0 . "LWPOLYLINE").

Para comenzar la ejecución del programa hay que teclear en la línea de comandos de Autocad el comando "estaciones" seguido de la tecla ENTER o retorno de carro, y comienza el proceso.

#### 3.1.3 Descripción del proceso

Al comenzar el proceso el primer dato requerido por el programa es la poligonal geo referenciada, la poligonal tiene que estar previamente trazada antes de ejecutar el programa, (en caso de tener una poligonal con muchos vértices, se recomienda trazar una poligonal mas simplificada para optimizar en redibujado del programa).

#### Selección adecuada de la zona horaria UTM

Una vez seleccionada la poligonal que abarca la zona de estudio, se verá un cuadro de dialogo como el mostrado en la figura 1, el cual tiene la función de simplificar la selección de la zona horaria en la que se encuentra dibujada nuestra poligonal, la forma de simplificar o ayudar en la selección de la zona horaria es simple, seleccionamos el estado y municipio de la zona de estudio y el cuadro de dialogo nos propone la zona horaria en que normalmente seria dibujada la poligonal, y en algunos casos podemos encontrar que el municipio abarca según su posición geográfica varias zonas horarias, en este caso el dialogo nos mostrara todas las zonas posibles en el municipio, en caso de saber en qué zona horaria esta dibujada la poligonal se puede omitir la selección de estado

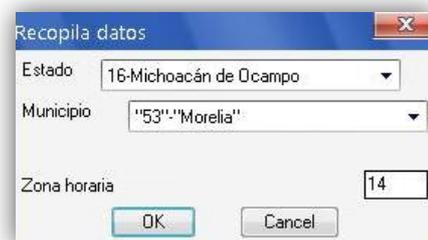


Figura 5. Cuadro de dialogo inicial de selección zona.



y municipio y solamente teclear la zona de horaria con un dato numérico de 2 dígitos ("14", en caso de Morelia, Mich.).

Una vez seleccionada la poligonal y la zona horaria comienza la búsqueda de las estaciones más cercanas con una distancia proporcional al tamaño de la zona en estudio, si la distancia de búsqueda es muy pequeña y no se localizan estaciones, el programa realizara algunas iteraciones modificando la distancia de búsqueda hasta encontrar una o más estaciones, en este caso el programa nos mostrara un dialogo avisando que se está ajustando la distancia de búsqueda, figura 2

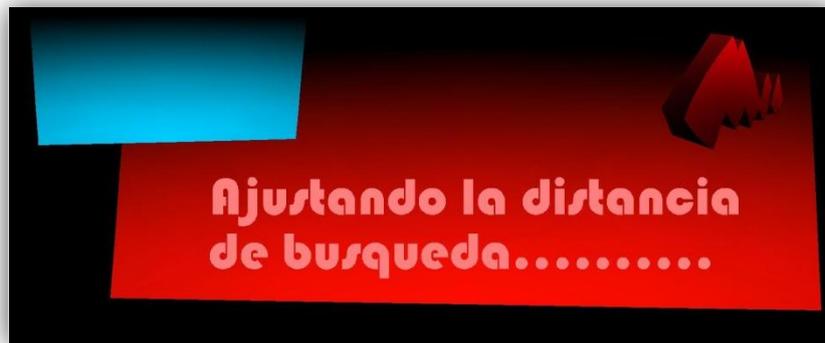


Figura 6, Cuadro de dialogo de ajuste de distancia.

#### **Primer interface (preselección y análisis superficial de estaciones).**

Al terminar el proceso de ajuste de la distancia, el programa comienza el trazado del polígono base en Autocad y la sobre posición de las estaciones climatológicas encontradas, la poligonal es trazada en color rojo y las estaciones se trazan con iconos de 2 tipos para diferenciar entre las estaciones registradas en la base de datos CLICOM o ERIC y un icono redondo para diferenciar la estación seleccionada. Como se muestra en la figura 3.3. Cabe mencionar que este dibujo es trazado en coordenadas geográficas es decir en grados decimales y puede variar a uno dibujado en coordenadas UTM.



Figura 7 Cuadro de dialogo inicial sobre posición de las estaciones.

Analizando la primera interface (figura 3) podemos observar que consta de dos zonas principales: una zona de dibujo y una zona de controles.

- En la zona de controles encontramos varias herramientas
  - Control del zoom en el dibujo
    - Nos permite alejar y acercar el dibujo.
  - Un control de la estación seleccionada
    - Nos permite pasar de 1 a 1 las estaciones en orden creciente y de una a otra base de datos.
  - un filtro de visualización de estaciones basado en la procedencia de los datos
    - nos permite seleccionar si queremos visualizar o no las estaciones de la base de datos, CLICOM o ERIC.
  - La distancia de búsqueda respecto al centroide del polígono
    - Un cuadro de texto que nos permite modificar la distancia de búsqueda de estaciones, en caso de querer redibujar la nueva distancia introducida se tiene que presionar de nuevo el botón “buscar estaciones”.

- botón para pasar al siguiente dialogo del proceso “terminar selección”.
- En la zona de dibujo encontramos también una serie de controles y zonas de despliegue de información. En general podemos clasificarlas en 5 zonas. Como se muestra en la figura 4

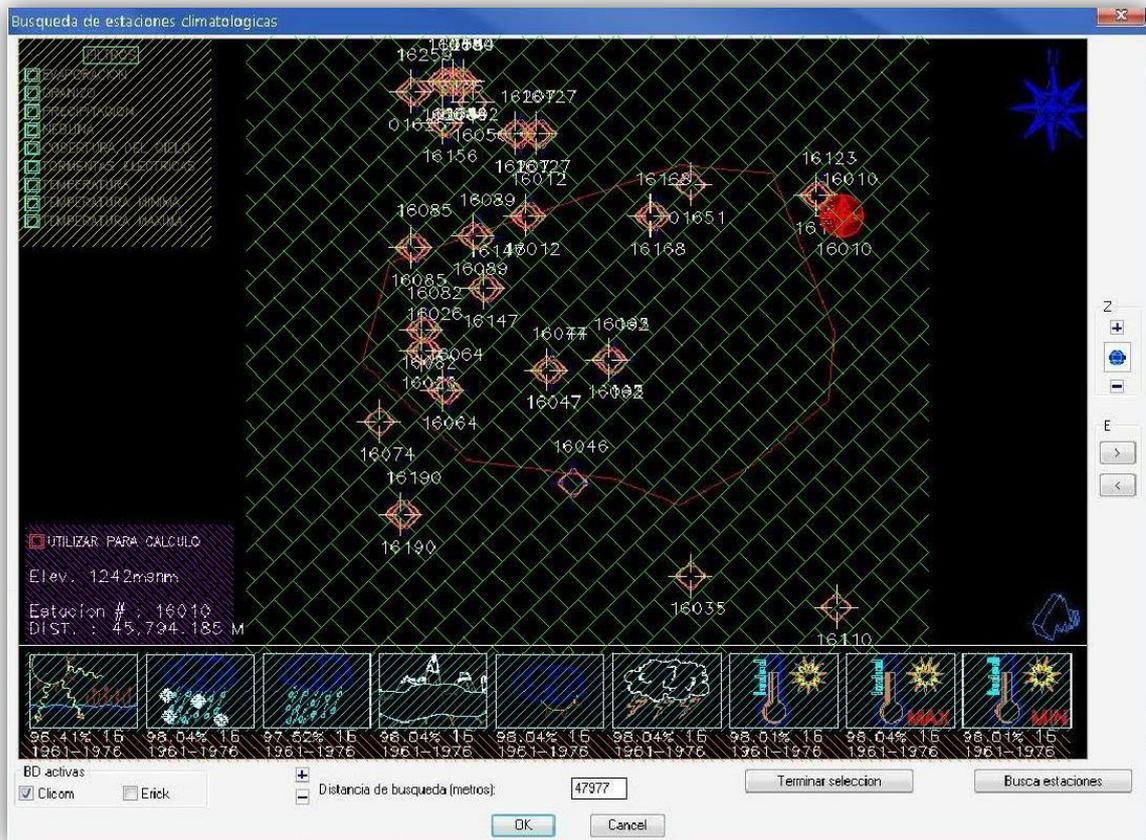


Figura 8, Zonificación de área de dibujo

La zona 1 es una zona de controles de filtro de búsqueda, es decir que podemos ver las estaciones que contengan información disponible del dato climatológico indicado.

La zona 2 es el área reservada para el dibujo y selección de las estaciones. Es decir en ella se dibujara la sobre posición de la poligonal en estudio y las estaciones más cercanas encontradas.

La zona 3 es la zona de despliegue y una pequeña zona de selección de estación, figura 7.



Figura 9, Simbología figura 4



La opción de selección de la estación (etiquetada como “UTILIZAR PARA CALCULO”), nos permite tomarla como definitiva y la lleva al siguiente paso que será el de la revisión de los datos.



Figura 10, Selección de filtros de búsqueda

Y la información desplegada es la elevación en metros sobre el nivel medio del mar, la procedencia de los datos, la clave de la estación y la distancia en metros al centroide del polígono trazado.

En la zona 4 y 5 son complementarias figura 8, la zona 4 es una zona de controles que nos permite acceder a la herramienta de graficas, habiendo 9 debido a que el programa controla 9 parámetros climatológicos. Y en zona 5 se despliega la cantidad de información que se tiene de cada elemento. El primer renglón se despliega el porcentaje total de datos que contenidos, en el segundo renglón el rango en que se encuentran distribuidos estos datos, o las fechas de registro que contienen las estaciones, y el número más grande se refiere a la cantidad de años registrados en la base de datos.

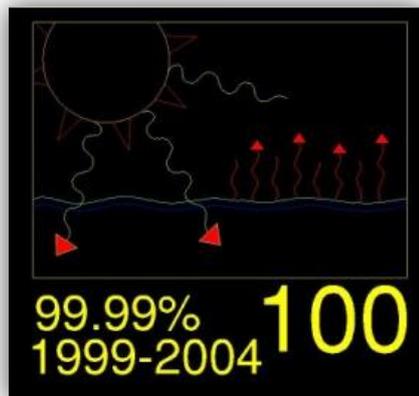


Figura 12, Resumen de datos y botón de acción.

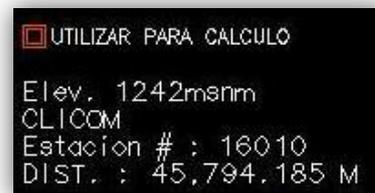


Figura 11, Despliegue de datos y opción de selección.

Al acceder a cualquiera de las graficas el área de las zonas 1, 2 y 3 es reemplazada por la herramienta de graficas figura 9.

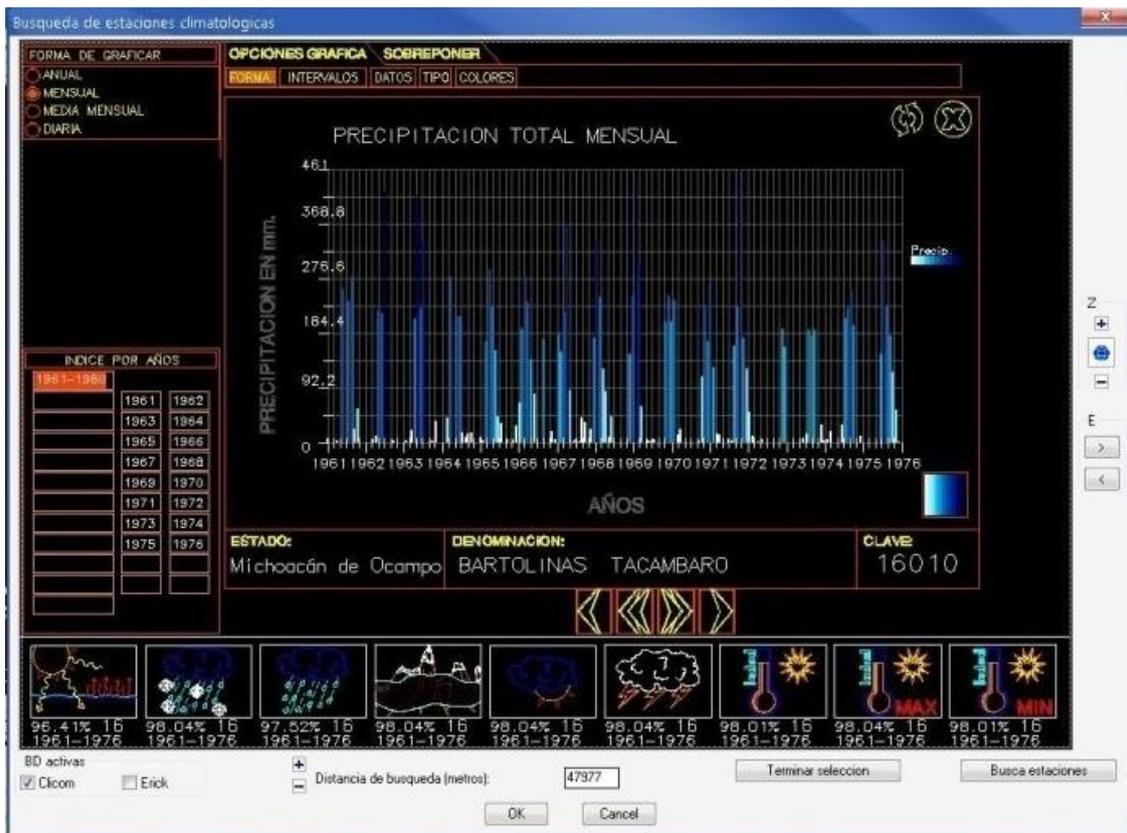


Figura 13, Herramienta de gráficos.

### Herramienta de gráficos

Esta herramienta contiene varias zonas, de despliegue de datos y varias zonas de controles, entre las cuales podemos ver:

#### ZONAS DE CONTROLES:

- Pestañas superiores.
  - Opciones de grafica
    - Muestra todas las opciones disponibles para graficar.
  - Sobreponer
    - Muestra todas las opciones disponibles para sobreponer datos
- Índice por años
  - Nos permite seleccionar el año a graficar de forma más rápida.
- Botones de opción de las pestañas
  - Estos botones se encuentran ubicados en la parte inferior de las pestañas y varían según sea la selección de la pestaña
- Zona de menús desplegables



- Se encuentra en la parte superior izquierda y ahí se muestran las opciones según sea el caso de nuestra selección en los paneles de pestaña superior y botones de opción de las pestañas.
- Botones de grafica
  - Botón de actualización de la grafica
    - Este botón actualiza la grafica según las opciones seleccionadas para graficar.
  - Botón cerrar grafica
    - No permite regresar a la visualización de todas las estaciones y el polígono de interés.

#### ZONAS DE DESPLIEGUE DE INFORMACION:

- Titulo de la grafica
  - Nos dice el tipo de grafica que realizamos, si es precipitación o evaporación, etc., y si son datos anuales mensuales o diarios, máximos mínimos o medios según sea el caso.
- Unidades a graficar
  - Escala vertical
    - Unidades medidas ya sean °C o mm, depende del dato climatológico seleccionado.
  - Escala horizontal
    - La fracción de tiempo a la que se refiere la grafica, etiquetada en años.
- Simbología de la grafica.
  - Los colores y el tipo de línea con los que se grafican cada dato.
- Estado
  - El estado en el que se encuentra ubicada la estación climatológica
- Denominación
  - El nombre con el que se ubica la estación climatológica
- Clave
  - La clave única con la cual se identifica la estación.

#### Opciones al graficar.

Dentro de las opciones a graficar tenemos 2 pestañas ya mencionadas anteriormente, opciones de grafica y sobre posición.

En las opciones para graficar tenemos 5 apartados:



- Forma
- Intervalos
- Datos
- Tipo
- Colores

En el apartado forma contiene las opciones de forma temporal a graficar y tiene 4 opciones que son:

- Anual
  - De esta forma se pueden graficar los resúmenes de los datos anuales.
- Mensual
  - Utilizando esta opción se pueden visualizar los resúmenes mensuales del dato requerido.
- Media mensual
  - Con esta opción podemos observar los resúmenes de los 12 datos 1 por cada mes según sean las demás opciones de la grafica.
- Diaria
  - Con esta opción podemos visualizar la serie temporal diaria.

En el apartado de intervalos podemos controlar en intervalo a graficar es decir la cantidad de datos que se mostraran en la graficar y puede ser anual, de 10 años, 20, 30, 40, o el total de los datos contenidos en la estación.

En la opción de datos se encuentran las opciones para controlar el tipo de datos a manejar que pueden ser:

- Máximos
  - Con esta opción podremos visualizar el dato máximo del intervalo de tiempo seleccionado en la opción forma de la pestaña de opciones de grafica.
- Medios
  - Con esta opción podremos visualizar el dato medio de todos los datos del intervalo de tiempo seleccionado en la opción forma de la pestaña de opciones de grafica.
- Mínimos
  - Con esta opción podremos visualizar el dato mínimo del intervalo de tiempo seleccionado en la opción forma de la pestaña de opciones de grafica.
- Totales



- Con esta opción podremos visualizar el dato total o la sumatoria de todos los datos del intervalo de tiempo seleccionado en la opción forma de la pestaña de opciones de grafica.
- % Vacios
  - Y con esta opción se visualizaran los porcentajes de vacios contenidos en el intervalo de tiempo seleccionado en forma.

Cabe mencionar que esta opción solo aplica en las opciones de forma anual, mensual y media mensual.

Pasando a las opciones de graficar en tipo podemos encontrar 3 opciones, las cuales nos permiten elegir qué tipo de grafica queremos, ya sea grafica de barras, dispersión de puntos, o líneas rectas. Y en las opciones de color podemos seleccionar los colores o el color con el cual se graficara.

En las opciones de sobre posición existen 2 apartados:

- Otros datos
- Otras estaciones

Y como su nombre lo dice estas opciones nos permiten sobreponer más datos climatológicos siempre y cuando estos se encuentren en las mismas unidades, y sobreponer datos de otras estaciones.

La herramienta de gráficos nos permite hacer una análisis muy superficial de las estaciones, y no nos dice si los datos son correctos o no solo nos da la posibilidad de escoger datos más que nada en base a la cantidad de datos contenidos en la estación y los años registrados.

Una vez descritas las herramientas disponibles para la selección primaria de las estaciones podemos pasar con la revisión de los datos.

## Segunda interface (revisión de los datos).

### Proceso de revisión de los datos.

Después del proceso de selección de estaciones y ya teniendo un grupo de estaciones seleccionadas, se procede a realizar una revisión en los datos, para terminar el proceso de selección y pasar al de revisión pulsamos el botón de comando “Terminar selección”, una vez pulsado tendremos otra ventana como la mostrada, figura 10



Figura 14: Cuadro de dialogo revisión de datos.

Este cuadro de dialogo al igual que el anterior cuenta con 2 zonas principales 1 zona de dibujo y otra de controles.

La zona de controles contiene solo 2 opciones una continuar y otra de cancelar.

La zona de dibujo a su vez contiene 4 zonas principales.

- Una zona variable reservada para distintos controles.
- Una zona de dibujo y despliegue de información.

- Una zona de filtros de selección de los datos y las estaciones participantes.
- Una zona de controles herramientas.

La zona reservada para distintos controles, es utilizada para desplegar controles de las distintas herramientas de análisis, según sea el caso de la herramienta activa.

En la zona de dibujo y despliegue de la información, de nuevo se despliegan las estaciones y algunos datos adicionales de la estación seleccionada.

En la zona de filtros de selección se encuentran los criterios de selección temporal numérica y de dato climatológico a utilizar, además de tener opción de datos fuente a utilizar y opción de la (las) estaciones participantes en los cálculos. En esta zona se pueden controlar los datos fuente para todo cálculo que se realizara, es decir si yo tengo seleccionados, datos anuales, de precipitación y máximos, al ejecutar cualquier herramienta estos datos será los que se tomen para el cálculo.

La zona de controles de herramientas se encuentra el control al acceso de las 4 tipos de herramientas:

- Herramientas de consistencia de datos.
- Herramientas de llenado de datos.
- Herramienta de graficas.
- Herramienta de tablas.

En el apartado de herramientas de consistencia de datos se encuentra el acceso a herramientas ya descritas para la revisión de los datos, en cuanto a pruebas básicas y pruebas específicas, figura 9.



Figura 16: Herramientas de llenado de datos.

En el apartado de llenado de datos se encuentra el acceso a las herramientas de llenado de datos, figura 10.

En el apartado de graficas nos da opción para que nos permite el acceso a las herramientas de de graficas descrita en el cuadro de dialogo anterior.

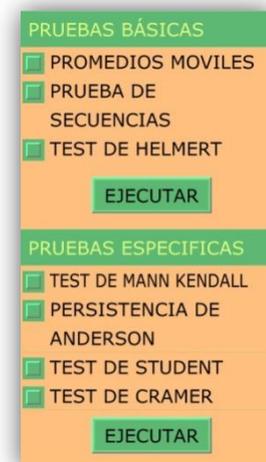


Figura 15: Herramientas de consistencia.

En el apartado de tablas se encuentran distintas opciones de generación de reportes directos a Excel, figura 11.

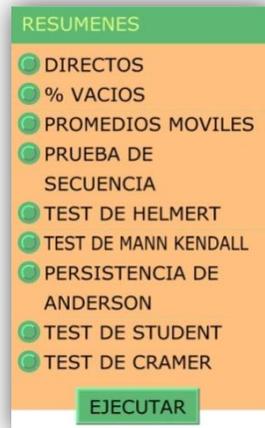


Figura 17: Herramientas de generación de reportes.

La generación de reportes y salida de resultados mucho dependen del método aplicado y en forma general se hace con un reporte en la forma como se muestra en la figura 12 y 13g.



Figura 18: Generación de reporte (prueba de secuencias).

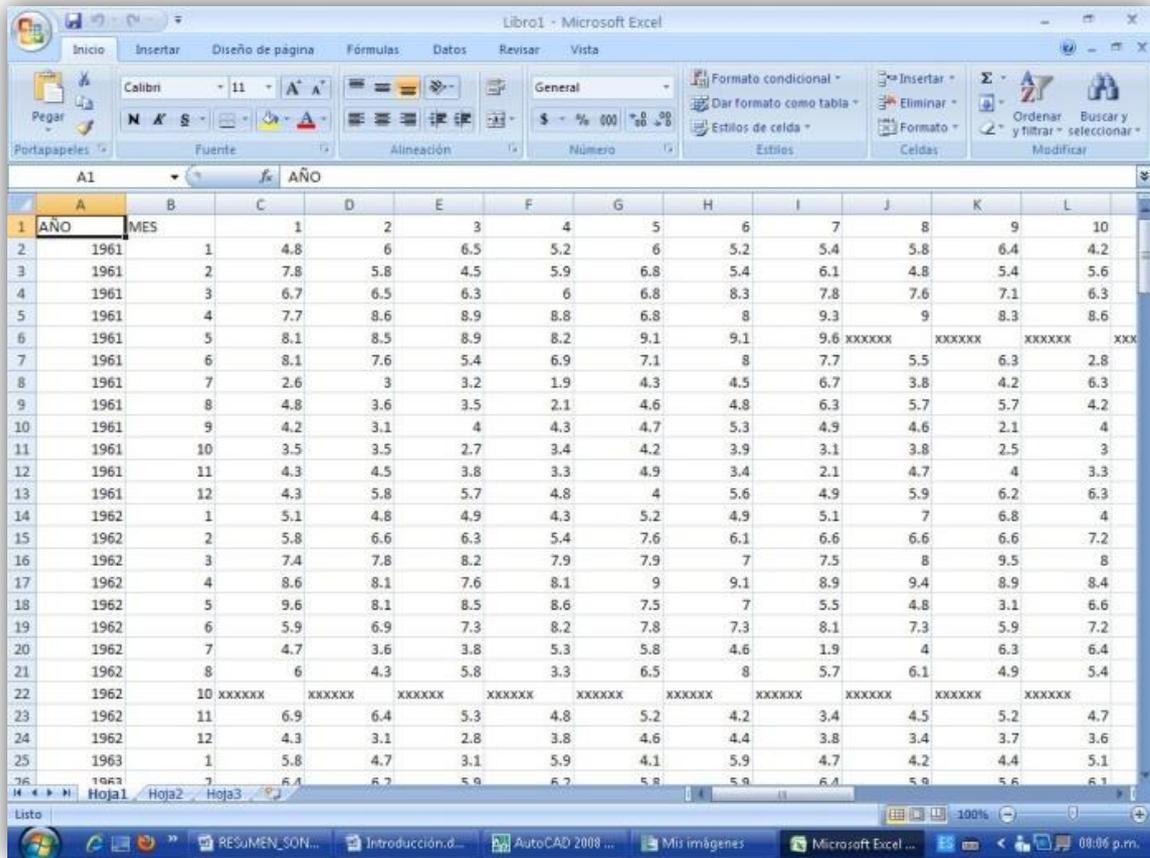


Figura 19: Generación de una tabla (datos diarios)

Dando en cada caso en particular datos específicos de cálculo y una conclusión de la aplicación del método, o en caso de una tabla generada directamente de las bases de datos los distintos tipos de datos posibles.

Tabla 6: Descripción detallada del análisis de los resultados y los datos mostrados en los reportes.

METODO	DATOS GENERADOS EN REPORTE	TIPO DE PRUEBA
Promedios móviles	Grafica	BASICA
Pruebas de secuencias	#datos, mediana, rango U, numero de cambios, conclusión	BASICA
Test de Helmert	#datos, media, numero de S, numero de C, conclusión	BASICA
Test de la T de Student	n1, n2, Sx, Tk, Td, conclusión	ESPECIFICA
Test de Cramer	n1, n2, d1, d2, Td, conclusión	ESPECIFICA
Test de Mann Kendall	#datos, Var(Z), Ecuación de ajuste, Ecuación de regresión lineal	ESPECIFICA
Regresión lineal	Ecuación de regresión lineal	ESPECIFICA



Persistencia de Anderson	#datos, rango95, rango99, datos correlograma, conclusión	ESPECIFICA
Método del U.S. Weather Service		LLENADO
Método de la relación normalizada		LLENADO

- **Promedios móviles**

El análisis con promedios móviles se realiza en forma grafica. Las series cronológicas de precipitación anual generalmente se analizan con  $N=5$ , si se analiza con un  $N$  muy bajo se pierde efectividad pues no se reduce la variabilidad, en cambio si  $N$  es muy grande se corre el riesgo que se oculten algunos componentes como la cíclica. Por medio de una línea recta se unen los datos de los periodos secos y de los periodos húmedos y los que varían mucho sobre la media, son probablemente errores en la toma o captura de los datos; analizando así de forma visual la homogeneidad de la serie.

- **Pruebas de secuencias**

Haciendo una comparación con base a la mediana. Las secuencias de A y B se contabilizan y si la cantidad de cambios esta dentro del intervalo que marca la tabla de que de acuerdo al número de datos contenidos en la serie, se concluye que no existen problemas en la serie, si se tiene un valor muy alto, se puede deber a un exceso de oscilación en la serie, por el contrario si el numero de u es muy alto se concluye que el problema de homogeneidad se puede deber a una tendencia existente o un cambio en la media.

- **Test de Helmert.**

La serie puede ser homogénea si la diferencia de secuencias S y C es igual a cero o está dentro de los límites probablemente dados por un error el cual depende de la longitud del registro.

Por tanto al aplicar el test podemos concluir si tenemos:  $S > C$  el problema de homogeneidad se puede deber a una tendencia existente o un cambio en la media.  $C > S$  puede haber un exceso de oscilación en la serie.

- **Test de la T de Student.**

El valor absoluto de la  $td$  se compara con el valor de la t de la distribución t Student de dos colas y con  $v = n_1 + n_2 - 2$  grados de libertad y con un 5% de nivel de significancia. Si el valor absoluto de  $td$  es mayor que t de la distribución se concluye que hay evidencia de falta de homogeneidad en las medias.



- **Test de Cramer.**

La estadística  $t_k$  tiene la distribución t Student de dos colas con  $v = n - 2$  grados de libertad, es utilizado de igual forma que la estadística  $td$  del método de la t de Studet.

- **Test de Mann Kendall.**

En base al valor obtenido de la pendiente de la ecuación y el valor de Z obtenido se puede analizar si las tendencias son o no significativas según se requiera.

- **Persistencia de Anderson.**

El base al correlograma, si los datos calculados en los 25 casos están dentro de los límites de Anderson se dice que la serie de datos es independiente de lo contrario se dice que tiene problemas de independencia.

En el caso de los métodos utilizados para el llenado de datos no se requiere un análisis de los resultados, solo se hace un análisis de la correlación lineal entre los datos de las estaciones.

Ya conociendo las pruebas, para un uso adecuado del programa se recomienda primero aplicar pruebas específicas y al detectar los posibles errores ir directamente a las pruebas específicas que nos podrán decir si efectivamente los errores existen.



**CAPITULO IV**  
**ANALISIS DE**  
**RESULTADOS Y**  
**REPORTE.**

CIBERMITANIOS.COM

#### 4.1 ANALISIS DE RESULTADOS Y REPORTES.

Al realizar una evaluación de resultados, o evaluación de los datos obtenidos por medio del programa, fue en base a un análisis realizado en los resultados obtenidos en un trabajo realizado en la cuenca del río Lerma en la zona ubicada en el meandro de la piedad, estos datos climatológicos no fueron revisados con los métodos estadísticos utilizados en el programa, se realizaron análisis de correlación en datos máximos anuales, y para la revisión de la homogeneidad se utilizó el método de curva masa media, con datos máximos anuales.

Para la localización de las estaciones se elaboró un mapa de las estaciones situadas en la zona de estudio:

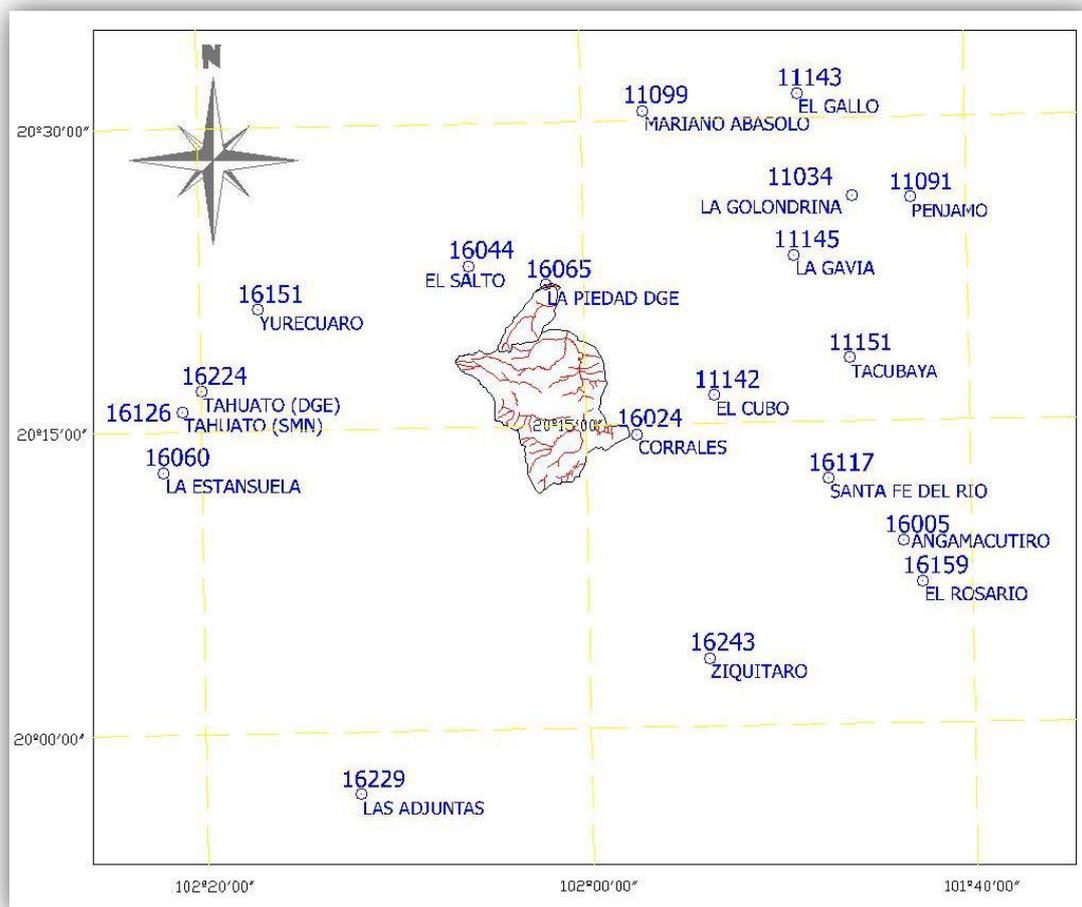


Figura 20: Mapa de localización de las estaciones.

Se realizó la búsqueda automática de las estaciones por medio del programa:





16117	STA. FE DEL RIO	bajo	S>C
16126	TANHUATO (SMN)	bien	c>s
16128	TEPALCATEPEC	bien	bien
16141	YURECUARO	bien	bien
16159	EL ROSARIO ANGACUTIRO	alto	c>s
16224	TANHUATO (DGE)	alto	bien
16243	ZIQUITARO PENJAMILLO	alto	c>s
16229	ADJUNTAS	alto	c>s
11014	CUERAMO	alto	c>s
11034	LA GOLONDRINA	bien	bien
11091	PENJAMO	bajo	S>C
11099	MARIANO ABASOLO	bien	c>s
11142	EL CUBO	bien	bien
11143	EL GALLO	bien	S>C
11145	LA GAVIA	alto	c>s
11151	TACUBAYA	bien	c>s

La tabla anterior muestra los resultados obtenidos al aplicar los métodos mencionados a los datos climatológicos contenidos en las estaciones climatológicas, utilizando los datos máximos anuales, las estaciones sombreadas en color naranja son las estaciones seleccionadas como definitivas en este estudio.

En caso de la prueba de secuencias al decir alto o bajo se refiere a que el número de secuencias cuantificadas es bajo o alto comparado con el rango permisible según el tamaño de la muestra. Si tenemos un número bajo de secuencias el resultado nos dice que puede existir errores en la tendencia o cambios de la media y si es muy alto puede haber un error de oscilación de la serie.

Y en caso de el test de Helmert se hace una comparación de C y S se refiere al número de cambios de la serie y en caso c>s esto nos indica que puede haber un error de oscilación de la serie y en caso contrario S>C puede existir un error en la tendencia de los datos.

Enfocándonos en las estaciones seleccionadas como definitivas el análisis de los datos nos dice:



Tabla 8: Resultado de las pruebas estadísticas estaciones usadas.

DATOS DE LA ESTACION		RESULTADO DEL TEST	
Clave est.	NOMBRE	Pruebas de secuencias	Test de Helmert
16024	CORRALES	-----	-----
16044	EL SALTO LA PIEDAD	-----	error de oscilación
16141	YURECUARO	-----	-----
16243	ZIQUITARO PENJAMILLO	error de oscilación	error de oscilación
11099	MARIANO ABASOLO	-----	error de oscilación
11142	EL CUBO	-----	-----

En la tabla anterior se puede apreciar claramente que puede haber un posible error de oscilación en los datos de 3 estaciones pero en una claramente se aprecia que es muy posible un error de oscilación de los datos en la estación 16243. Se puede observar que las estaciones no muestran errores en la tendencia esto debido a que la revisión de la consistencia en este estudio fue por medio de la curva masa media que en esta se puede apreciar cambios en la media o tendencias.

En este caso no aplicaría el utilizar los métodos para la revisión de los cambios en la media y la tendencia en los datos. Solo se aplicara el método para la revisión de la independencia.

Tabla 9: Resultado de las la persistencia de Anderson

DATOS DE LA ESTACION		Persistencia de Anderson
Clave est.	NOMBRE	
16024	CORRALES	-----
16044	EL SALTO LA PIEDAD	Error
16141	YURECUARO	-----
16243	ZIQUITARO PENJAMILLO	Error
11099	MARIANO ABASOLO	Error
11142	EL CUBO	-----

Al aplicar la persistencia de Anderson, test utilizado para la revisión de la independencia de los datos, se encontró efectivamente que existe un error en la independencia de los datos. La revisión de la independencia de los datos se realizo con datos mensuales y anuales máximos y en los 2 casos en resultado fue negativo. Cabe destacar que al encontrar un error de independencia no necesariamente nos dice que los datos no sirven solo nos indica un error en la independencia de los datos y la aplicación depende de criterios muy específicos para la aplicación de los mismos.



Adicionalmente a esta revisión de los datos se hizo una revisión de la correlación en los datos de las estaciones seleccionadas. La correlación se calculó con los datos máximos mensuales.

Tabla 10: Matriz de correlaciones

	<b>11142</b>	<b>11099</b>	<b>16243</b>	<b>16141</b>	<b>16044</b>	<b>16024</b>
<b>16024</b>	66%	70%	56%	65%	66%	100%
<b>16044</b>	61%	68%	62%	75%	100%	66%
<b>16141</b>	64%	71%	65%	100%	75%	65%
<b>16243</b>	42%	64%	100%	65%	62%	56%
<b>11099</b>	71%	100%	64%	71%	68%	70%
<b>11142</b>	100%	71%	42%	64%	61%	66%

La matriz de correlación es una herramienta que nos da la pauta para la selección de los tipos de métodos utilizados para el llenado de datos.

La correlación en este caso fue muy baja en este caso el método más recomendado para el llenado de los datos es el método del U.S. WEATHER SERVICE. En caso de contener una correlación del 90% o mayor se puede utilizar el método de la relación normalizada.



## CONCLUSIONES

El programa a pesar de mostrar cierta facilidad en el uso y aplicación de las múltiples herramientas estadísticas, su correcto uso interpretación y aplicación requiere un cierto grado de conocimiento sobre el tema.

La última actualización de esta versión del programa (1.0), de los datos climatológicos se realizo a finales del septiembre del 2009.

Las múltiples herramientas estadísticas solo son elementos que simplifican los criterios de la selección de los datos, es decir si son o no útiles para el propósito requerido.

La herramienta estadística para la revisión de la tendencia, evalúa una tendencia en forma de pendiente de una line recta ajustada a los datos el valor de Z nos permite decir si son tendencias validas o no.

## RECOMENDACIONES.

Para la correcta aplicación del las herramientas se requiere un conocimiento mínimo del marco teórico en que está basado el programa.

El programa fue totalmente probado en Autocad 2008 y trabajando en un sistema operativo Windows XP (SP3), para asegurar su mejor rendimiento se recomienda usarlo bajo estas características. Para generación de reportes el programa fue probado utilizando Office 2007.

Se recomienda para agilizar la lectura de los datos, utilizar un ordenador con un rápido acceso a los datos del disco duro, y un bus de datos con una traza de transferencia de datos alta.

Los resultados generados por los reportes y las herramientas pueden no ser idénticos a algunos generados con los datos fuente originales, esto debido a la depuración realizada en las bases de datos originales.



## REFERENCIAS

- I. Daniel Francisco Campos Aranda, (2007), Estimación y aprovechamiento del escurrimiento, San Luis Potosí.
- II. Daniel Francisco Campos Aranda, (1998), Procesos del ciclo hidrológico, San Luis Potosí.
- III. J.D. Salas, J.W. Delleur, V. Yevjevich and W.L. Lane, (1985), Applied Modeling of Hydrologic time Series.
- IV. Timo Salmi, Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's slope estimates –the Excel template application, August 2002, MAKESENS, [http://www.fmi.fi/kuvat/MAKESENS\\_MANUAL.pdf](http://www.fmi.fi/kuvat/MAKESENS_MANUAL.pdf)
- V. William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery, (1986-1992), Numerical Recipes, Cambridge Universe.
- VI. Ronald E. Walpole Raymond H. Myers, (2010), Probabilidad y estadística. 09810 D.F. México.
- VII. N. Plummer, W. Lipa, S. Palmer, G. Prank, J. Shortridge, D. Stuber, (2007) Directrices sobre la gestión de datos climáticos, Ginebra, organización meteorológica mundial.