



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO
INSTITUTO INTERAMERICANO DE TECNOLOGÍA Y CIENCIAS
DEL AGUA

TESIS

DISEÑO DE UNA HERRAMIENTA HIDRO-INFORMÁTICA PARA EL
ANÁLISIS DE CALIDAD Y COMPLEMENTO DE DATOS DE
ESTACIONES METEOROLÓGICAS.

QUE PARA OBTENER EL GRADO DE:

MAESTRA EN CIENCIAS DEL AGUA

PRESENTA:

ING. NANCY SALGADO ÁLVAREZ

DIRECTOR DE TESIS:

DR. CARLOS
DÍAZ DELGADO

CO-DIRECTORA

DRA. ALEIDA YADIRA
VILCHIS FRANCÉS

CO-DIRECTOR

DR. KHALIDOU
MAMADOU. BÂ

Diciembre 2020

Los datos meteorológicos de calidad generan beneficios al usuario, ya que reducen la incertidumbre al ser utilizados en actividades como climatología, agricultura, etc. A nivel mundial la Organización meteorológica mundial (OMM), en conjunto con los servicios meteorológicos nacionales (SMN) no cuenta con una normativa estandarizada para analizar la calidad de los datos meteorológicos antes ponerlos a disposición de los usuarios. En México se encuentra vigente la norma NMX-AA-166/1-SCFI-2013 (NMX-AA-166/1-SCFI, 2013) sobre estaciones meteorológicas, pero ésta no refleja una metodología para el análisis de calidad de datos.

La falta de esta normativa ha sido tema de investigación de diversos autores, proponiendo diversas metodologías de análisis de calidad de datos meteorológicos. Por lo cual, en el presente proyecto se desarrolló de una herramienta hidro- informática para el análisis de calidad y complemento de datos meteorológicos. Con base en técnicas y métodos actualmente utilizados a nivel global y reportados en bibliografía especializada para que con el uso de esta herramienta se facilite la verificación de la calidad de los datos que se están utilizando. Esta metodología consta de la revisión de calidad por medio de umbrales físicos, coherencia interna, coherencia espacial, homogeneidad, y el complemento de datos por medio de la interpolación espacial. Adecuando cada uno de estos métodos para las variables meteorológicas de temperatura, precipitación, dirección y velocidad del viento, radiación solar, y presión atmosférica.

Esta herramienta lleva por nombre SICDAMET que quiere decir “Sistema de control de calidad de datos meteorológicos” y contiene una metodología de análisis de calidad basada en métodos estadísticos que han sido probados para evaluar la calidad de la información de la red de estaciones meteorológicas automáticas y convencionales del Estado de México. El

resultado obtenido consiste en la evaluación de la calidad de datos que ofrecen las diversas dependencias que operan estas estaciones.

Adicionalmente, los beneficios que presenta la herramienta hidro-informática son: la organización automatizada de los diversos formatos de registro de variables de meteorológicas que proporcionan las distintas dependencias. Lo anterior, facilita el análisis de calidad y complemento de información meteorológica, proporcionando al usuario diversos reportes y gráficas que disminuyen la incertidumbre de la calidad de los datos meteorológicos.

Abstract

Quality meteorological data generates benefits for the user, since they reduce uncertainty when used in activities such as climatology, agriculture, etc. At the global level, the World Meteorological Organization (WMO), together with the national meteorological services (SMN), does not have a standardized regulation to analyze the quality of meteorological data before making it available to users. In Mexico, the NMX-AA-166/1-SCFI-2013 (NMX-AA-166/1-SCFI, 2013) standard on meteorological stations is in force, but it does not reflect a methodology for data quality analysis.

The lack of these regulations has been the subject of research by several authors, proposing various methodologies for analyzing the quality of meteorological data. Therefore, in this project a hydro-informatics tool was developed in order to perform the quality analysis and complement of meteorological data. Based on techniques and methods currently used globally and reported in specialized bibliography so that it can make sure of the quality of the data it is using. This methodology consists of a quality review through physical thresholds, internal coherence, spatial coherence, homogeneity, and the complement of data through spatial interpolation. Adapting each of these methods for the meteorological variables of temperature, precipitation, wind direction and speed, solar radiation, and atmospheric pressure.

This tool is called SICDAMET which means "Meteorological data quality control system" and contains a quality analysis methodology based on statistical methods which have been tested to evaluate the quality of the information of the meteorological network of the State of Mexico taking into account the two types of weather stations (automatic and conventional). Additionally, the benefits of the hydro-informatics tool are the automated organization of the various meteorological variables record formats provided by the different agencies. Thus,

facilitating the quality analysis and complement of meteorological information, providing the user with various reports and graphs that reduce the uncertainty of the quality of meteorological data.

Agradecimientos	1
Resumen.....	3
Abstract	5
Índice	8
Índice de figuras.....	12
Índice de tablas	16
Acrónimos.....	18
Simbología.....	20
Capítulo 1: Introducción	22
1.1.2 Antecedentes	23
1.1.1 Meteorología	24
1.1.2 Análisis de calidad de datos meteorológicos.....	24
1.1.3 Deficiencias de la información meteorológica en México.....	27
1.2 Justificación contextual y científica.....	29
1.3 Hipótesis	30
1.4 Objetivos	31
1.4.1 Objetivo general.....	31
1.4.2 Objetivos específicos	31
Capítulo 2: Gestión de calidad de los datos meteorológicos	32
2.1 Normatividad para la gestión de los datos meteorológicos	32
2.2 Fuentes de información meteorológica.....	34
2.2.1 Servicio Meteorológico Nacional (SMN).....	34
2.2.2 Instituto Nacional de Investigación Forestal, Agrícolas y Pecuarias (INIFAP)..	35

2.2.3 Otras dependencias que brindan información meteorológica.....	36
2.3 Características e instrumentos de las estaciones meteorológicas	36
2.3.1 Estación meteorológica convencional (EMC)	36
2.3.2 Estación meteorológica automatizada (EMA)	37
2.3.3 Estación sinóptica meteorológica (ESIME).....	38
2.3.4 Instrumentos de las EMC, EMA y ESIME.....	39
2.4 Aplicación de las variables meteorológicas.....	40
2.5 Zona de estudio	41
Capítulo 3: Marco teórico	45
3.1 Fases del análisis de calidad de los datos meteorológicos.....	45
3.2 Métodos de umbrales físicos.....	47
3.2.1 Umbrales del sensor	47
3.2.2 Desviación estándar	48
3.2.3 Boxplot	48
3.3 Métodos de coherencia	50
3.3.1 Coherencia interna.....	50
3.3.1.1 Umbrales físicos	50
3.3.1.2 Desviación estándar temporal	51
3.3.2 Coherencia espacial	51
3.3.2.1 Comparación con otras estaciones.....	52
3.3.2.2 Comparación con datos de huracanes.....	53
3.4 Métodos de homogeneidad	54
3.4.1 Distribución Normal	55
3.4.2 Distribución Log-Normal	57
3.4.3 Distribución Gamma	57
3.4.4 Distribución Log-Pearson III	58
3.4.6 Distribución T-Student	59
3.5 Métodos para complementar series.....	60
3.5.1 Método de Kriging	60
3.5.1.1 Kriging con factor de cambio externo.....	62
3.5.2 Validación cruzada de vecindario único	63

3.6	Métodos de aplicación para cada variable meteorológica y tipo de estación	65
3.7	Información relevante de los datos meteorológicos	67
3.7.1	Gráficas	67
3.7.2	Resumen estadístico	68
3.8	Matlab	70
Capítulo 4: Materiales y métodos		71
4.1	Base de datos meteorológicos.....	71
4.1.1	Bases de datos de estaciones meteorológicas convencionales (EMC's)	72
4.1.2	Bases de datos de estaciones meteorológicas automáticas (EMA's).....	74
4.1.2.1	Bases de datos del SMN.....	74
4.1.2.2	Bases de datos del INIFAP	76
4.1.2.3	Bases de datos de CAEM.....	77
4.1.2.4	Bases de datos de PEMBU	79
4.1.2.3	Bases de datos del IITCA	80
4.1.2.4	ID oficial de las variables meteorológicas.....	81
4.1.3	Formatos de entrada la herramienta hidro-informática	82
4.2	Base de datos de huracanes.....	84
4.3	Desarrollo y programación de la herramienta hidro-informática	84
4.3.1	Interfaz de la herramienta hidro-informática.....	85
Capítulo 5: Resultados		87
5.1	Herramienta hidro-informática "SICCDAMET"	87
5.1.2	Reportes de resultados	92
5.1.2.1	Reporte por método	92
5.1.6.2	Reporte general	97
5.2	Validación de la herramienta hidro informática SICCDAMET	100
5.2.1	Organización de los datos	100
5.2.2	Métodos de análisis	102
5.3	Análisis a la base de datos	107
5.3.1	Base de datos de las EMA's.....	109
5.3.2	Base de datos de las EMC's.....	118

5.3.3 Interpolación de datos diarios.....	120
5.4 Discusión de resultados	129
Capítulo 6: Conclusiones y recomendaciones	131
6.1 Conclusiones generales.....	131
6.2 Recomendaciones	134
Bibliografía	136
Anexos digitales.....	143
Anexo 4.1 Clave de los estados de la República	143
Anexo 4.2” Diagramas de flujo por métodos de la herramienta hidro-informática “SICCDAMET.....	143
Anexo 4.3” Diagramas de flujo de la interfaz de la herramienta hidro-informática “SICCDAMET.....	143
Anexo 5.1 Manual de usuario de la herramienta hidro informática SICCDAMET	143
Anexo 5.2 Ejemplos de reportes de resultados para la estación Observatorio UAEM...	143
Anexo 5.3 Validación del análisis de calidad a la base de datos de la estación Observatorio UAEM.....	143
Anexo 5.4 Listado de estaciones consideradas en el proyecto	143
Anexo 5.5 Resultados del análisis de calidad a la base de datos de las EMA’s	143
Anexo 5.6 Imágenes de resultados del análisis de calidad a la base de datos de las EMA’s	143
Anexo 5.7 Resultados del análisis de calidad a la base de datos de las EMC’s	143
Anexo 5.8 Imágenes de resultados del análisis de calidad a la base de datos de las EMC’s	144
Anexo 5.9 Resultados de la interpolación de T _{máx} , T _{mín} y P	144
Anexo 5.10 Resultados de la simulación de P para el día 01/07/2016	144

Índice de figuras

	Nombre	Pág.
Figura 2.1	Estación meteorológica convencional (EMC).....	36
Figura 2.2	Estación meteorológica automática.....	37
Figura 2.3	Estructura de una ESIME.....	38
Figura 2.4	Estaciones meteorológicas del mapex del Estado de México.....	42
Figura 3.1	Gráfico boxplot.....	49
Figura 3.2	Método de comparación con estaciones cercanas.....	52
Figura 3.3	Trayectoria de huracanes con estaciones meteorológicas.....	54
Figura 3.4	Campana de Gauss.....	55
Figura 3.5	Intervalos de confianza de la distribución Normal.....	56
Figura 3.6	Ejemplo muestreo para la interpolación con Kriging.....	61
Figura 4.1	Formato de registro CLICOM.....	73
Figura 4.2	Formato tipo 1 estación meteorológica “ALAMOS, en el Estado México.....	74
Figura 4.3	Formato tipo 2 estación meteorológica “ALAMOS, en el Estado México.....	75
Figura 4.4	Formato tipo 3 estación meteorológica “ALAMOS, en el Estado México.....	75
Figura 4.5	Formato para las ESIME’s del SMN estación “CEMCAS”).....	76
Figura 4.6	Formato 1 de la base de datos del INIFAP.....	77
Figura 4.7	Formato 2 de la base de datos del INIFAP.....	77
Figura 4.8	Formato 1 de la base de datos de CAEM; estación “Presa Allende”...	78

Figura 4.9	Formato 2 de la base de datos de CAEM; estación “Presa Allende”...	78
Figura 4.10	Formato 3 de la base de datos de CAEM; estación “Presa Allende”...	79
Figura 4.11	Formato 4 de la base de datos de CAEM; estación “Presa Allende”.....	79
Figura 4.12	Formato de la base de datos de PEMBU.....	80
Figura 4.13	Formato 1 de la estación del IITCA (2001-2007).....	80
Figura 4.14	Formato 2 de la estación del IITCA (2008-actualidad).....	81
Figura 4.15	Formato 3 de las estaciones 2 y 3 del IITCA.....	81
Figura 4.16	1 Formato 1 de la herramienta hidro informática.....	83
Figura 4.17	Formato 2 de la herramienta hidro informática.....	83
Figura 4.18	Formato para la base de datos de huracanes.....	84
Figura 4.19	Diagrama de flujo de las etapas generales de la herramienta hidro-informática.....	86
Figura 5.1	Logotipo y nombre de la herramienta hidro-informática.....	88
Figura 5.2	Interfaz de bienvenida al usuario en la herramienta hidro-informática.....	89
Figura 5.3	Selección de la estación a analizar.....	90
Figura 5.4	Interfaz básica de la herramienta hidro informática SICCDAMET...	90
Figura 5.5	Interfaz de los métodos de análisis.....	91
Figura 5.6	Ejemplo de reportes de resultados por método.....	93
Figura 5.7	Ejemplo de reportes de resultados para el método de coincidencia con huracanes.....	94
Figura 5.8	Reporte de validación de la interpolación.....	95
Figura 5.9	Reporte de resultados del variograma de la interpolación.....	96
Figura 5.10	Ejemplo de archivo *.tiff de interpolación para T _{máx}	97
Figura 5.11	Formato del reporte general de análisis de calidad de datos.....	98
Figura 5.12	Selección de la estación a analizar	100
Figura 5.13	Porcentaje promedio de resultados de precipitación [P] para EMA’s.....	110
Figura 5.14	Porcentaje promedio de resultados de temperatura ambiente [T _{amb}] para EMA’s.....	111

Figura 5.15	Porcentaje promedio de resultados de humedad relativa [Hr] para EMA's.....	112
Figura 5.16	Porcentaje promedio de resultados de dirección del viento [Dv] para EMA's.....	113
Figura 5.17	Porcentaje promedio de resultados de velocidad del viento [Vv] para EMA's.....	114
Figura 5.18	Porcentaje promedio de resultados de presión atmosférica [Pb] para EMA's.....	115
Figura 5.19	Porcentaje promedio de resultados de radiación solar [Rs] para EMA's.....	116
Figura 5.20	Mapex del Estado de México con porcentajes obtenidos para el método de desviación estándar aplicado a la variable de T _{máx} de las EMA's.....	117
Figura 5.21	Gráfica de los porcentaje promedio de los valores aprobados en la base de datos CLICOM.....	119
Figura 5.22	Mapex del Estado de México con porcentajes obtenidos para el método de desviación estándar aplicado a la variable de T _{máx} de las EMC's.....	120
Figura 5.23	Ejemplos de archivos de entrada para la interpolación.....	121
Figura 5.24	Mapa digital de elevaciones del mapex del Estado de México con la ubicación de las estaciones seleccionadas para la interpolación.....	121
Figura 5.25	Imagen de la interpolación de T _{máx} para el día 15 de agosto del 2016	123
Figura 5.26	Variograma del día 15 de agosto para T _{máx}	123
Figura 5.27	Imagen de la interpolación de T _{mín} para el día 15 de agosto del 2016	124
Figura 5.28	Variograma del día 15 de agosto para T _{mín}	124
Figura 5.29	Gráficas de R ² de ajuste del variograma y regresión de las variables T _{máx} y T _{mín}	126
Figura 5.30	Ejemplo de interpolación de precipitación 15/08/2016.....	127
Figura 5.31	Ejemplo de interpolación de precipitación 01/07/2016.....	128
Figura 5.32	Resultados de la simulación del día 01/07/2016.....	129

Índice de tablas

	Nombre	Pág.
Tabla 1.1	Métodos de análisis de calidad de datos meteorológicos.....	26
Tabla 2.1	Normativas destacadas para la gestión de datos meteorológicos.....	33
Tabla 2.2	Variables de las estaciones meteorológicas.....	39
Tabla 2.3	Número de estaciones meteorológicas del mapex del Estado de México.....	43
Tabla 2.4	Estaciones meteorológicas consideradas.....	44
Tabla 3.1	Rangos de medición del sensor.....	47
Tabla 3.2	Métodos de aplicación para cada variable meteorológica.....	66
Tabla 3.3	Gráficas más usuales para representar datos meteorológicos.....	67
Tabla 4.1	ID de las variables en CLICOM.....	73
Tabla 4.2	ID oficial de las variables metrológicas.....	82
Tabla 5.1	Validación de la extracción de datos.....	101
Tabla 5.2	Validación de la estación Observatorio UAEM.....	102
Tabla 5.3	Validación manual de los resultados coherencia con estaciones cercanas para la estación Observatorio UAEM.....	104
Tabla 5.4	Validación manual de los resultados de coherencia con huracanes para la estación Observatorio UAEM.....	105
Tabla 5.5	Validación cruzada de la interpolación del día 15/08/2016 de la variable T _{máx}	106
Tabla 5.6	Total de registros de la base de datos del mapex del Estado de México.....	108

Tabla 5.7	Porcentaje de datos aprobados de las Ema's para la variable de precipitación [P].....	110
Tabla 5.8	Porcentaje de datos aprobados de las Ema's para la variable de temperatura ambiente [Tamb].....	111
Tabla 5.9	Porcentaje de datos aprobados de las Ema's para la variable de humedad relativa [Hr].....	112
Tabla 5.10	Porcentaje de datos aprobados de las Ema's para la variable de dirección del viento [Dv].....	113
Tabla 5.11	Porcentaje de datos aprobados de las Ema's para la variable de velocidad del viento [Vv].....	114
Tabla 5.12	Porcentaje de datos aprobados de las Ema's para la variable de presión atmosférica [Pb].....	115
Tabla 5.13	Porcentaje de datos aprobados de las Ema's para la variable de radiación solar [Rs].....	116
Tabla 5.14	Porcentaje de datos aprobados de la base de datos CLICOM.....	118

Acrónimos

CAEG	Comisión del Agua del Estado de Guanajuato
CAEM	Comisión del Agua del Estado de México
Chaak	Computational Hydrology and Hydraulics Adier Kit
CLICOM	Sistema CLIma COMputarizado
CONACyT	Consejo Nacional de Ciencia y Tecnología
CONAGUA	Comisión Nacional del Agua
DEM	Modelo Digital de Elevaciones
EMA's	Estaciones Meteorológicas Automáticas
EMC's	Estaciones Meteorológicas Convencionales
ESIME's	Estaciones Sinópticas Meteorológicas
GUI	Interfaz gráfica
ID	Identificador
I-D-Tr	Intensidad-Duración-Periodo de retorno
INAI	Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos
IITCA	Instituto Interamericano de Tecnología y Ciencias del Agua
INIFAP	Instituto Nacional de Investigación Forestal, Agrícola y Pecuaria
NOAA	National Oceanic and Atmospheric Administration
NOF	Norma Oficial
OCAVM	Organismo de Cuencas del Valle de México

OCLSP	Organismo de Cuencas Lerma Santiago Pacífico
OMM	Organización Meteorológica Mundial
PC Guanajuato	Protección Civil de Guanajuato
PEMBU	Programa de Estaciones Meteorológicas del Bachillerato Universitario
SEMARNAT	Secretaría de Medio Ambiente y Recursos Naturales
SICCDAMET	Sistema de Control de Calidad de Datos Meteorológicos
SIG	Sistema de Información Geográfica
SMN	Servicio Meteorológico Nacional
UAEM	Universidad Autónoma del Estado de México
UAEMEX	Comisión del Agua del Estado de México
UNE	Asociación Española de Normalización
UTC	Tiempo universal coordinado
WMO	World Meteorological Organization
MATLAB	Matrix Laboratori

Símbolo	Representación
Dv	Dirección del viento
Ev	Evaporación
Hr	Humedad relativa
P	Precipitación
Pb	Presión atmosférica
Rs	Radiación solar
Tamb	Temperatura ambiente
Tmáx	Temperatura máxima
Tmed	Temperatura media
Tmín	Temperatura mínima
Vv	Velocidad del viento
°	Grados
m.s.n.m	Metros sobre el nivel del mar
mm	Milímetros
%	Porcentaje
Bar	Bar
W	Watts
°C	Grados Celsius
°C	Grados Celsius
°C	Grados Celsius
m/s	Metros por segundo
Km	Kilometro
*.csv	Extensión de archivo tipo Microsoft Excel

- *.txt Extensión de archivo tipo texto
- *.pdf Extensión de archivo tipo Portable Document Format
- *.tiff Extensión de archivo de imagen tipo Tagged Image File Format
- *.xlsx Extensión de archivo tipo Microsoft Excel
- *.mat Extensión de archivo tipo MATLAB

Capítulo 1

Introducción

La generación de datos meteorológicos confiables produce beneficios al usuario cuando esta información se utiliza en actividades como climatología, agricultura, diseño y operación de infraestructura, etc. Sin embargo, a nivel global la Organización Meteorológica Mundial (OMM), en conjunto con los servicios meteorológicos de cada región, no cuentan con una normativa estandarizada para analizar la calidad de los datos meteorológicos antes de publicarlos. La falta de esta normativa ha sido tema de investigación de diversos autores, proponiendo para ello diferentes metodologías de análisis (Serrano et al., 2016).

En México se encuentra vigente la norma NMX-AA-166/1-SCFI-2013 (NMX-AA-166/1-SCFI, 2013) sobre estaciones meteorológicas automáticas (EMA's). Pero se destaca que esta norma no es obligatoria en su aplicación, y no contiene un método explícito para el análisis de calidad de datos.

Por lo cual, analizar la calidad de los datos meteorológicos es uno de los principales aspectos en la meteorología. Y donde la automatización de los procesos de verificación de calidad facilitará el uso y aplicación de esta información. Para ello, en el presente proyecto se presenta el desarrollo una herramienta hidro-informática que tiene la capacidad de analizar la calidad y complementar las series de datos de las estaciones meteorológicas automáticas y convencionales.

Esta herramienta se basa en técnicas y métodos utilizados a nivel global y reportados en bibliografía especializada, dividiendo el análisis en:

- Revisión de umbrales físicos
- Revisión de coherencia interna
- Revisión de coherencia espacial
- Revisión de homogeneidad
- Facilitar el complemento de datos faltantes en las series

También, tiene la capacidad de ordenar y dar un formato uniforme a las bases de datos que analice y complemente. Para ejemplificar los métodos de análisis y corroborar el correcto funcionamiento de la herramienta, se revisa la calidad de la base de datos de las estaciones meteorológicas del Estado de México. Este análisis ha considerado las variables meteorológicas de precipitación (P), temperatura máxima, mínima y media (Tmáx, Tmín, Tmed, respectivamente), velocidad y dirección del viento (Vv y Dv), presión atmosférica (Pa), humedad relativa (Hr), evaporación (Evp) y radiación solar (Rs). Obteniendo como resultados, un resumen estadístico de los datos, calificación del análisis, series de datos completas hasta donde la información lo permitió y algunas recomendaciones de aplicación.

1.1.2 Antecedentes

La meteorología es una disciplina compleja, que requiere una observación frecuente de estados del tiempo en todo el planeta; y está relacionada con la evolución de las ciencias físicas, químicas y astronómicas, así como con el desarrollo tecnológico. La investigación meteorológica se basa en el registro y observación del cambio que sufren los diferentes periodos de tiempo y aplicados a diversas actividades (Castro, 2008). El desarrollo científico y tecnológico alcanzado a la fecha, permite disponer y utilizar herramientas como las estaciones meteorológicas, las cuales conforman un sistema de monitoreo climático que representa un avance trascendente, ya que permite generar información que apoya la toma de decisiones en diferentes ámbitos de aplicación como la aeronáutica, agricultura, climatología, hidrología, predicción de eventos naturales, etc. (Maldonado & Aravena, 2006).

1.1.1 Meteorología

El monitoreo meteorológico consiste en una red de estaciones meteorológicas localizadas en diferentes puntos geográficos. Esta red está compuesta por Estaciones Meteorológicas Convencionales (EMC's), Estaciones Meteorológicas Automáticas (EMA's) y Estaciones Sinópticas Meteorológicas (ESIME's) las cuales hacen registros diarios y semihorarios respectivamente. Esta red es monitoreada por el Servicio Meteorológico Nacional (SMN) de cada región o país, tomando en cuenta normas y criterios establecidos por la OMM.

Cabe resaltar que el objetivo de la OMM es asegurar y facilitar la cooperación entre los SMN, así como proveer normatividad para la operación y calibración de instrumentos meteorológicos. También establece las normas para obtener y procesar sus bases de datos, para que lleguen a los usuarios con una buena calidad (OMM, 2012). Las diversas redes de EMC's, EMA's y ESIME's recopilan gran cantidad de información atmosférica. Toda esta información no sólo sirve para estudios históricos retrospectivos, también forma parte fundamental de los datos de ingreso a los modelos de pronósticos meteorológicos. Desde mediados del siglo XX los modelos de pronóstico numérico han sido de gran ayuda para la toma de decisiones ante diversos tipos de contingencias ambientales. En la actualidad son la principal herramienta ante la proximidad de fenómenos que pueden ser muy destructivos, entre los que se encuentran los ciclones tropicales y los frentes fríos (Prieto, 2008), así como en actividades de agricultura, infraestructura e investigación (Castro, 2008)

1.1.2 Análisis de calidad de datos meteorológicos

Cuando se tiene que trabajar con datos meteorológicos es necesario asegurar que sean de buena calidad, ya que muchos de éstos generalmente no son sometidos a un control de calidad. Estos datos son recuperados y almacenados de forma instantánea, para ser recibidos con suficiente antelación y ser utilizados por los usuarios. La calidad del dato climático está relacionada con la ubicación de la estación meteorológica, la calibración de sus instrumentos de medición, la experticia del operador que realiza las observaciones, entre otros aspectos (OMM, 2012).

La mayoría de los errores se pueden verificar en el momento de la observación y archivado de los datos; sin embargo, persisten debilidades en los protocolos de revisión de observaciones meteorológicas de origen humano. Además, pueden existir datos faltantes, o bien valores extraños que no cumplen con el comportamiento acostumbrado de la variable meteorológica. A nivel internacional se ha comprobado que datos erróneos pasan inadvertidos cuando no se aplican protocolos especializados para el control de calidad; sobre todo valores atípicos infiltrados que afectan la homogeneidad y análisis espacio-temporal de las series. Así mismo, se ha observado que las series incompletas suelen producir consecuencias negativas en la generación de conocimiento (N., P., & Pomares P., 2015). Para lo anterior, es necesario evaluar la confiabilidad de estos datos mediante la aplicación de técnicas estadísticas, las cuales permiten obtener resultados rápidos y confiables (Oviedo & Rojas, 2012). En 1992 la OMM sugiere que el control de calidad de datos fuese dividido en varias fases de acuerdo con la función y requerimientos de los datos meteorológicos. Estas fases comprenden la verificación de:

- I. Errores gruesos (que no tengan una relación aproximada con los demás datos).
- II. Coherencia interna
- III. Coherencia temporal.
- IV. Coherencia espacial

Este análisis de control de calidad de datos, fue propuesto para EMC's y sigue vigente para las EMA's y ESIME's (OMM, 2012). Sin embargo, parece no ser confiable ya que los métodos para la toma de datos son diferentes en cada tipo de estación. Por ello, diversos países han optado por el desarrollo de nuevos modelos de control de calidad, realizando una combinación de controles manuales y automáticos para evaluar y validar los datos meteorológicos con la finalidad de aplicarlos en actividades como la agricultura, elaboración de pronósticos meteorológicos, etc. (Ramchandra, 2010).

Estos modelos se basan en métodos estadísticos-probabilísticos y contemplan la revisión de umbrales, coherencia, homogeneidad, correlación y el complemento de datos faltantes. La

tabla 1.1 muestra los métodos que contemplan algunos países para el control de calidad, éstos se marcan con un * (asterisco).

Tabla 1.1 Métodos de análisis de calidad de datos meteorológicos

Bibliografía		Comprobación de equipo	Umbrales			Coherencia		Homogeneidad					Correlación	Completar serie				Corrección de datos	Gráficas	
País	Cita		Análisis visual	Datos de la zona	Desviación estándar	Boxplot	Interna	Espacial	Gama Normal	Pearson III	F-Fisher	Log Normal		T-Student	Interpolación lineal	Promedios aritméticos	Espacial (Kriging)			Regresión
EE.UU	(Hansen & Poulain, 1995)													•						
Suecia	(Hans A. & Mober A., 1997)								•											
USSR	(Gandin L. S., 1988)								•					•						
España	(Aguilar et al., 1999)			•			•	•									•			
Noruega	(Vejen., et al., 2002)		•																	
España	(Abaurrea, et al., 2004)													•						
EE.UU	(Hubbard K .G., et al., 2006)			•													•			
India	(Machival, et al., 2006)								•				•	•						
Canadá	(Quarda & Khaliq , 2007)								•											
Costa Rica	(Araya J. & Alfaro E.J., 2009)			•	•															
Argentina	(Avarado & Espinoza, 2013)			•	•									•						
Recomendación mundial	OMM(2012)	•	•														•			
Colombia	(Oviedo & Rojas (2012)							•	•	•	•	•	•							
Perú	(Boshell & León, 2013)			•	•												•			
América del sur	(Aguilar A., 2013)				•									•			*			
Venezuela	(Olivares., et al., 2013)			•	•	•			•					•	•					
España	(Miró, et al., 2014)								•											
España	(Bakirón, 2014)													•			•			
Colombia	(Lanos. H., 2018)				•		•	•	•	•	•	•	•	•	•	•	•	•		
Europa	(SHULSKI, ET. AL., 2014)				•				•					•						
España	(Vicente Serrano., et al., 2016)			•		•	•													
México	(CICOM-2018)																	•		

Como se muestra en la tabla anterior, existen diferentes métodos empleados para el análisis de calidad de datos y complemento de series de datos meteorológicos. En México, las dependencias que reportan datos meteorológicos no aplican un análisis de control de calidad, o al respecto no se encuentra nada documentado en las fuentes consultadas. Por lo cual, son datos sin validación que transmiten la incertidumbre al ser aplicados en procesos y toma de decisiones en los que son utilizados.

1.1.3 Deficiencias de la información meteorológica en México

La red meteorológica de México se encuentra administrada por distintas dependencias como son: Comisión Nacional del Agua (CONAGUA), Instituto Nacional de Investigación Forestal, Agrícola y Pecuaria (INAFAP) y dependencias no gubernamentales como el Programa de Estaciones Meteorológicas del Bachillerato Universitario (PEMBU), Comisión del Agua del Estado de México (CAEM), Instituto Interamericano de Tecnología Ciencias del Agua (IITCA), entre otras. Estas dependencias cuentan con una red de monitoreo meteorológico que transmite la información a través del internet, llegando a los usuarios sin tener una certificación de calidad en las mediciones (Araya & Alfaro, 2009), (INIFAP, 2019). Por lo cual, la información no debería ser considerada como confiable.

Según Prieto (2008), el marco jurídico que se aplica en México para actividades relacionadas con los servicios de información meteorológica tiene los siguientes defectos:

- No responde ni refleja una política clara y precisa.
- Es insuficiente, ya que no refleja una normativa integradora con autoridades e instancias articuladoras.
- No existe un órgano, instancia, comisión, grupo de trabajo, eje conductor, autoridad única.
- Carece de auditorías para certificar el mantenimiento de las estaciones, así como su correcto funcionamiento.
- No tiene metodologías para el análisis de calidad de datos, tanto en la medición como en la base de datos.

- Su aplicación y efectividad son limitadas, al no abordarse la información en forma integral y transversal, su ámbito de aplicación resulta limitado.
- Ausencia de la participación de los estados y municipios, dichos entes se constituyen sólo como receptores de datos y no como actores de la generación y aplicación de la información para la toma de decisiones.

Para lo cual se requiere de:

- La certificación de la calidad de las mediciones y de la cobertura temporal de los datos a través de auditorías continuas que supervisen el correcto funcionamiento de las estaciones.
- Análisis en la calidad de datos, tanto en las estaciones meteorológicas, como en la base de datos.
- Mayor detalle espacial en cuanto a los pronósticos (regionalización a nivel estatal y municipal). Esto último, sería alcanzable mediante el uso de modelos de pronóstico meteorológico calibrados localmente.
- Automatización de los procedimientos de verificación de la calidad de los datos, mediante la creación e implementación de sistemas de cómputo expertos.
- Elaboración de modelos estadísticos para la corrección de errores sistemáticos.
- Uso de los resultados de los modelos para pronósticos extendidos.
- Evitar la multiplicidad de las mismas funciones entre las instituciones nacionales.
- Compra de equipos de cómputo de alto rendimiento para la ejecución de modelos de pronóstico de forma rutinaria con alta resolución.
- El desarrollo completo del pronóstico estacional y climático en México, para lo cual, es necesario generar capital humano capaz de implementar pronósticos estacionales y climáticos utilizando las herramientas técnicas y científicas más modernas.
- Adaptación de modelos de pronóstico estacional a las condiciones nacionales.
- Publicación de la información disponible y su nivel de confiabilidad.

1.2 Justificación contextual y científica

Los datos meteorológicos adquieren un alto valor socioeconómico cuando se les considera en la toma de decisiones. Éstos se emplean frecuentemente en los pronósticos del estado del tiempo y en la prevención de fenómenos como sequías e inundaciones. También son de uso frecuente en todos los países para proyectar actividades productivas a corto y mediano plazo. En pocas palabras, esta información utilizada en la toma de decisiones puede salvar vidas, mitigar desastres económicos y garantizar la seguridad alimentaria.

En México, la mayoría de la información generada en el área de meteorología es repetida entre las diversas instituciones que se tiene acceso. Sin embargo, esta información generalmente carece de un control de calidad, especialmente en datos numéricos crudos (Prieto, 2008). Por lo cual, el aseguramiento de la calidad de los datos de una red meteorológica resulta ser un aspecto fundamental que debe ser tomado en cuenta para minimizar la aparición de datos falsos y con ello reducir la incertidumbre en la toma de decisiones (Araya & Alfaro, 2009).

En estudios anteriores, realizados por Salgado (2018), se comprobó que el análisis de calidad datos meteorológicos facilita la detección de las estaciones que generan información de baja calidad. Por ello, se considera que el desarrollo de una herramienta hidro-informática basada en métodos estadísticos adaptados para el procesamiento de datos meteorológicos de estaciones automáticas y convencionales, mejorará la confiabilidad de los datos, facilitando su uso y aplicación en proyectos que en los que se requieren como insumo.

La herramienta, debe tener la capacidad de ordenar los datos meteorológicos, realizar un análisis de calidad adecuado a cada variable meteorológica y para cada estación, complementar datos faltantes (diarios para la variable de temperatura) y detectar datos erróneos, colocando banderillas de alerta de acuerdo con los resultados del análisis, así como recomendaciones a los usuarios sobre resultados finales. También, será de apoyo a las dependencias que la utilicen, para identificar estaciones que generan información errónea, con la finalidad de programar su mantenimiento. Todo ello con el objetivo de contar con una red meteorológica de mayor calidad.

En este proyecto, se toman en cuenta las estaciones que se encuentran dentro del mapex del Estado de México. El propósito de elegir este mapex como zona de estudio, es para comparar

los resultados del análisis con la información y el conocimiento de las características meteorológicas de la zona que se tienen en estudios anteriores realizados en el IITCA.

Cabe resaltar que el presente proyecto de investigación forma parte del proyecto denominado “Análisis de sequía y diseño de una herramienta hidrogeomática como soporte para la toma de decisiones” con clave UAEMEX: 4212/2016E, patrocinado por la *Fiducie de Recherche en hydologie* (Quebec-Canadá) desarrollado en el IITCA.

Los métodos estadísticos de análisis de calidad que se proponen están basados en estudios y metodologías de diferentes autores y han sido aplicados en diversas zonas del planeta. Éstos en su mayoría, verifican los umbrales físicos, realizan la revisión de la coherencia interna y espacial, efectúan pruebas de homogeneidad y completan series de datos. Los métodos han sido adaptados para cada variable meteorológica, tipo de estación meteorológica y características de la zona.

1.3 Hipótesis

El desarrollo de una herramienta hidro-informática que realice un análisis de control de calidad de los datos, así como complemento de series basados en métodos estadísticos, mejorará la confiabilidad de la información que las estaciones meteorológicas proporcionan. De esta forma, se obtendrá mayor certeza para los tomadores de decisiones de gobierno, académicos y población en general que desarrollen proyectos en materia hídrica como: hidrología urbana, agricultura, inundaciones y sequías, etc., para coadyuvar en la gestión de recursos naturales de la región de análisis.

1.4 Objetivos

1.4.1 Objetivo general

Desarrollar una herramienta hidro-informática con base en métodos estadísticos adaptados para el control de calidad, validación y complemento de series de datos meteorológicos de las estaciones automáticas y convencionales, con el propósito de mejorar la confiabilidad de la información que éstas transmiten y almacenan, para obtener mayor certeza en la toma de decisiones de proyectos hídricos y coadyuvar en la gestión de los recursos naturales de la región de análisis.

1.4.2 Objetivos específicos

1. Recopilar y estructurar las bases de datos meteorológicos de las estaciones automáticas y convencionales de las diferentes dependencias (INIFAP, SMN, CONAGUA, PEMBU, IITCA, CAEM) en un formato único para llevar a cabo el análisis de la calidad de datos.
2. Evaluar y seleccionar los métodos de análisis de calidad y complemento de series, para su aplicación en cada una de las variables meteorológicas de estaciones automáticas y convencionales.
3. Diseñar y desarrollar una herramienta hidro-informática de fácil uso que permita leer, organizar, analizar la calidad y complementar los datos meteorológicos con los métodos estadísticos seleccionados para cada variable y estación meteorológica.
4. Aplicar los métodos de análisis de la herramienta hidro-informática a las estaciones meteorológicas que se encuentran en el mapex del Estado de México, para determinar la calidad de la información que éstas proporcionan, así como comprobar la funcionalidad de la herramienta.

Capítulo 2

Gestión de calidad de los datos meteorológicos

El término calidad, ha evolucionado desde el tradicional “cumplimiento de especificaciones” hasta ser una nueva forma de “gestión” (ISO-9001, 1947). La gestión de calidad consiste en una planificación estratégica y sistemática basada en normas que se deben cumplir. Es aplicable a todo tipo de empresas y servicios, ya que su objetivo principal es brindar calidad en sus productos (COMET, 2019)..

En el caso de la meteorología, diversas organizaciones han creado normas, guías y manuales para gestionar diversas actividades. Siendo estos productos de gran ayuda para brindar a los usuarios información de calidad y una guía para su aplicación.

2.1 Normatividad para la gestión de los datos meteorológicos

La meteorología es una ciencia que observa, estudia y registra los cambios que ocurren en la atmósfera (Rodríguez et al., 2008). Estas observaciones son normadas por distintas organizaciones de cada país y éstas a su vez por la OMM. El objetivo principal de estas organizaciones, es promover y asegurar la calidad de los datos mediante reglamentos, normas y manuales que indiquen el procedimiento, los instrumentos, etc., adecuados para actividades meteorológicas y climatológicas (OMM-485, 2010). Algunas de estas organizaciones son la OMM, Asociación Española de Normalización (UNE), Organización Internacional Estándar (por sus siglas en inglés ISO), proyectos y manuales de México y Costa Rica, éstas se

muestran en la tabla 2.1, donde se describen los objetivos de las normas, manuales, guías o reglamento para el control de calidad de datos meteorológicos y climatológicos, así como la automatización de estos procesos.

Tabla 2.1 Normativas destacadas para la gestión de datos meteorológicos

Norma	Organización / País	Objetivo
Guía sobre gestión de datos de la Vigilancia Meteorológica Mundial	OMM	Proporcionar datos para las necesidades de gestión de datos meteorológicos y cómo pueden satisfacerse de forma óptima a largo plazo las necesidades de los usuarios (OMM-788, 1993)
Guía para la aplicación de un sistema de gestión de la calidad para los SM	OMM	Orientar sobre la elaboración y la aplicación de un sistema de gestión de la calidad, para garantizar la calidad de los productos y servicios que proporcionan los SMN. Asimismo, detalla los pasos necesarios para obtener el certificado de conformidad de acuerdo con la norma ISO 900 (OMM-100, 2011)
ISO 9001 “Norma de gestión de calidad”	ISO	Mencionar las actividades y pasos detallados para implementar con eficacia un sistema de gestión de calidad mediante el cumplimiento de requisitos (ISO-9001, 2008).
UNE 500510	UNE	Brinda aspectos de las redes de estaciones meteorológicas automáticas. Así como aspectos generales de nomenclatura de las variables meteorológicas (UNE-500510-IN, 2005)
UNE 500520	UNE	Describir los criterios para la localización, emplazamiento e instalación de sensores de las EMA’s (UNE-500520, 2002)
UNE 500530	UNE	Proporcionar las características de las EMA’s y los sensores e las mismas. También brinda los criterios para el mantenimiento y calibración de los sensores (UNE-176101, 2010)
UNE 500540	UNE	Proporcionar directrices para la validación en tiempo real de registros meteorológicos procedentes de EMA’s (UNE-500540, 2004)
NMX-AA-166/2-SCFI-2015	México	Especificar las técnicas que debe cumplir el emplazamiento y exposición de los instrumentos de las EMA’s y EMC’s (PROY-NMX-AA-166, 2015)
Manual de buenas prácticas para la generación de información climática.	Costa Rica	Brindar las técnicas para la generación de información climática para la producción agropecuaria, identificando las variables meteorológicas y los métodos de control de calidad para su correcta aplicación (Basualdo, 2015)

2.2 Fuentes de información meteorológica

Todos los sectores de la sociedad se ven afectados por los fenómenos meteorológicos y climáticos (SMN, 2012). Por ello, se han creado diversas organizaciones que monitorean los cambios meteorológicos y proporcionan a la sociedad esta información. En México, algunas de estas organizaciones son el SMN, INIFAP y otras independientes las cuales se describen a continuación.

2.2.1 Servicio Meteorológico Nacional (SMN)

Es el organismo encargado de proporcionar información sobre el estado del tiempo a escala nacional y local en el país. Este organismo depende de la CONAGUA, la cual forma parte de la Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) (SMN, 2012). Los objetivos del SMN se concentran en la vigilancia continua de la atmósfera para identificar los fenómenos meteorológicos que pueden afectar las distintas actividades económicas y sobre todo originar la pérdida de vidas humanas. Así como, el acopio de la información climatológica nacional (Villarreal, Reyes, Carrillo, Velázquez, & Paredes, 2014). Después del acopio de información, el SMN difunde la información al público a través de su página de <http://smn.cna.gob.mx/es/observando-el-tiempo/estaciones-meteorologicas-automaticas-ema-s> (CONAGUA, 2019). Y a su vez esta información es integrada en la Gerencia Estatal de la CONAGUA, la cual efectúa una actualización en el proyecto de computación climática (por sus siglas en inglés CLICOM). Este sistema contiene una base histórica que data desde 1890, correspondiente a EMC (SMN, 2012).

Para cumplir con las funciones y objetivos, el SMN cuenta con la siguiente infraestructura para la observación de los fenómenos atmosféricos (CONAGUA, 2019)

- Red de EMA's y ESIME's: consta de 188 EMA's y 90 ESIME's distribuidas en toda la República Mexicana, transmitiendo en tiempo real las variables meteorológicas.
- Red sinóptica de superficie: integrada por 79 observatorios meteorológicos, que transmiten en tiempo real las condiciones atmosféricas.

- Red sinóptica de altura: consta de 16 estaciones de radio sondeo, cuya función es la observación de las capas altas de la atmósfera. Cada estación realiza mediciones de presión, temperatura, humedad y viento mediante una sonda que se eleva por medio de un globo dos veces al día.
- Red de 13 radares meteorológicos distribuidos en el territorio nacional.
- Estación terrena receptora de imágenes del satélite meteorológico GOES-8. Con esta estación se reciben imágenes cada 30 minutos de cinco diferentes bandas: una visible, tres infrarrojas y una de vapor de agua. Cada imagen cubre la región meteorológica número IV, la cual abarca México, Canadá, Estados Unidos, el Caribe y Centro América.

2.2.2 Instituto Nacional de Investigación Forestal, Agrícolas y Pecuarias (INIFAP)

El INIFAP (2019) a través de la red nacional de estaciones agroclimáticas que opera en 29 entidades federativas que son: Aguascalientes, Baja California Sur, Campeche, Chiapas, Chihuahua, Coahuila, Colima, Durango, México, Guanajuato, Guerrero, Hidalgo, Jalisco, Michoacán, Morelos, Nayarit, Nuevo León, Oaxaca, Puebla, Quintana Roo, San Luis Potosí, Sonora, Sinaloa, Tabasco, Tamaulipas, Tlaxcala, Veracruz, Yucatán y Zacatecas. Cuenta con una infraestructura de 1046 estaciones distribuidas en las entidades mencionadas, ofreciendo información meteorológica en tiempo real enviada cada 15 minutos al laboratorio nacional de modelaje y sensores remotos del INIFAP, ubicado en el Campo Experimental Pabellón, Aguascalientes, en donde se procesa la información para su difusión a través de un portal de internet. La información meteorológica que se genera en esta red, es aplicada a los procesos de producción agrícola, ayudando así a la detección temprana de heladas, aplicación de riego agrícola, control de plagas, etc (Medina, Grageda, Corral, & Báez, 2008).

2.2.3 Otras dependencias que brindan información meteorológica

Existen dependencias que tiene una pequeña red de monitoreo climatológico para una zona determinada. Entre las más importantes son: CAEM, PEMBU, IITCA, Comisión del Agua del Estado de Guanajuato (CAEG), Organismo de Cuencas Lerma Santiago Pacífico (OCLSP), Organismo de Cuencas del Valle de México (OCAVM) y Protección Civil de: Chiapas, Guanajuato, Guerrero, Oaxaca, Tabasco (CONAGUA, 2019). Cada una de estas dependencias publica su información en sus sitios de internet y otras son privadas para sus estudios o aplicaciones (SMN, 2012).

2.3 Características e instrumentos de las estaciones meteorológicas

2.3.1 Estación meteorológica convencional (EMC)

Una estación meteorológica convencional es un conjunto de dispositivos mecánicos que realizan mediciones de las variables meteorológicas de forma semi manual (SMN, 2012). Colocados de forma estratégica para cumplir su función como se muestra en la figura 2.1 (OMM-100, 2011). Están conformadas por un anemómetro, pluviómetro y/o pluviógrafo, pirómetro, termómetro de máxima, termómetro de mínima, tanque evaporímetro, distribuidos como se muestra en la figura 2.1.

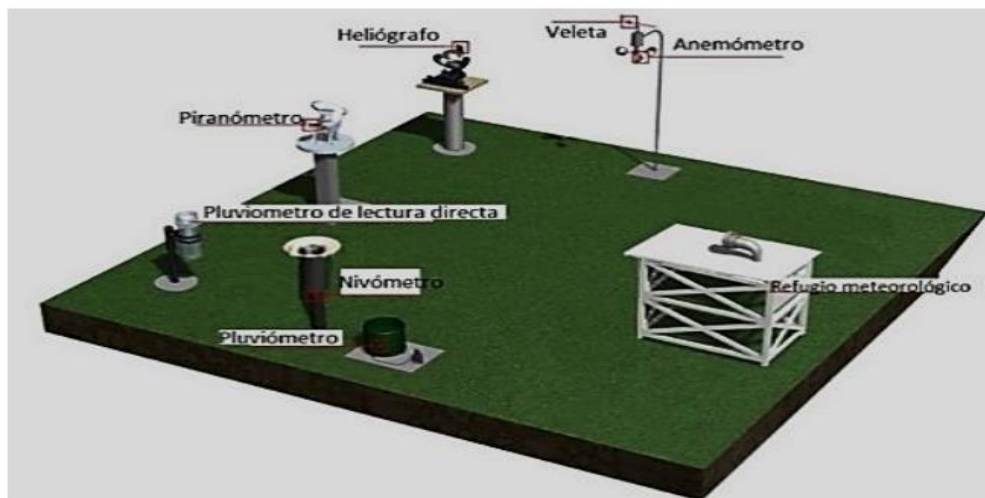


Figura 2. 1 Estación meteorológica convencional (EMC) (OMM-168, 2011)

2.3.2 Estación meteorológica automatizada (EMA)

Una estación meteorológica automática es un conjunto de dispositivos eléctricos, electrónicos y mecánicos que realizan mediciones de las variables meteorológicas de forma automática, colocadas de forma estratégica para cumplir su función (OMM-100, 2011). Están conformadas por un grupo de sensores que puede registrar hasta 20 variables meteorológicas en periodos de 10 segundos o menos. Transmitiendo esa información en tiempo real y creando archivos del promedio a cada 10 minutos, que son enviados a una estación principal que puede publicar los datos vía satélite en intervalos de una a tres horas por estación (SMN, 2012). El área representativa que cubren las estaciones es de 5 km de radio aproximadamente, en terreno plano, excepto en terreno montañoso (OMM, 2012).

La configuración de las EMA's puede con base en el propósito del sistema, pero típicamente consiste en una estructura modular que contiene una caja para intemperie, con el data logger, batería recargable, sensores meteorológicos, panel solar o generador eólico y el mástil que soporta los sensores, como se muestra en la figura 2.2 (PROY-NMX-AA-166, 2015).

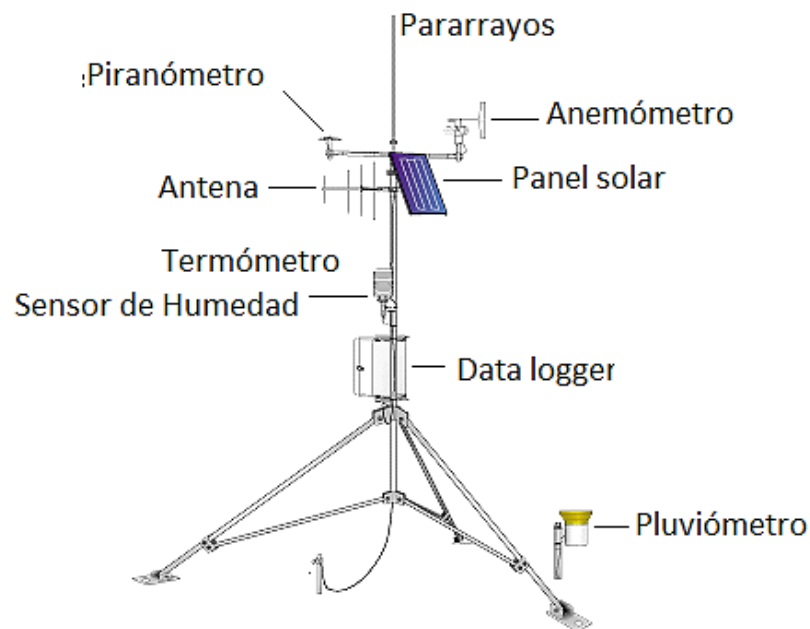


Figura 2.2 Estación meteorológica automática (CONAGUA, 2019)

2.3.3 Estación sinóptica meteorológica (ESIME)

Una estación sinóptica meteorológica es un conjunto de dispositivos eléctricos que realizan mediciones de las variables meteorológicas de manera automática (Mercado-Mancera et al., 2014). Como se muestra en la figura 2.3, las ESIME's son iguales a las EMA's de torre triangular. Las ESIME's se encuentran ubicadas exclusivamente en los observatorios meteorológicos y generan una base de datos y un mensaje sinóptico cada tres horas (CONAGUA, 2019). Los mensajes sinópticos son reportes que presentan información meteorológica de tiempo presente y pasado de manera codificada. Por lo cual, para generar este mensaje es necesario un software que permite hacer los cálculos correspondientes de las variables meteorológicas. Estos mensajes se rigen por el Tiempo Universal Coordinado (por sus siglas en inglés UTC), también llamado hora Z, el cual toma como referencia el meridiano de 0° (OMM-100, 2011).

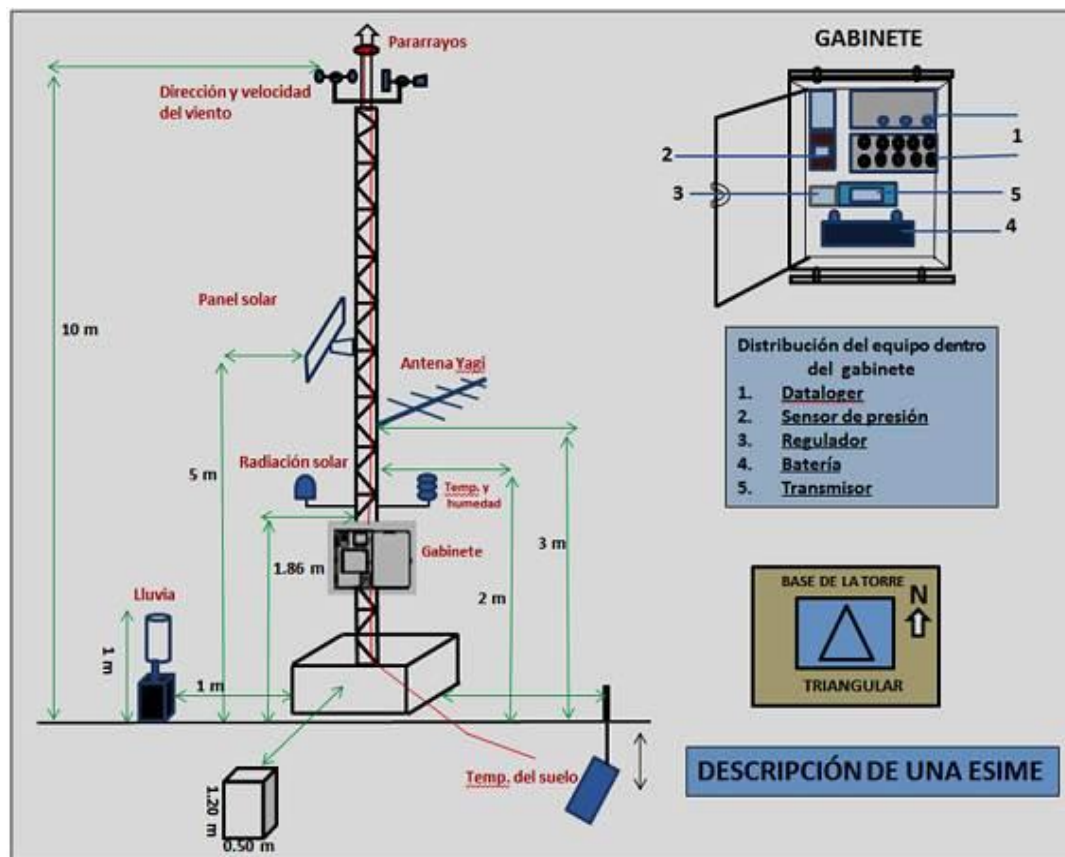


Figura 2.3 Estructura de una ESIME (CONAGUA, 2019)

2.3.4 Instrumentos de las EMC, EMA y ESIME

Una estación meteorológica normalmente dispone de varios instrumentos, el número de ellos depende del tipo de estación e información que se requiera. Por ejemplo, una estación convencional básica puede registrar temperatura (T), precipitación (P) y dirección y velocidad del viento (Dv, Vv). En la tabla 2.2, se muestran los sensores y las variables que las EMA's, ESIME's y EMC's pueden registrar, marcado con un asterisco * los instrumentos correspondientes a cada una (PROY-NMX-AA-166, 2015), (OMM-100, 2011).

Tabla 2.2 Variables de las estaciones meteorológicas

Instrumento	Unidad de medición	Variable meteorológica	Tipo de estación que lo dispone		
			EMA	ESIME	EMC
Termómetro	°C	Temperatura (T)	•	•	
Termómetro de máxima	°C	Temperatura (T)		•	•
Termómetro de mínima	°C	Temperatura (T)		•	•
Barómetro	hPa	Presión atmosférica (Pa)	•	•	
Pluviómetro	mm	Precipitación (P)	•	•	•
Pluviógrafo	mm	Precipitación (P)		•	•
Higrómetro	%	Humedad relativa (Hr)	•	•	
Tanque evaporímetro	mm de agua evaporada	Evaporación (Ev)			•
Anemómetro	m/s	Velocidad del viento (Vv)	•	•	•
Veleta	°	Dirección del viento (Dv)	•	•	•
Heliógrafo	h	Horas de Sol			•
Radiómetro	W/m ²	Radiación solar	•	•	

Todos estos datos son registrados, almacenados y están disponibles para los usuarios. En todo el mundo se proporciona constantemente gran cantidad de información meteorológica. Toda esta información se vuelve una serie de indicadores que aporta una mejor visión en la toma de decisiones (Medina et al., 2008).

2.4 Aplicación de las variables meteorológicas

Independientemente de la actividad que se empleen los datos de las estaciones meteorológicas lo que realmente interesa es extraer información, analizarla y aplicarla en distintas actividades (Serrano et al., 2016). Algunas de las principales actividades en que se emplean los datos meteorológicos, se enlistan a continuación, dando una breve descripción de su importancia.

- **Logística:** Este es uno de los sectores más sensibles a la ocurrencia de fenómenos meteorológicos atípicos, ya que todo su proceso productivo llegando hasta el consumidor consta de las rutas óptimas de los medios de transporte como son aviones, barcos, camiones, almacenes, furgonetas, etc. Sus principales variables de interés son la precipitación, olas de calor, nevadas, tornados, granizadas. Ya que de ellas dependen los posibles ahorros económicos por la selección de mejores rutas (Boshell & León, 2013; UNE-176101, 2010) .
- **Industria productiva:** observa los cambios climatológicos extremos o variables no controladas en el proceso productivo ya que éstos pueden afectar directamente a la línea de producción o a los materiales que se encuentran en ella (Llanos, 2015).
- **Agricultura y ganadería:** son claros ejemplos del uso de las estaciones meteorológicas para el aprovechamiento de la información climatológica, esto es debido a que las condiciones climáticas son fundamentales para el desarrollo de la producción adecuada sobre todo en la agricultura (Medina et al., 2008). Esta información le permitirá al usuario tomar decisiones más efectivas con sus cultivos o bien establecer patrones de sucesos para asegurar la máxima calidad de la cosecha. En la ganadería el control y gestión de las condiciones meteorológicas puede ayudarle a detectar también patrones de comportamiento y ver cómo afecta al desarrollo de su ganado (Llanos, 2015).
- **Deportes al aire libre:** se ven muy afectados en los siguientes ámbitos: Rendimiento de los jugadores en función de temperaturas u otras condiciones, distancias recorridas, recuperación, explosividad, calorías quemadas, etc. Existen apps capaces de captar información sobre el rendimiento del deportista con la información meteorológica durante los entrenamientos (Machiwal, 2009).

- **Infraestructura:** Es una de las actividades más comunes y mayormente afectada por los fenómenos extremos que se presentan (Machiwal, 2009). La precipitación, granizadas y viento, son las variables más aplicadas en esta actividad, ya que las estructuras ya sean hidráulicas y edificaciones, deben soportar esos fenómenos (Medina et al., 2008).
- **Investigación:** Es una de las actividades más usuales y con mayor impacto, ya que estudian principalmente los cambios que éstas sufren con el paso del tiempo. Mostrando las catástrofes que ocurren o pueden ocurrir en la Tierra (GEM & SMA, 2009).

2.5 Zona de estudio

Según el Programa Hidráulico Integral del Estado de México (SEMARNAT, 2002) en la entidad se cuenta con 134 estaciones meteorológicas automáticas y convencionales, de las cuales, 59 están desmanteladas. De las 75 restantes, siete EMA's y dos ESIME's son operadas por el SMN. También se tenían 66 EMA's que estaban operadas por CAEM, pero recientemente se implementó un proceso de logística inversa con objetivo de recuperar estas estaciones, desconociéndose los alcances del proceso. De las estaciones del Estado de México, el 64% de ellas se encuentran en buenas condiciones y lo único que requieren es un mantenimiento del sistema. El 15% requiere reparaciones menores, el 13% requiere un mantenimiento mayor y el 9% restante es necesario sustituir el equipo.

De los equipos meteorológicos disponibles en el Estado de México, en el presente proyecto se analizan las bases de datos de las estaciones administradas por CONAGUA, SMN, INIFAP, PEMBU y CAEM que se encuentren en la región del Estado de México y sus estados circunvecinos. A esta región se le denominará en lo consecutivo: “mapex del Estado de México”, que es la zona de estudio. Las coordenadas extremas de esta región de análisis serán: -98.325833° a -100.72361° de longitud Oeste y 18.2663889° a 20.3336111° de latitud norte como se muestra en la figura 2.4. También se muestra la distribución de las estaciones de cada dependencia contempladas en el proyecto.

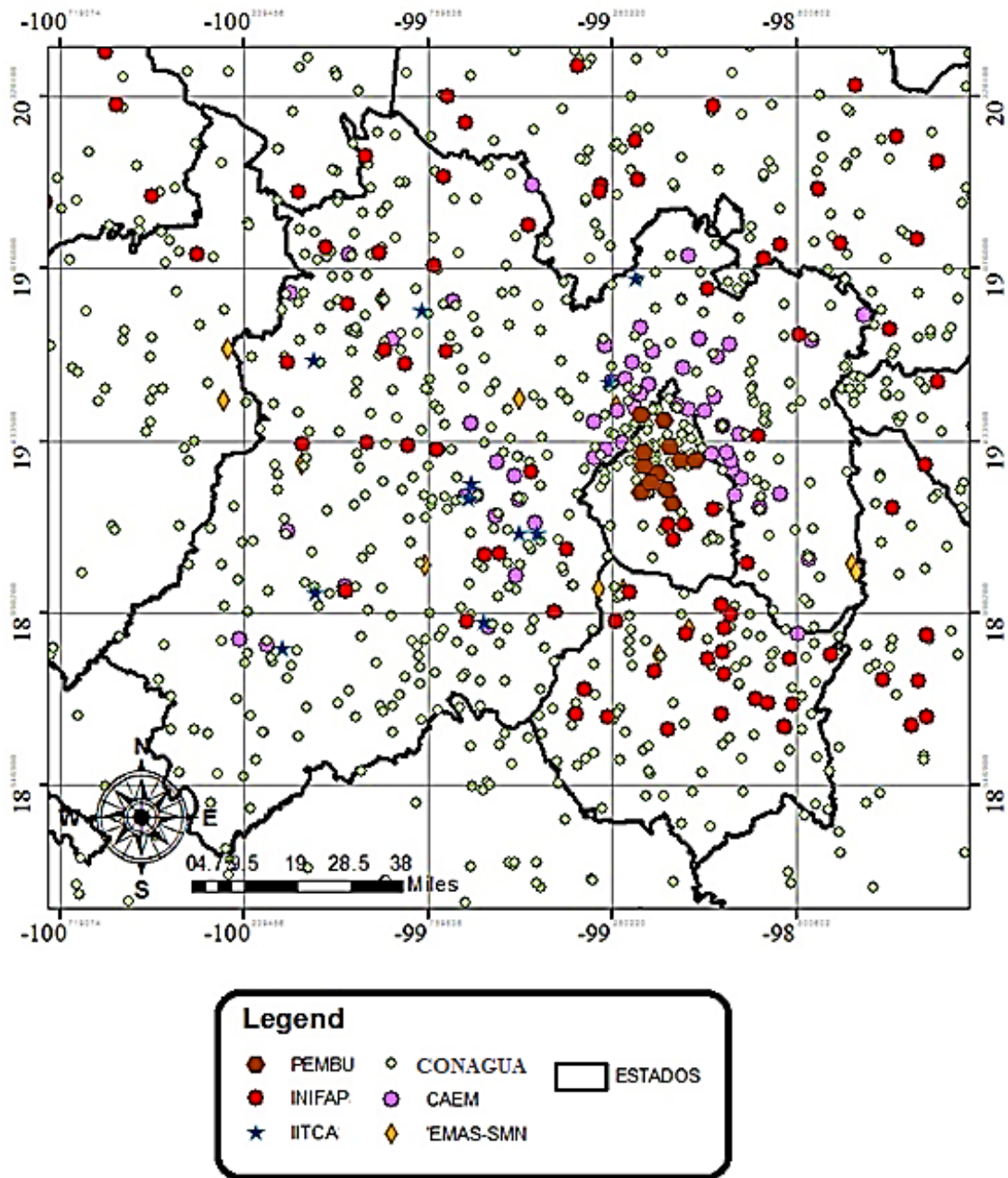


Figura 2.4 Estaciones meteorológicas del mapex del Estado de México

En la tabla 2.3 muestra el número de estaciones que conforman el mapex del Estado de México. De ahí, se derivan las correspondientes a cada uno de los estados y las dependencias que las administran. También se muestra el número total de estaciones y porcentaje correspondiente a EMA's y EMC's. Se puede apreciar que se tiene mayor número de EMC's

(84% del total de estaciones). Se observa igualmente, que se considerarán todas las estaciones ubicadas en Morelos, Estado de México y Ciudad de México, así como sólo algunas estaciones de los estados de Guerrero, Guanajuato, Hidalgo, Michoacán, Puebla, Querétaro y Tlaxcala.

Tabla 2.3 Número de estaciones meteorológicas correspondientes al mapex del Estado de México

Estado	EMC	EMA					Total
	CONAGUA	SMN	IITCA	INIFAP	PEMBU	CAEM	
Guanajuato	12	0	0	3	0	0	15
Guerrero	37	0	0	0	0	0	37
Hidalgo	87	0	0	9	0	0	96
Michoacán	38	1	0	1	0	0	40
Morelos	82	1	0	15	0	0	98
Puebla	32	0	0	3	0	0	35
Querétaro	6	0	0	1	0	0	7
Tlaxcala	24	0	0	1	0	0	25
Ciudad de México	60	2	0	2	13	0	77
Estado de México	345	4	3	20	0	58	430
Suma	723	8	3	55	13	58	860
Porcentaje	84%	16%					100%

Del total de estaciones mencionadas, para el presente proyecto sólo se consideraron 729 descartando las estaciones con periodos de registro cortos y/o discontinuadas. El total de estaciones analizadas se muestran en la tabla 2.4.

Se puede apreciar en la tabla 2.4 el número correspondiente de estaciones por cada dependencia que la administra. También se observan el periodo de registro aproximado y las variables meteorológicas consideradas con los intervalos de registro correspondientes. Por ejemplo, una EMA del SMN tiene 144 registros diarios con intervalos a cada 10 minutos, registrando datos de 8 variables meteorológicas y analizando un periodo del 2000 al 2018.

Tabla 2.4 Estaciones meteorológicas consideradas

Dependencia	Tipo de estación	Nº de estaciones	Intervalos de registro	Registros por día	Variables meteorológicas	Periodo de análisis promedio	
CONAGUA	EMC	643	24 h	1	Tmáx, Tmín, Ev, P	1920	2018
SMN	EMA	8	10 min	144	Tamb, Hr, P, Dv, Vv, Pb, Rs	2000	2018
IITCA	EMA	1	10 min	144	Tamb, Hr, P, Dv, Vv, Pb	2008	2017
INIFAP	EMA	19	15 y 30 min	96-48	Tamb, Hr, P, Dv, Vv, Pb	2000	2015
PEMBU	EMA	13	15 min	96	Tamb, Hr, P, Dv, Vv, Pb, Rs	1996	2018
CAEM	EMA	45	30 min	48	Tamb, Hr, P, Dv, Vv, Pb	2000	2015
Total de estaciones:		729					

Capítulo 3

Marco teórico

El control de calidad de datos meteorológicos es dividido en varias fases en función de la profundidad de la verificación y de los datos disponibles. Por ello, objetivo principal del presente trabajo es diseñar una herramienta hidro-informática con métodos robustos y por fases para el análisis, control de calidad, y validación de los datos meteorológicos. Estos datos se obtienen de las EMA's, ESIME's y EMC's, de las dependencias de SMN, INIFAP, CONAGUA y CAEM, que se encuentren ubicadas en el mapex del Estado de México.

Con base en la revisión de bibliografía, los sistemas de control de calidad de datos meteorológicos en diferentes países constan de diversas fases (Abaurrea et al., 2004; COMET, 2019; Cruz-Saco et al., 2010) (Prieto, 2008). Estas fases, dependen de diferentes métodos y estrategias para su correcto funcionamiento, las cuales se describen a continuación.

3.1 Fases del análisis de calidad de los datos meteorológicos

Como se ha mencionado, el control de calidad se considera como un proceso de comprobación y validación de los datos, pero no la garantía de calidad en términos generales. Y en estaciones automáticas se tiene que tener un mayor control y cuidado en el análisis debido a la cantidad de datos que éstas generan (OMM-100, 2011). Este control de calidad adopta métodos para el análisis de las variables meteorológicas. Las principales etapas del

control de calidad según las normas y manuales encontrados,(ISO-9001, 1947),(Llanos, 2015) (OMM-100, 2011) consisten en:

1. Calibración de instrumentos y obtención los datos en la estación
2. Análisis visual y detección de datos erróneos
3. Organización de la base de datos
4. Análisis de calidad de datos con métodos estadísticos.
5. Categorización del análisis
6. Identificación de la calidad del dato.

Donde las etapas uno, dos y tres son llevadas a cabo por personal calificado de manera semi manual. Mientras que en las etapas cuatro, cinco y seis se realizan con ayuda de softwares especializados. En la etapa cuatro se encuentran los métodos estadísticos y numéricos para el análisis de calidad de datos. Y en las etapas cinco y seis se realiza una valoración del análisis de calidad de los datos y con ello se obtiene una calificación de los datos.

La etapa cuatro se puede encontrar una subdivisión de acuerdo con el análisis que se desea hacer. Con base en la bibliografía encontrada (Alvarez, Montaña, Quentin, Maldonado, & Solano, 2013; Llanos, 2015; Martinez & Amoreo, 2015; OMM-100, 2011; OMM-168, 2011; Ortiz, Escobar, & Sepúlveda, 2018), los métodos de análisis se subdividen en:

- a) Revisión de umbrales físicos
- b) Revisión de coherencia interna
- c) Revisión de coherencia espacial
- d) Revisión de homogeneidad de las series
- e) Complementar datos faltantes de las series

En cada una de estas subdivisiones, se tienen diversos métodos de análisis, que diferentes países y dependencias generadores de información meteorológica utilizan para el control de calidad en sus bases de datos. Entre los métodos recopilados, los que se aplican en el presente proyecto se describen a continuación.

3.2 Métodos de umbrales físicos

Estos métodos consisten en verificar que el dato meteorológico se encuentre en un rango predeterminado y factible. Este rango puede ser la capacidad de medición de los sensores (Estévez G., 2008), datos climatológicos de la zona y se utilizan métodos estadísticos como el boxplot y la desviación estándar. Estos métodos se describen en apartados ulteriores.

3.2.1 Umbrales del sensor

Este método se establece con base en las normas o manuales de fábrica de las estaciones, donde se indica la capacidad de medición de los sensores. Si uno de los datos no se encuentra dentro de estos rangos, éste es considerado automáticamente erróneo, ya que no es posible que el sensor pueda registrarlo (Flemming, et al., 2002). Los rangos de las principales variables meteorológicas que se encontraron se muestran en la tabla 3.1. Donde se tiene las variables meteorológicas: dirección del viento (Dv), velocidad del viento (Vv), humedad relativa (Hr), temperatura (T), precipitación (P), radiación solar (Rs), presión atmosférica (Pa) y los respectivos rangos que deben de cumplir según algunos autores (NMX-AA-166/1-SCFI, 2013; Llanos, 2015; Medina, Grageda, Corral, & Báez, 2008; OMM-100, 2011; OMM-306, 2018).

Tabla 3.1 Rangos de medición del sensor

Variable	Rango del sensor	
	Superior	Inferior
Dv	0°	360°
Vv	0 m/s	75 m/s
Hr	0 %	100 %
T	-40 °C	60 °C
P	0 mm	10 mm (en 10 min aprox.)
Rs	0 W/m ²	3000 W/m ²
Pa	500 hPa	1030 hPa

3.2.2 Desviación estándar

La desviación estándar (s) indica el grado de dispersión de las observaciones con respecto al centro de gravedad de la información (la media (\bar{x})). Mientras mayor sea la desviación estándar, mayor será la dispersión de los datos (Sanchez, 2008). Para estimar la variación general de los datos meteorológicos se emplea el método basado en la desviación estándar como se muestra en la ecuación 3.1 (OMM-305, 2000) (Hubbard et al., 2006).

$$\bar{x} - (3s) \leq x_i \leq \bar{x} + (3s) \quad (3.1)$$

Donde \bar{x} es la media y s la desviación estándar de la muestra. El rango varía de ± 3 desviaciones estándar, el cual engloba aproximadamente un 99,7% todos los posibles datos (Walpole & Mye, 2007).

3.2.3 Boxplot

El Boxplot, o gráfico de caja, describe visualmente las características importantes de una serie de datos (Walpole, Myers, Myers, & Ye, 2012). Este gráfico de caja representa los tres cuartiles, valores mínimos y máximos de una serie de datos. Los cuartiles son los tres valores que dividen al conjunto de datos ordenados en cuatro partes iguales, éstos se calculan como se muestra en las ecuaciones siguientes

$$\text{Primer cuartil } Q_1 \quad Q_1 = x_{(1+(n-1)(0.25))} \quad (3.2)$$

$$\text{Segundo cuartil } Q_2 \text{ es la mediana} \quad Q_2 = x_{(1+(n-1)(0.50))} \quad (3.3)$$

$$\text{Tercer cuartil } Q_3 \quad Q_3 = x_{(1+(n-1)(0.75))} \quad (3.4)$$

Bigote superior: corresponde a la observación $A = Q_3 + 1.5(Q_3 - Q_1)$ (3.5) inmediatamente inferior al valor calculado por la expresión A cuando éstas se ordenan de menor a mayor.

Bigote inferior: corresponde a la observación $B = Q_1 - 1.5(Q_3 - Q_1)$ (3.6) inmediatamente superior al valor calculado por la expresión B cuando éstas se ordenan de menor a mayor.

Donde n es el número total de los datos ordenados de menor a mayor y X es el valor que se encuentra en la posición dada por el subíndice. Este valor delimita el lugar del cuartil. En la figura 3.1, se muestra la ubicación de los cuartiles y bigotes.

Los rangos de este método están referenciados por lo bigotes y si algún dato se encuentra fuera de los bigotes, esta observación se considera atípica y deberá ser cotejada.

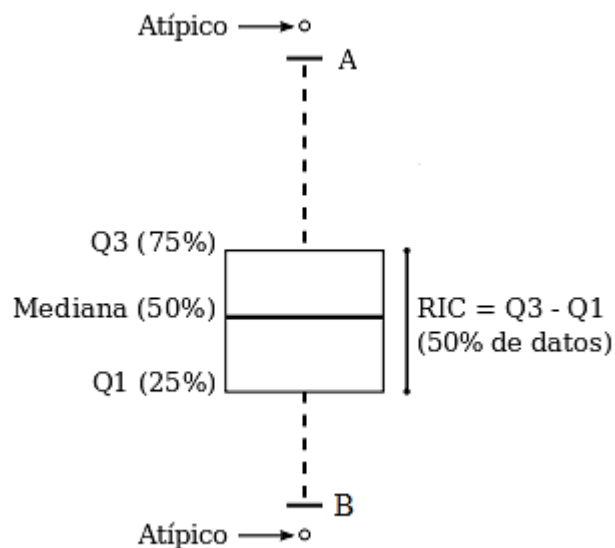


Figura 3.1 Gráfico boxplot

3.3 Métodos de coherencia

La coherencia, es la relación lógica entre dos o más elementos, de tal modo, que no se producen cambios bruscos entre ellos. En meteorología estos métodos, de análisis de calidad de datos, verifican la concordancia entre los registros meteorológicos, ya sea interna o espacialmente, como se describe a continuación.

3.3.1 Coherencia interna

Este método consiste en comparar el cambio entre observaciones consecutivas. Si la diferencia excede un valor o umbral factible, este dato es considerado como “atípico” o “sospechoso”. Para la revisión de coherencia interna, se han encontrado dos métodos: desviación estándar y umbrales físicos.

3.3.1.1 Umbrales físicos

En este método de análisis de datos meteorológicos, toma en cuenta las relaciones meteorológicas que existen entre las observaciones efectuadas en la misma estación (Flemming et al., 2007). Es decir, los valores medidos al mismo tiempo y en el mismo lugar no pueden ser inconsistentes entre ellos. Según Estévez (2008), las comprobaciones que se realizan en esta validación son las siguientes:

- El valor de V_v no puede ser igual a cero m/s.
- El valor de Dirección del Viento (D_v) no puede ser igual a cero grados.
- El valor de Humedad relativa (H_r) no puede ser inferior a un 50% cuando el valor del dato de Precipitación es superior a 0,2 mm.
- El valor mínimo de Humedad relativa es del 20% sin presencia de precipitación.
- El valor de Temperatura Mínima (T_{min}) no puede ser igual o superior al de Temperatura Máxima (T_{max}) ni al de Temperatura (T). Del mismo modo, el valor de T no puede ser igual o superior al de T_{max} (Para valores diarios)

- El valor de Radiación solar (R_s) no puede ser igual a cero W/m^2 .
- La diferencia de dos registros consecutivos de temperatura no puede ser mayor a $10^\circ C$ (aplicable en EMA's o EMC's)
- La diferencia de dos registros consecutivos de precipitación no puede ser mayor a 10 mm. (aplicable en EMA's)

3.3.1.2 Desviación estándar temporal

Este nivel de validación evalúa la coherencia del valor de un dato con respecto a los valores de los datos precedentes y los datos posteriores. Para realizar la validación retoma el método de la desviación estándar como se muestra en la expresión 3.7, en donde se halla la media de los 5 registros precedentes y de los 5 registros posteriores (Balairón, 2014). Del mismo modo, aparte de la media, también se calcula la desviación estándar de estos datos. Si el valor de un dato rebasa cualquiera de las medias calculadas ± 3 veces la desviación estándar de los datos implicados en esas medias, se considerará un dato sospechoso y, por tanto, inválido en cuanto a su consistencia temporal (UNE-500540, 2004).

$$\bar{x}_{\pm 5} - 3s_{\pm 5} \geq x_i \leq \bar{x}_{\pm 5} + 3s_{\pm 5} \quad (3.7)$$

Donde $s_{\pm 5}$ es la desviación estándar de los 5 datos posteriores y precedentes del valor a analizar, x_i es el elemento i de la muestra y $\bar{x}_{\pm 5}$ es la media de los 5 datos posteriores y anteriores del valor a analizar.

3.3.2 Coherencia espacial

Este tipo de método sirve para detectar posibles errores comparando los datos de una estación con los valores espaciales de otras estaciones, tomando en cuenta la información de instrumentos como radar y satélite, así como la presencia de fenómenos naturales en dicha zona (Flemming et al., 2007). Para el presente proyecto sólo se consideran los métodos de comparación con otras estaciones y comparación con registros de huracanes.

3.3.2.1 Comparación con otras estaciones

Este método es uno de los más simples para la revisión de coherencia espacial. Consiste en cotejar los datos de una estación con respecto a las estaciones vecinas correlacionadas (Flemming et al., 2007). Se toman en cuenta sólo las estaciones situadas a una distancia inferior a 10 Km (Shulskii et al., 2014). A manera de ejemplo, se muestra en la figura 3.2 el círculo en color azul tiene un radio de 10 km teniendo como centro la estación meteorológica ubicada en Rectoría de la UAEMex (ubicada en color rojo con un cero al centro). Como se observa, dentro de este radio se encuentran dos estaciones, una ubicada en Calixtlahuaca (marcada en color amarillo y el número 1) y otra en Metepec (marcada en color amarillo y el número 2).

Cabe destacar que para el presente proyecto se tomó la distancia de 20km como radio máximo, debido a la falta de estaciones que puedan tenerse en la región.

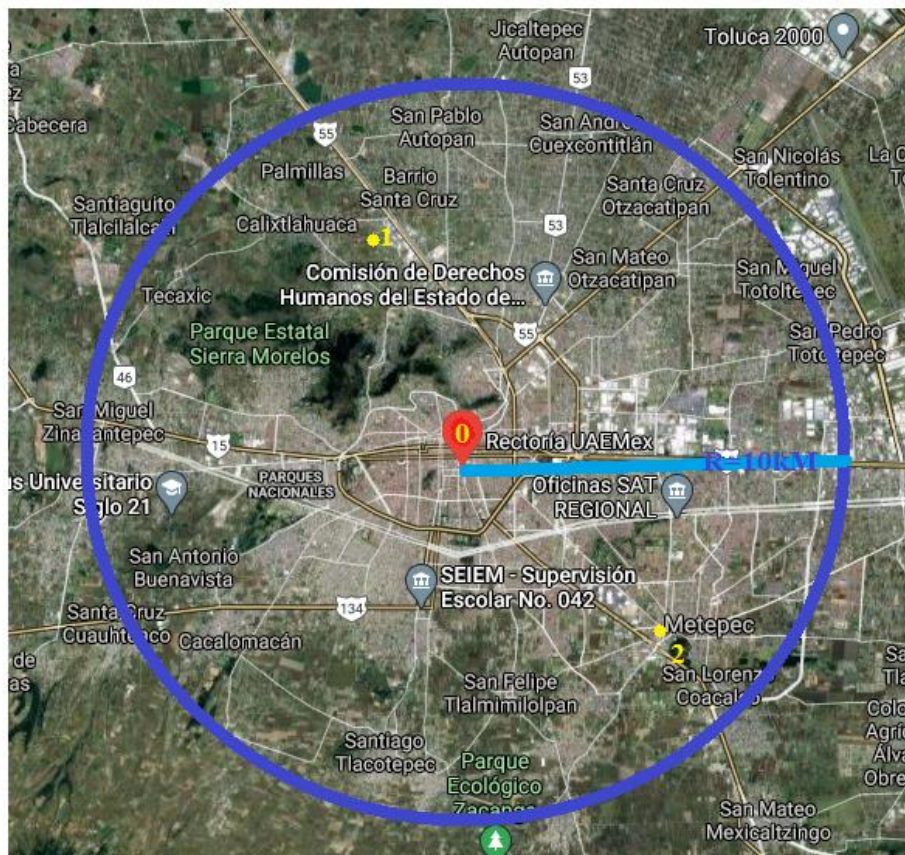


Figura 3.2 Método de comparación con estaciones cercanas.

Este método consiste en verificar y relacionar los datos de las estaciones cercanas que se elijan. Por ejemplo, para el caso de la precipitación, si el fenómeno se presenta en varias estaciones, este tipo de lluvia es consecuencia de un proceso pseudo-adiabático por elevación orográfica, o bien por frente frío, es decir, por el choque entre dos masas de aire con diferente humedad y temperatura. Si se presenta algún fenómeno como huracanes, es lluvia tipo ciclónica. Y si sólo se presenta en una estación puede ser lluvia conectiva.

3.3.2.2 Comparación con datos de huracanes

Este método consiste en comparar y validar los registros de las variables meteorológicas, con la presencia simultánea de un fenómeno ciclónico. Es una comparación cualitativa la cual sólo indica la ocurrencia concomitante de un huracán en determinada zona. Las estaciones que se encuentren dentro de zona pueden presentar registros de P, Vv y Hr elevadas, la Dv se puede inclinar a una dirección, y la Rs, T y Ev bajan. Por lo cual, para llevar a cabo este método es necesario contar con una base de datos de huracanes con fechas y lugares en los que se manifestó el fenómeno.

Por ejemplo, en la figura 3.3 se muestra en negro la trayectoria de un huracán, en ella se delimita su área de influencia o afectación. Esta área, fecha y huracán quedan registrados en la base de datos. Mientras que los puntos rojos, muestran la ubicación de las estaciones meteorológicas.

Los datos de las estaciones meteorológicas que se encuentren dentro del área de afectación del huracán se comparan con las fechas en que éste se presentó. Si coincide que la estación registró datos inusuales en las variables meteorológicas con la ubicación y fecha igual al del fenómeno, estos registros se consideran como válidos.

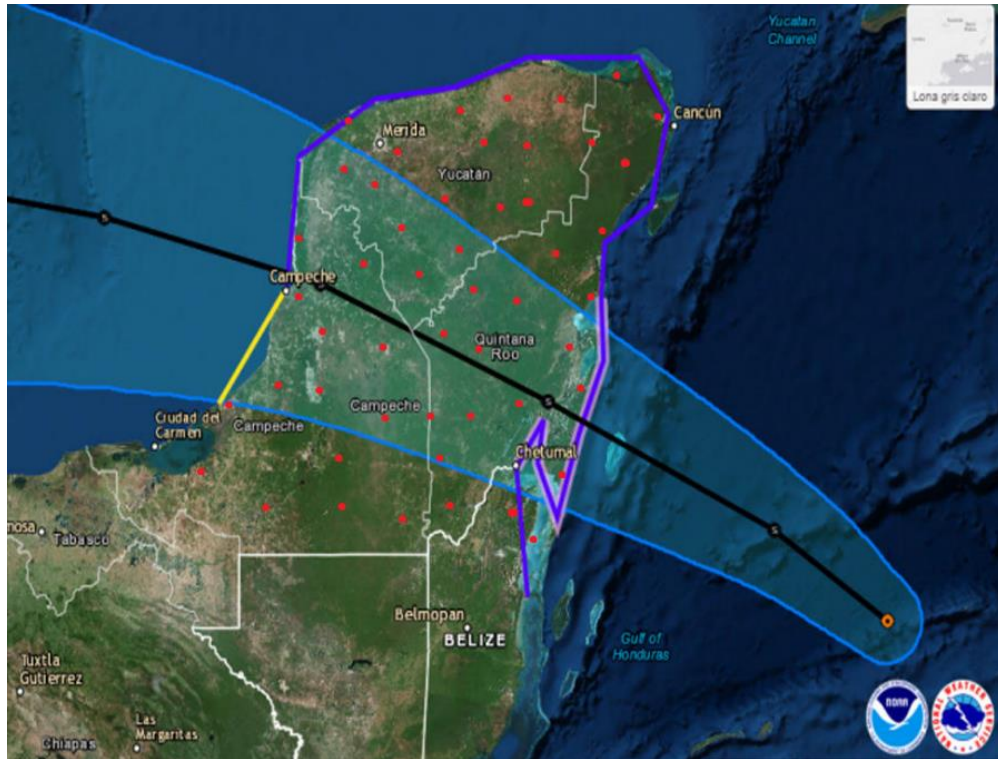


Figura 3.3 Trayectoria de huracanes con estaciones meteorológicas (NOAA, 2019)

3.4 Métodos de homogeneidad

La homogeneidad de las series de datos meteorológicos, es un análisis básico para la detección de la variabilidad climática (Štěpánek, Zahradníček, & Farda, 2013). Estos métodos de análisis ponen en evidencia los cambios bruscos en los datos, mismos que se mantienen posteriormente. Si esto ocurre, se dice entonces que estas series no son homogéneas y puede ser necesario tratar estos fenómenos como dos poblaciones estadísticas diferentes. Igualmente, para este tipo de análisis, diversos autores también utilizan funciones de distribución de probabilidad.

3.4.1 Distribución Normal

La distribución de probabilidad conocida como distribución Normal o campana de Gauss, es la más importante de las distribuciones estadísticas por la cantidad de fenómenos que explica (Walpole, et al., 2012). Una de sus características más importantes es que casi cualquier distribución de probabilidad, tanto discreta como continua, se puede aproximar por una normal bajo ciertas condiciones (Bidegain, Alvaro, & Diaz, 2011)

La distribución Normal está completamente determinada por dos parámetros que son: la media (\bar{x}) que nos indica la posición de la campana y la desviación estándar (s) que determina el grado de achatamiento de la curva como se muestra en la figura 3.4 (Walpole, et al., 2012). Y su función de densidad está dada por la siguiente ecuación:

$$N(x; \bar{x}, s) = \frac{e^{-\frac{(x-\bar{x})^2}{2s^2}}}{\sqrt{2\pi}s} \quad -\infty < x < \infty \quad (3.8)$$

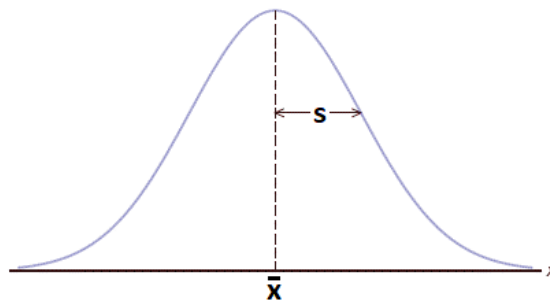


Figura 3.4 Campana de Gauss

Para determinar la probabilidad o área bajo la curva, es necesario conocer Z la cual está dada por la ecuación 3.9

$$Z = \frac{x - \bar{x}}{s} \quad (3.9)$$

Posteriormente, se determina el valor de α el cual corresponde al porcentaje de incertidumbre que se desea tener. Con base en este valor se calcula la probabilidad del valor buscado y está dada por:

$$P(-\alpha/2 < Z < \alpha/2) = 1 - \alpha \quad (3.10)$$

Si un valor se encuentra fuera de esa probabilidad se considera atípico o dudoso.

El valor de α indica un porcentaje de error permitido en la muestra. A partir de ello, se establecen los intervalos de confianza. Estos intervalos, establecen un rango con cierto grado de seguridad, en el cual se encuentra determinado porcentaje de la muestra. Dentro de este intervalo, se debe encontrar el valor a analizar. Estos intervalos dependen de la desviación estándar como describen a continuación y gráficamente en la figura 3.5.

- Se encuentran aproximadamente, el 68,26% de los datos de la muestra

$$\bar{x} - s \quad a \quad \bar{x} + s \quad (3.11)$$

- Se encuentran aproximadamente, el 95,44% de los datos de la muestra

$$+\bar{x} - 2s \quad a \quad \bar{x} + 2s \quad (3.12)$$

- Se encuentran aproximadamente, el 99,74% de los datos de la muestra

$$\bar{x} - 3s \quad a \quad \bar{x} + 3s \quad (3.13)$$

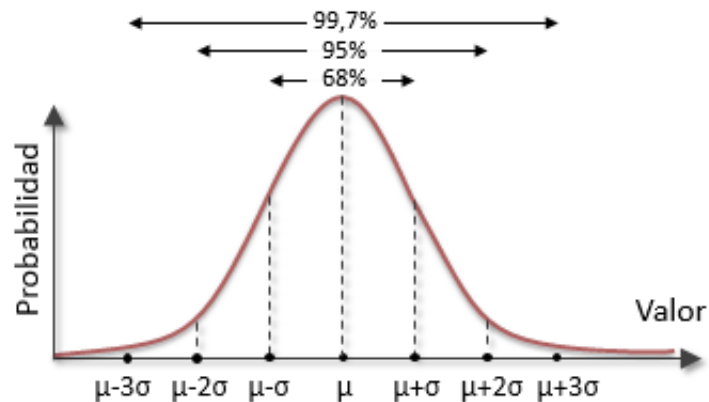


Figura 3.5 Intervalos de confianza de la distribución Normal

3.4.2 Distribución Log-Normal

Esta distribución, consiste en una transformación logarítmica a la variable de interés y luego utilizar el modelo de distribución Normal para la variable transformada. La distribución así obtenida se denomina log-Normal, donde su función de densidad de probabilidad es $N(\mu_y; \sigma_y)$ y está dada por la siguiente ecuación

$$f(y) = \frac{1}{x\sigma_y\sqrt{2\pi}} e^{-\left(\frac{1}{2\sigma_y^2}\right)(y-\mu_y)^2} \quad (3.14)$$

Donde $y=\ln(x)$, μ_y es el promedio de y y σ_y es la desviación estándar de la muestra

$$\mu_x = e^{\left(\mu_y + \frac{\sigma_y^2}{2}\right)} = \bar{y} \quad (3.15)$$

$$\sigma_y^2 = (\mu_y)^2 (e^{\sigma_y^2} - 1) = s_y^2 \quad (3.16)$$

Donde $(e^{\sigma_y^2} - 1)$ corresponde al coeficiente de variación C_v .

3.4.3 Distribución Gamma

Muchas de las variables meteorológicas solo pueden asumir valores positivos. Es por ello por lo que muchas de las variables tienen un comportamiento como la distribución Gamma, ya que tiene como límite inferior cero. La función de densidad de probabilidad de la distribución Gamma es de la siguiente forma:

$$f(x) = \frac{\lambda}{\Gamma k} \lambda x^{k-1} e^{-\lambda x} \quad \text{para } x \geq 0 \quad (3.17)$$

Donde $\Gamma(k)$ es la función Gama, λ es el parámetro de escala, k es el parámetro de forma y son los parámetros de la función.

$$\Gamma(k) = \int_0^{\infty} t^{k-1} e^{-t} dt = (k - 1)! \quad (3.18)$$

$$\lambda = \frac{x}{s^2} \quad k = \frac{x^2}{s^2} \quad (3.19)$$

Para el cálculo de los parámetros de la función, es necesario calcular s^2 que es desviación estándar de la muestra y μ que es la media.

$$\mu = \frac{k}{\lambda} \quad s^2 = \frac{k}{\lambda^2} \quad (3.20)$$

3.4.4 Distribución Log-Pearson III

Esta distribución dice, que si los logaritmos de una variable aleatoria X se ajustan a una distribución Pearson III, se dice que la variable aleatoria X se ajusta a una distribución Log-Pearson III. Esta distribución es usada en el análisis de frecuencia de caudales máximos. Se trabaja igual que una distribución Pearson tipo III pero con \bar{x}_y y s_y como la media y desviación estándar de los logaritmos de la variable original X (Valleumbroso, 2014). Siendo su función de densidad como se muestra en la ecuación 3.21.

$$f(x) = \frac{1}{x |\alpha| \Gamma\beta} \left(\frac{\ln(x) - y_0}{\alpha} \right)^{\beta-1} \exp\left(- \frac{\ln(x) - y_0}{\alpha} \right) \quad (3.21)$$

Donde a y b son parámetros de escala, $\ln(x)$ y forma, respectivamente. Y y_0 es el parámetro de localización. Sus parámetros se calculan como:

$$\beta = \left(\frac{2}{Cs} \right) \quad (3.22)$$

$$\alpha = s_y \frac{Cs}{2} \quad (3.23)$$

$$x_0 = \bar{x}_y - \alpha\beta \quad (3.24)$$

Donde Cs es el coeficiente de asimetría, \bar{x}_y y s_y son la media y la desviación estándar de los logaritmos de la muestra respectiva

3.4.6 Distribución T-Student

La distribución t ó T-Student es una distribución de probabilidad que surge del problema de estimar la media de una población o muestra cuando ésta es pequeña (Montgomery & Runger, 1994). Esta distribución determina la diferencia entre dos medias muestrales y un intervalo de confianza. La definición la distribución T-Student es:

Sea Z una variable aleatoria normal estándar y V una variable aleatoria chi cuadrada con v grados de libertad. Si Z y V son independientes, entonces la distribución de la variable aleatoria T está dada por $T = \frac{Z}{\sqrt{V/v}}$

Esta distribución de probabilidad tiene la siguiente función de densidad:

$$h(t) = \frac{\Gamma\left[\frac{(v+1)}{2}\right]}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi v}} \left(1 + \frac{t^2}{v}\right)^{-(v+1)/2} \quad -\infty < t < \infty \quad (3.25)$$

A esta función se le conoce como la distribución T con v grados de libertad. Con X_1, X_2, \dots, X_n como las variables aleatorias independientes y media μ y desviación s^2

$$\mu = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (3.26)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (3.27)$$

Por lo tanto, el intervalo de confianza basado en la distribución T-Student, consiste en estimar \bar{x} y s de la muestra. Con ello se el intervalo de confianza de $(1 - \alpha)$ y con $\nu = n - 1$ grados de libertad, se busca el área bajo la curva de la función de probabilidad. En donde la probabilidad de T se debe encontrar dentro de estos dos valores como se muestra en la ecuación siguiente

$$P(-t_{\alpha/2} < T < t_{\alpha/2}) = 1 - \alpha \quad (3.28)$$

3.5 Métodos para complementar series

Es importante tener series de datos meteorológicos lo más completas posible, para ello, se han adaptado métodos estadísticos de complemento de series que se basan en la información que aportan los datos registrados en estaciones vecinas y a partir de ellos estimar los valores faltantes en la estación, o región, de interés.

3.5.1 Método de Kriging

El método de Kriging es un método geoestadístico de estimación de la variable de interés en la región de estudio el cual utiliza un modelo de semivariograma para la obtención de los ponderadores que se darán a cada punto de referencia usados en la estimación. La ventaja del kriging sobre otros métodos de interpolación es la estimación de la varianza y del error de predicción. Sin embargo, no proporciona información acerca de la relación que pueda existir entre dichos puntos (Štěpánek et al., 2013). El método consiste en que se tienen las mediciones $Z(s_1), Z(s_2), Z(s_3)$ y $Z(s_4)$, en los puntos s_1, s_2, s_3 y s_4 respectivamente, y se requiere predecir el valor $Z(s_0)$ como se muestra en la figura 3.6. El valor por estimar se ubica más cerca de s_2 que de cualquier otra ubicación donde se tenga medición; por lo tanto,

es lógico pensar que $Z(s_0)$ es más parecido a $Z(s_2)$ que a cualquiera de los otros tres valores medidos.

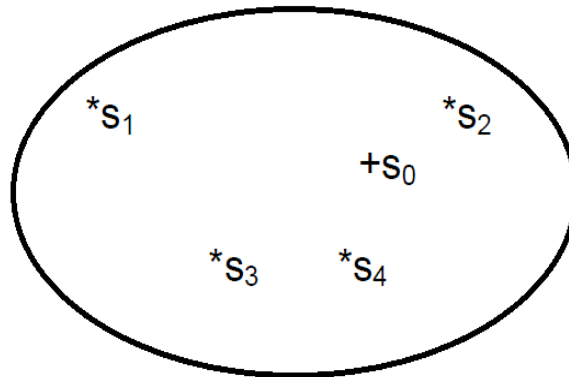


Figura 3.6 Ejemplo de un muestreo para la interpolación con Kriging

Así, la función para la aproximación del valor está dada por

$$\hat{Z}(s_0) = \lambda_1 Z(s_1) + \lambda_2 Z(s_2) + \lambda_3 Z(s_3) + \lambda_4 Z(s_4) \quad (3.29)$$

Donde s_1, s_2, s_3, s_4 son los valores de la muestra y $\lambda_i = 1, 2, 3, 4$ son los factores de ponderación o pesos tales que $\sum_{i=1}^n \lambda_i = 1$.

En general, para obtener una estimación de $Z(s_0)$, se realiza una combinación lineal de los valores con los factores de ponderación

$$\hat{Z}(s_0) = \sum_{i=1}^n \lambda_i Z(s_i) \quad (3.30)$$

De acuerdo con el método, para obtener un estimador insesgado, la esperanza matemática E y el semivariograma $\gamma(h)$ están dadas por:

Esperanza matemática

$$E[Z(s+h) - Z(s)] = 0 \quad (3.31)$$

Semivariograma

$$\gamma(h) = \frac{1}{2} \text{var}[Z(s+h) - Z(s)] \quad (3.32)$$

Donde $\gamma(h)$ es el semivariograma, el cual se utiliza para la estimación de los puntos faltantes. Y h es la variable independiente de la muestra. Estas funciones deben dar cumplimiento a las siguientes condiciones del método.

1. Condición de insesgamiento (3.33)

$$E(\hat{Z}) = Z$$

2. Condición de mínima varianza (3.34)

$$\text{var}[\hat{Z}(s_0) - Z(s_0)] = E[\hat{Z}(s_0) - Z(s_0)]^2$$

Donde $\hat{Z}(s_0)$ es el valor ponderado, aproximado con Kriging, E es la esperanza matemática o media y var es la varianza de los valores.

3.5.1.1 Kriging con factor de cambio externo.

Este método es una derivación del kriging ordinario, el cual asume que existen condiciones externas que están físicamente correlacionadas con la variable a estimar. Para el caso de las variables meteorológicas, éstas se ven modificadas por la altitud principalmente. Por ejemplo, para interpolar temperaturas en lugares donde la altitud no cambia bruscamente kriging ordinario es adecuado, ya que la temperatura sólo se ve afectada por el tiempo y el cambio de estaciones y no cambia de una estación a otra. Pero en zonas con variación topográfica la temperatura difiere de una estación a otra, por lo cual kriging ya no es adecuado (Arévalo. M, 2019).

Kriging con cambio externo divide una función aleatoria en una combinación lineal de funciones deterministas, una tendencia de variación no estacionaria $\mu(x)$ llamada factor de cambio y un componente aleatorio $Y(x_i) = Z(x_i) - \mu(x)$ que representa la función aleatoria residual del sistema. De esta forma se introduce una tendencia en la esperanza matemática

con un valor estimado, similar al que se obtiene con kriging ordinario. Este modelo se expresa en la ecuación 3.41, muy similar a la expresión de kriging ordinario, su diferencia radica en expandir el sistema k veces la función $f(x)$ y se ajusta al variograma con los residuales de la regresión multivariable.

$$\left[\left(\begin{array}{ccc|ccc} \tilde{z}_{1,1} & \dots & \tilde{z}_{1,n} & 1 & f_0(x_1) & \dots & f_0(x_1) \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{z}_{n,1} & \dots & \tilde{z}_{nn} & 1 & f_k(x_1) & \dots & f_k(x_n) \\ \hline 1 & \dots & 1 & 0 & 0 \dots & 0 & 0 \\ f_0(x_1) & \dots & f_0(x_n) & 0 & 0 \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ f_k(x_1) & \dots & f_k(x_n) & 0 & 0 \dots & 0 & 0 \end{array} \right) * \begin{bmatrix} w \\ w_1 \\ \vdots \\ l \\ \mu_0 \\ \vdots \\ \mu_{f_k} \end{bmatrix} = \begin{bmatrix} \tilde{z}_{00} \\ \vdots \\ \tilde{z}_{n0} \\ 1 \\ f_0(x_0) \\ \vdots \\ f_k(x_0) \end{bmatrix} \quad (3.35)$$

Donde \tilde{Z} es el valor ponderado, aproximado con kriging, μ es la esperanza matemática.

Además de la esperanza matemática y la varianza, Kriging con cambio externo proporciona la desviación estándar para medir el error de cálculo de la interpolación. Existen otros métodos para validar los resultados como: el método del vecino cercano, el vecino móvil o una validación cruzada (Pérez-Planells, 2015). En el presente proyecto, se validará Kriging con cambio externo mediante el método de validación cruzada de vecindario único propuesto por Dubrule (1983)

3.5.2 Validación cruzada de vecindario único

En los métodos de regresión geoestadística la validación cruzada es muy útil ya que genera el grado de calidad de la inferencia estadística. Para el caso del método de kriging, los mapas proporcionan buenos resultados en sus aproximaciones (Dubrule, 1983). En un vecindario único, los resultados de kriging se interpretan como puntos utilizados para la validación cruzada. Entonces la matriz de kriging se interpreta como:

$$\begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1n} & f_1^0 & \dots & f_1^m \\ \gamma_{21} & & & \gamma_{2n} & f_2^0 & \dots & f_2^m \\ \dots & \gamma_{n1} & \gamma_{n2} & \gamma_{nn} & f_n^0 & \dots & f_n^m \\ f_1^0 & \dots & f_1^m \\ f_2^0 & \dots & f_2^m \\ \dots \\ f_n^0 & \dots & f_n^m \end{bmatrix} = \begin{bmatrix} b^1 \\ b^2 \\ \dots \\ b^n \\ c_0 \\ \dots \\ c_m \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ \dots \\ z_n \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (3.36)$$

Donde f es el monomio de grado inferior a m conocido por kriging, γ es el multiplicador de lagrange, y b es un coeficiente que se deduce como:

$$\left. \begin{cases} \sum_{\alpha=1}^n b^\alpha \gamma(x_\alpha - x_\beta) + \sum_{l=0}^m c_l f^l(x_\beta) = z(x_\beta) & (\forall \beta \in \{1, \dots, n\}) \\ \sum_{\alpha=1}^n b^\alpha f^l(x_\alpha) = 0 & (\forall l \in (0, \dots, m)) \end{cases} \right\} \quad (3.37)$$

Para realizar la validación cruzada, se toma el sistema de matrices (3.41) y se invierte, deduciendo el siguiente sistema de ecuaciones.

$$\begin{bmatrix} \gamma_{11} & \gamma_{1\alpha_0-1} & \gamma_{1\alpha_0+1} & \dots & \gamma_{1n} & f_1^0 & \dots & f_1^m \\ \gamma_{\alpha_0-1,1} & \dots & \dots & & \gamma_{\alpha_0-1,n} & f_{\alpha_0-1}^0 & \dots & f_{\alpha_0-1}^m \\ \gamma_{\alpha_0+1,1} & \dots & \dots & & \gamma_{\alpha_0+1,n} & f_{\alpha_0+1}^0 & \dots & f_{\alpha_0+1}^m \\ \gamma_{n1} \dots & \gamma_{n,\alpha_0-1} & \gamma_{n,\alpha_0+1} & & \gamma_{nn} & & & \\ f_1^0 \dots & f_{\alpha_0-1}^0 & f_{\alpha_0+1}^0 & \dots & f_n^0 & & & \\ \dots f_1^m & f_{\alpha_0-1}^m & f_{\alpha_0+1}^m & \dots & f_n^m & & & \end{bmatrix} = \begin{bmatrix} b'^1 \\ b'^{\alpha_0-1} \\ b'^{\alpha_0+1} \\ \dots \\ b'^n \\ c'_0 \\ \dots \\ c'_m \end{bmatrix} = \begin{bmatrix} z_1 \\ z_{\alpha_0-1} \\ z_{\alpha_0+1} \dots \\ z_n \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad (3.38)$$

Esta matriz inversa se utiliza para seleccionar el variograma del sistema de la matriz 3.41. A partir de ello, se define la ecuación resumida 3.44, la cual es empleada para la solución del sistema. Obteniendo con ella, los datos desconocidos y pueden ser comparados con los resultados de kriging original.

$$z^*(s_{\alpha_0}) = - \sum_{\alpha \neq \alpha_0} \frac{b^{\alpha\alpha_0}}{b^{\alpha_0\alpha_0}} z(s_{\alpha}) \quad (3.39)$$

Donde s_{α_0} son los puntos del sistema de kriging.

3.6 Métodos de aplicación para cada variable meteorológica y tipo de estación

Los métodos matemáticos descritos anteriormente, se emplean para el análisis de calidad de datos de las variables meteorológicas del presente proyecto. Algunos de ellos se deben ajustar a las características de las variables meteorológicas y en algunos casos el método no es apto para algunas variables. En la tabla 3.2, los métodos, las variables meteorológicas y su temporalidad de registro. En él, se marca con una “•” los métodos que resultan adecuados para cada situación.

Tabla 3.2 Métodos de aplicación para cada variable meteorológicas

Método	Temperatura máxima [T _{máx}]	Temperatura mínima [T _{mín}]	Evaporación [Ev]	Precipitación [P]		Temperatura ambiente [T _{amb}]	Dirección del viento [D _v]	Velocidad del viento [V _v]	Radiación solar [Rs]	Presión atmosférica [P _h]
	Diario	Diario	Diario	Diario	Semi-horario	Semi-horario	Semi-horario	Semi-horario	Semi-horario	Semi-horario
Umbrales físicos	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
Coherencia interna	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
Coherencia espacial	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
Homogeneidad	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•	•	•
Resumen estadísticos	•	•	•	•	•	•	•	•	•	•
Curvas IDIT				•						
Complemento de datos	•	•								
Kriging										

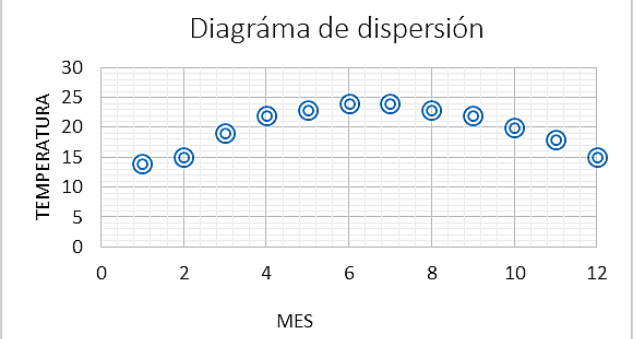
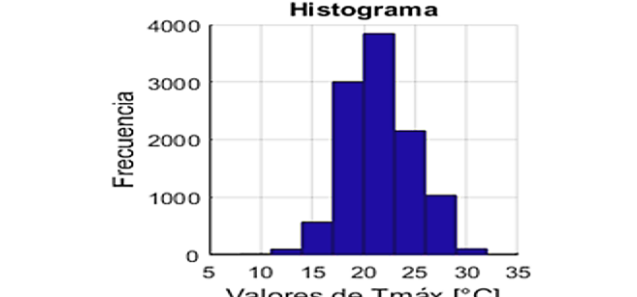
3.7 Información relevante de los datos meteorológicos

En los análisis estadísticos de bases de datos, es frecuente utilizar gráficas y resúmenes estadísticos para dar una visión amplia del comportamiento de los datos. Esta información transmite los resultados de forma rápida, directa y comprensible (Montgomery & Runger, 1996) (Lind, 2015)

3.7.1 Gráficas

Es una representación visual que describe la relación entre variables. Estos gráficos se utilizan de acuerdo con la naturaleza de los datos y la información que se desea representar (Walpole, et al., 2012). Los gráficos más comunes para representar de forma clara datos meteorológicos se muestran en la tabla 3.3, donde se tiene el tipo de gráfico con una breve descripción y una figura que lo representa.

Tabla 3.3 Gráficas más usuales para representar datos meteorológicos

Gráfico	Ejemplo
<p>Diagramas de dispersión: muestran los valores de sobre un sistema de ejes cartesianos, cuando la variable es discreta o cualitativa.</p>	<p style="text-align: center;">Diagrama de dispersión</p> 
<p>Histogramas: formas especiales de diagramas de barras para distribuciones cuantitativas continuas</p>	<p style="text-align: center;">Histograma</p> 

3.7.2 Resumen estadístico

Es la forma numérica de representar de forma resumida toda la información de la base de datos y con ello identificar su comportamiento. Está dado por medidas descriptivas, obtenidas a partir de la muestra, las cuales se describen a continuación (Walpole et al., 2012)

- Media (\bar{x}): es el promedio aritmético de las observaciones.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (3.40)$$

Donde n es el número total de datos de la muestra y x_i es el valor i de la muestra

- Mediana (Me): es el valor que separa por la mitad las observaciones ordenadas de menor a mayor. De tal forma que el 50% de estas son menores que la mediana y el otro 50% son mayores.
- Moda (Mo): es el valor de la variable que más veces se repite, es decir, aquella cuya frecuencia absoluta es mayor. No tiene por qué ser única.
- Varianza (s^2): medida de dispersión que indica la variabilidad de los datos respecto a la media

$$s^2 = \frac{\sum_{i=1}^n (x_n - \bar{x})^2}{n - 1} \quad (3.41)$$

Donde n es el número total de datos de la muestra y \bar{x} es la media de la muestra

- Desviación estándar (s): medida de dispersión que se define como la raíz cuadrada positiva de la varianza

$$s = \sqrt{s^2} \quad (3.42)$$

Donde s^2 es la varianza de la muestra

- Coeficiente de variación de Pearson (Cv): representa el grado de dispersión de los datos, se define como el cociente entre la desviación típica y el valor absoluto de la media aritmética

$$Cv = \frac{s}{\bar{x}} \quad (3.43)$$

Donde s y \bar{x} es la varianza y media de la muestra

- Máximo (x_{max}): Valor mayor de la muestra
- Mínimo (x_{min}): Valor menor de la muestra
- Rango muestra (Re): Es la diferencia entre el valor de las observaciones mayor y el menor.

$$Re = x_{max} - x_{min} \quad (3.44)$$

Donde x_{max} y x_{min} es el valor máximo y mínimo de la muestra

3.8 Matlab

Matrix Laboratory, laboratorio matricial (MATLAB) es una herramienta informática que surgió para realizar cálculos matemáticos, especialmente operaciones con matrices. Ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio llamado lenguaje M. Entre sus prestaciones básicas se hallan la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware (Matlab, 2019). Permite crear gráficos de diversos tipos y presenta grandes ventajas al trabajar con matrices de gran tamaño. Sin embargo, no es la mejor herramienta para usar en una simple tarea de programación.

En este proyecto, se eligió MATLAB como software para programar la herramienta hidro-informática, ya que éste tiene la facilidad de trabajar con matrices de gran tamaño. En este proyecto, son las bases de datos meteorológicos. También contiene funciones ya predeterminadas las cuales facilitan la programación de las metodologías de análisis, así como la generación de gráficos que ayudan a comprender el comportamiento de los datos. Para el caso de las interfaces de la herramienta, MATLAB cuenta con la función GUI que es un desarrollador de interfaz gráfica fácil de programar y amigable para el usuario.

Capítulo 4

Materiales y métodos

Para el desarrollo de la herramienta hidro-informática es necesario conocer las características de las bases de datos que generan las EMA's, ESIME's y EMC's. Así como el comportamiento de las variables meteorológicas. A partir de ello, se genera una base de datos uniforme y de fácil uso la cual es analizada por los métodos numéricos descritos en el capítulo 3.

4.1 Base de datos meteorológicos

Desde el siglo XX empezó a ser habitual el intercambio de datos meteorológicos en forma digital. Esto permitió que los Servicios Meteorológicos Nacionales construyeran bases de datos, generando gran avance en la recopilación y procesamiento de datos meteorológicos. Con base en lo anterior, la OMM creó directrices y normas internacionales sobre gestión de datos climáticos e intercambio de datos (OMM-100, 2011). En la actualidad, la gestión de registros climáticos responde a una metodología sistemática que abarca registros en papel, registros en microfilm o en microficha y registros digitales, que a su vez pueden incorporar tanto imágenes como representaciones alfanuméricas tradicionales (OMM-306, 2018).

La estandarización para la estructuración de una base de datos consiste en almacenar de manera redundante y acceder eficazmente a los datos. Esto con el propósito de evitar incoherencias, pérdida y redundancia en los datos (PROY-NMX-AA-166, 2015). En México las diferentes dependencias que observan, almacenan y transmiten datos meteorológicos

utilizan distintos formatos de registro. Dependiendo del tipo de estación y dependencia cambia el número de variables que se registran, los periodos de registro y su formato.

En el presente proyecto se organizaron las bases de datos meteorológicos del mapex del Estado de México, donde se ubican las dependencias del INIFAP, CAEM, PEMBU, IITCA, SMN y CLICOM. Las características de las bases de datos de cada una de ellas se describen a continuación.

4.1.1 Bases de datos de estaciones meteorológicas convencionales (EMC's)

Como ya se mencionó en el capítulo 2, las EMC's generan datos diarios, los cuales obtenemos de CLICOM, que es un software de manejo de datos climatológicos desarrollado por las Naciones Unidas y administrado en México por el SMN. CLICOM presenta los datos recopilados durante las últimas 24 horas, finalizando a las 08:00 am. Cada una de las diferentes estaciones del país contiene diferentes periodos de información, pero se pueden encontrar datos desde 1920 a la actualidad.

El formato oficial de este software es como el que se muestra en la figura 4.1, el cual es un archivo con formato *.csv, que contiene datos como el identificador (ID) del estado y de la estación, la clave de la variable meteorológica, el año y mes del registro y posteriormente los datos diarios registrados en el mes. Para este ejemplo la clave del estado es 15 que corresponde al Estado de México (las demás claves de los estados se muestran en el anexo 4.1, "ID de los Estados de la República Mexicana"). Las claves de las variables meteorológicas básicas se describen en la tabla 4.1 con las unidades que estas presentan.

DATASET-ID	Station-ID	ELEMENT-CC	YEAR-MONT	VALUE-1	VALUE-2	VALUE-3	VALUE-4
15	15001	1	1957-05	15	13.5	13.5	13
15	15001	2	1957-05	26.5	27.5	24	27.5
15	15001	3	1957-05	7.5	7.5	6.5	7.5
15	15001	5	1957-05	0	7	0	0.5
15	15001	18	1957-05	7.1	9.6	8.8	7.7
15	15001	30	1957-05	0	0	0	0
15	15001	31	1957-05	0	0	0	0
15	15001	32	1957-05	0	0	0	0
15	15001	43	1957-05	1	1	0	1
15	15001	1	1957-06	16.5	13.5	14	14
15	15001	2	1957-06	21.5	22	24.5	22
15	15001	3	1957-06	8.5	7.5	8	9.5
15	15001	5	1957-06	11.5	3	17.5	17

Figura 4.1 Formato de registro CLICOM

Tabla 4.1 ID de las variables en CLICOM

Variable	Abreviación	ID de la variable	Unidades
Temperatura ambiente	Ta	1	°C (a las 8:00 am)
Temperatura máxima	Tmáx	2	°C
Temperatura mínima	Tmín	3	°C
Precipitación en 24 hrs	P	5	mm (de 8:00 am a 8:00 am)
Evaporación en 24 hrs	Ev	18 o 17	mm (de 8:00 am a 8:00 am)
Días con tormenta	Dt	30	0 no hay, 1 si hay
Días con graniza	Dg	31	0 no hay, 1 si hay
Días con niebla	Dn	32	0 no hay, 1 si hay
Cobertura nubosa	Cn	43	0 despejado, 1 medio nublado, 2 nublado
Días con heladas	Dh	91	0 no hay, 1 si hay

En el presente proyecto sólo se tomarán en cuenta las variables de Temperatura máxima y mínima (Tmáx y Tmín), Precipitación (P) y Evaporación (Ev)

4.1.2 Bases de datos de estaciones meteorológicas automáticas (EMA's)

Las estaciones automáticas generan distintos formatos de registro, esto se debe al tipo de estación y el software que contenga. Es por ello que es más difícil tener un formato único (UNE-176101, 2010). Para la base de datos de EMA's y ESIME's que se trabajó en este proyecto, de acuerdo con la zona de estudio, se encuentran las dependencias del SMN, INIFAP, CAEM, PEMBU e IITCA. Para esta base de datos, se encontraron distintos formatos de registro, los cuales se describen brevemente a continuación.

4.1.2.1 Bases de datos del SMN

La red de EMA's y ESIME's del SMN, cuenta con observaciones en algunas estaciones desde el año 2000 hasta hoy en día. Éstas son registradas con intervalos de 10 minutos para distintas variables. El orden en que éste registra es distinto y depende de la estación meteorológica. Por ejemplo, en las figuras 4.2, 4.3 y 4.4, se muestran los tres formatos encontrados en la base de datos del SMN para la zona de estudio.

Las imágenes que se presentan son de la estación meteorológica “ALAMOS, en el Estado México”. Se puede observar que en los formatos cambia la forma en registrar la fecha y hora, el nombre y la información básica de la estación.

Date	Time	Dir	WSMDir	WSK	WSMK	AvgTemp	AvgRh	AvgBP	Rain	AvgSR	Batt
		deg	deg	kph	kph	C	%	mbar	mm	W/m^2	V
2006 Abr 01	00:00	138	152	14.7	24.1	29	22	965.1	0	258	
2006 Abr 01	00:10	140	148	15.4	23.5	28.3	24	965.3	0	111	
2006 Abr 01	00:20	150	128	13.6	18.3	28.1	25	965.4	0	104	
2006 Abr 01	00:30	150	127	14.2	20.6	27.9	28	965.4	0	140	
2006 Abr 01	00:40	146	147	16.4	23.5	28	28	965.3	0	159	
2006 Abr 01	00:50	156	155	12.9	18.9	27.8	28	965.6	0	83	
2006 Abr 01	01:00	141	143	12.9	20.6	27.6	28	965.7	0	57	
2006 Abr 01	01:10	139	143	14.2	20.6	27.3	25	965.7	0	32	
2006 Abr 01	01:20	117	118	10.9	16	26.9	24	965.8	0	15	
2006 Abr 01	01:30	102	113	11.6	16	26.4	23	965.9	0	4	
2006 Abr 01	01:40	90	89	13	16.6	26	23	966.2	0	0	
2006 Abr 01	01:50	81	68	13.1	18.3	25.8	24	966.4	0	-2	

Figura 4.2 Formato tipo 1 estación meteorológica “ALAMOS, en el Estado México”

	DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativa	PresBarometric	Precipitacion	RadSolar	nombre_estacion
01/04/2011 00:00	307	306	20.5	27.5	31	17	959.5	0	262	ALAMOS, SON. (EHCA)
01/04/2011 00:10	311	302	18	25.2	30.8	17	959.5	0	227	ALAMOS, SON. (EHCA)
01/04/2011 00:20	319	326	16.1	24.1	31.1	17	959.6	0	251	ALAMOS, SON. (EHCA)
01/04/2011 00:30	316	320	19.4	24.1	30.9	18	959.6	0	211	ALAMOS, SON. (EHCA)
01/04/2011 00:40	321	331	16	24.6	30.8	19	959.7	0	170	ALAMOS, SON. (EHCA)
01/04/2011 00:50	322	314	18.6	24.1	30.5	21	959.7	0	131	ALAMOS, SON. (EHCA)
01/04/2011 01:00	323	327	23	31	29.8	22	959.8	0	94	ALAMOS, SON. (EHCA)
01/04/2011 01:10	315	313	25.1	31.6	29.1	24	959.9	0	58	ALAMOS, SON. (EHCA)
01/04/2011 01:20	312	306	22.2	31	28.5	24	960	0	13	ALAMOS, SON. (EHCA)
01/04/2011 01:30	320	315	21.2	28.1	28	24	960.2	0	5	ALAMOS, SON. (EHCA)
01/04/2011 01:40	318	312	22	28.7	27.7	24	960.3	0	0	ALAMOS, SON. (EHCA)
01/04/2011 01:50	324	315	19.5	25.2	27.4	24	960.4	0	0	ALAMOS, SON. (EHCA)
01/04/2011 02:00	334	338	18.7	24.1	27.2	23	960.5	0	0	ALAMOS, SON. (EHCA)
01/04/2011 02:10	333	334	16	20.6	26.9	23	960.5	0	0	ALAMOS, SON. (EHCA)
01/04/2011 02:20	343	349	15.5	21.8	26.7	22	960.6	0	0	ALAMOS, SON. (EHCA)
01/04/2011 02:30	347	354	17.1	22.3	26.5	22	960.8	0	0	ALAMOS, SON. (EHCA)

Figura 4.3 Formato tipo 2 estación meteorológica “ALAMOS, en el Estado México”

nombre_estacion	fecha	DirVient	DirRafaga	RapVient	RapRafaga	TempAii	HumRel	PresBaro	Precipit	RadSolar
ALAMOS, SON. (EHCA)	01/04/2017 00:00	187	198	23.8	35	27.4	15	959.7	0	314
ALAMOS, SON. (EHCA)	01/04/2017 00:10	180	204	23.1	35	27.1	14	959.8	0	276
ALAMOS, SON. (EHCA)	01/04/2017 00:20	172	151	22.1	31.6	27.1	12	959.9	0	236
ALAMOS, SON. (EHCA)	01/04/2017 00:30	169	151	20.4	33.3	26.8	12	959.9	0	194
ALAMOS, SON. (EHCA)	01/04/2017 00:40	189	194	20.6	31.6	26.7	13	960	0	154
ALAMOS, SON. (EHCA)	01/04/2017 00:50	187	198	23.1	37.9	26.2	14	960	0	113
ALAMOS, SON. (EHCA)	01/04/2017 01:00	180	162	18.2	27.5	25.8	15	960.1	0	76
ALAMOS, SON. (EHCA)	01/04/2017 01:10	168	167	20.6	33.3	25.5	15	960.2	0	45
ALAMOS, SON. (EHCA)	01/04/2017 01:20	173	171	15.3	26.9	25.1	16	960.4	0	14
ALAMOS, SON. (EHCA)	01/04/2017 01:30	172	175	16.5	25.2	24.7	18	960.4	0	5
ALAMOS, SON. (EHCA)	01/04/2017 01:40	172	183	17.2	28.1	24.4	20	960.4	0	0
ALAMOS, SON. (EHCA)	01/04/2017 01:50	171	167	16.1	23.5	24.1	22	960.4	0	0
ALAMOS, SON. (EHCA)	01/04/2017 02:00	166	162	18.4	29.8	23.8	27	960.4	0	0

Figura 4.4 Formato tipo 3 estación meteorológica “ALAMOS, en el Estado México”

El SMN, también cuenta con ESIME’s, las cuales tienen sus formatos uniformes, pero con más información meteorológica como se muestra en la figura 4.5. En ella se observa el formato de la estación “CEMCAS” en el Estado de México (la parte que dice continuación se alarga en el formato hacia la derecha).

ID	Fecha-Tiempo	SR(W/m^2)	Rain(mm)					DP	RH %	GTempC(C)		
			24hrs	6Hrs	3Hrs	1Hr	10Min			12Hrs	TS(C)	
76111	01/01/2013 00:00	4	0	0	0	0	0	6.8	52	1.9	17.2	
76111	01/01/2013 00:10	-1	0	0	0	0	0	6.7	52	1.9	17.3	
76111	01/01/2013 00:20	-1	0	0	0	0	0	6.6	52	1.9	17.3	
76111	01/01/2013 00:30	-1	0	0	0	0	0	6.3	51	2	17.4	
76111	01/01/2013 00:40	-1	0	0	0	0	0	7	54	2.1	17.4	
76111	01/01/2013 00:50	-1	0	0	0	0	0	7	55	2.1	17.5	
76111	01/01/2013 01:00	-1	0	0	0	0	0	7.2	57	2.4	17.5	
76111	01/01/2013 01:10	-1	0	0	0	0	0	7.3	58	2.8	17.5	
76111	01/01/2013 01:20	-1	0	0	0	0	0	7.5	59	3.4	17.6	
76111	01/01/2013 01:30	-1	0	0	0	0	0	8.1	63	3.8	17.6	
76111	01/01/2013 01:40	-1	0	0	0	0	0	7	59	4.2	17.6	
Continuación	ATCmax(C)		ATCmin(C)		ATC(C)		WSmax(m/ WSmin(m/s)		WDmax(m/ WDmin(m/		QFF	BP(mbar)
							s)		s)			
	23.4	2.7	16.7	9.6	3.2	6.9	178	143	159	1017.4	781.8	
	23.4	2.7	16.6	9.7	4.3	7.1	177	146	163	1017.6	781.8	
	23.4	2.7	16.4	9.2	3.5	6.8	177	154	165	1017.9	781.9	
	23.4	2.7	16.4	9.7	4.3	7.3	180	152	165	1018	782	
	23.4	2.8	16.2	8.8	3.7	6.5	178	134	164	1018.3	782	
	23.4	2.9	15.9	7.9	3.6	5.6	181	154	167	1018.5	782.1	
	23.4	2.9	15.7	6.5	3.5	5.1	181	152	168	1018.8	782.1	
	23.4	3.4	15.5	6	3.2	4.8	180	157	171	1019	782.1	
	23.4	3.8	15.4	5.1	2.4	3.9	189	160	173	1019.2	782.2	
	23.4	4.3	15.1	4.7	3.1	3.9	184	162	172	1019.6	782.3	
23.4	4.5	14.9	5.3	2.8	4	176	158	167	1019.9	782.4		

Figura 4.5 Formato para las ESIME's del SMN estación "CEMCAS"

4.1.2.2 Bases de datos del INIFAP

La red INIFAP, cuenta con observaciones de estaciones agroclimáticas, con periodos de registro desde el año 2000 hasta hoy en día para algunas estaciones. Éstas, son registradas con intervalos de 15 y 30 minutos para distintas variables. Los formatos que maneja la dependencia son dos. El primero se muestra en la figura 4.6 el cual muestra los datos que envía la dependencia al realizar la solicitud en línea, éste sólo contiene las variables que se solicitan con intervalos de registro a cada 15 minutos, el nombre y clave de la estación.

El segundo formato se muestra en la figura 4.7, y se obtiene por medio de la solicitud en línea al Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos (INAI). En este formato se tiene el registro de todas las variables meteorológicas de la estación con periodos de registro a cada 30 minutos.

ESTACIÓN SAN JOSE DEL RIO					
CLAVE 31					
ESTADO DE MÉXICO					
INIFAP					
Fecha	Prec	HumR	Temt	DirV	VelV
20/12/2013 23:45	NULL	NULL	NULL	NULL	NULL
21/12/2013 00:00	0	83	10.7	322.5	6.6
21/12/2013 00:15	0	83	10.7	279.6	5.7
21/12/2013 00:30	0	83	10.8	241.7	7.7
21/12/2013 00:45	0	84	10.2	138.9	5
21/12/2013 01:00	0	87	9.2	124.1	3.8
21/12/2013 01:15	0	87	8.9	57.7	4
21/12/2013 01:30	0	91	8.2	121.2	5.2
21/12/2013 01:45	0	91	7.9	3.2	4.5
21/12/2013 02:00	0	93	7.4	322.6	3.7
21/12/2013 02:15	0	93	7.6	46.8	5.8
21/12/2013 02:30	0	93	6.8	316.9	3.1
21/12/2013 02:45	0	95	6.8	325.4	3.9
21/12/2013 03:00	0	93	7.7	226.5	5.6
21/12/2013 03:15	0	84	10.2	230.9	9.8
21/12/2013 03:30	0	82	10.6	143.1	7.6
21/12/2013 03:45	0	86	9.3	28.1	4.3

Figura 4.6 Formato 1 de la base de datos del INIFAP

Date	Time	Temp	Hi	Low	Out	Dew	Wind	Wind	Wind	Hi	Hi	Wind	Heat
		Out	Temp	Temp	Hum	Pt.	Speed	Dir	Run	Speed	Dir	Chill	Index
01/01/2009	00:30	13.2	14.9	13.2	57	4.9	0.4	NW	0.8	2.2	NNW	13.2	12.4
01/01/2009	01:00	12.3	13.2	12.3	62	5.3	0.4	NW	0.8	2.7	N	12.3	11.7
01/01/2009	01:30	11.6	12.3	11.6	66	5.5	1.3	NW	2.41	3.6	W	11.6	11.2
01/01/2009	02:00	10.6	11.6	10.6	69	5.2	1.3	NW	2.41	3.6	WNW	10.5	10.3
01/01/2009	02:30	10	10.6	10	71	5	1.3	NW	2.41	3.6	NW	9.8	9.8
01/01/2009	03:00	8.6	10	8.6	74	4.2	0.9	NW	1.61	3.1	NNW	8.6	8.4
	Continuación	Index	Bar	Rain	Rate	D-D	D-D	Temp	Hum	Samp	Tx	Recept	Int.
		12.4	999.2	0	0	0.106	0	21.7	32	684	1	100	30
		11.7	1000.1	0	0	0.125	0	21.8	32	680	1	99.4	30
		11.1	999.7	0	0	0.14	0	21.9	32	669	1	97.8	30
		10.1	1000.9	0	0	0.161	0	22.1	32	661	1	96.6	30
		9.6	999.4	0	0	0.174	0	22.1	32	661	1	96.6	30
	8.4	1000.3	0	0	0.204	0	22.1	32	661	1	96.6	30	

Figura 4.7 Formato 2 de la base de datos del INIFAP

4.1.2.3 Bases de datos de CAEM

La red de EMA's de CAEM, cuenta con observaciones registradas desde el año 2000 a la actualidad para todo el Estado de México. Las observaciones son registradas con intervalos de 30 minutos para todas las variables meteorológicas. Para esta base de datos se encontraron

cuatro formatos distintos dentro de cada una de las estaciones, los cuales cambian por periodos inesperados, es decir cambian en distintas fechas de formato.

En las figuras 4.8 a la 4.11, se muestran los cuatro formatos para la estación “Presa Allende” en Guanajuato. Se puede observar que en los formatos cambia la forma en registrar la fecha y hora, el nombre y la información básica de la estación, así como la cantidad de información que se tiene.

Date	Time	Dir	WSMDir	WSK	WSMK	AvgTemp	AvgRh	AvgBP	Rain	AvgSR	Batt	SPanel
2001 Ene 01	00:00	224	231	6.8	9.1	18.8	45	816.1	0	13	13.8	11.8
2001 Ene 01	00:10	230	213	5.7	8.5	18.2	47	816.2	0	5		
2001 Ene 01	00:20	216	235	6.9	8.5	17.7	49	816.3	0	-2		
2001 Ene 01	00:30	225	217	5.1	7.4	17.1	51	816.5	0	-5		
2001 Ene 01	00:40	242	238	6.1	7.4	16.8	51	816.5	0	-5		
2001 Ene 01	00:50	250	253	5.1	6.8	16.5	52	816.6	0	-5		
2001 Ene 01	01:00	298	288	4	5.6	16	55	816.8	0	-5		
2001 Ene 01	01:10	234	210	3.3	6.8	15.4	58	816.8	0	-5		
2001 Ene 01	01:20	211	228	3	6.2	14.8	63	816.9	0	-5		
2001 Ene 01	01:30	224	221	1.9	2.8	14.1	64	817	0	-6		
2001 Ene 01	01:40	289	25	1.8	2.8	13.7	64	817.1	0	-5		
2001 Ene 01	01:50	126	129	1.6	2.2	13.4	69	817.2	0	-5		
2001 Ene 01	02:00	144	235	1.6	2.2	13	68	817.3	0	-5		
2001 Ene 01	02:10	197	226	2.4	6.2	12.7	71	817.4	0	-5		
2001 Ene 01	02:20	240	240	4.1	6.2	12.7	69	817.4	0	-4		
2001 Ene 01	02:30	218	272	3.1	5.6	12.7	69	817.4	0	-5		
2001 Ene 01	02:40	191	221	3.1	7.4	12.4	70	817.6	0	-5		
2001 Ene 01	02:50	177	203	5.6	10.2	12.2	70	817.7	0	-4		
2001 Ene 01	03:00	237	250	4.3	7.9	12	69	817.8	0	-4	13.3	0
2001 Ene 01	03:10	321	337	6	8.5	12.3	67	818	0	-3		
2001 Ene 01	03:20	288	287	5.7	8.5	12.6	64	818.2	0	-2		
2001 Ene 01	03:30	179	68	2.7	5.6	12.8	67	818.3	0	-3		
2001 Ene 01	03:40	80	353	3.2	7.9	12.7	70	818.5	0	-3		
2001 Ene 01	03:50	148	168	3.2	5.6	12.6	73	818.6	0	-2		
2001 Ene 01	04:00	277	286	5.2	10.8	12.6	71	818.7	0	-2		

Figura 4.8 Formato 1 de la base de datos de CAEM; estación “Presa Allende”

nombre_estacion	fecha	DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativa	PresBaromet	Precipitacion
PRESA ALLENDE, GTO (ECHA)	01/01/2010 00:00	267	284	7.8	11.4	22.3	31	818.1	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 00:10	245	262	6.5	9.7	21.9	33	818.3	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 00:20	230	227	5.5	6.8	21.4	34	818.4	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 00:30	227	217	6	7.9	20.9	35	818.5	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 00:40	241	242	6.3	8.5	20.4	35	818.5	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 00:50	237	242	6.7	7.9	20.1	36	818.5	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 01:00	234	228	6.5	7.4	19.8	38	818.7	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 01:10	239	247	5.2	6.8	19.5	39	818.8	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 01:20	247	242	3.8	6.2	19.3	40	818.9	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 01:30	139	80	2.7	4.5	18.8	46	818.9	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 01:40	65	76	2.5	4.5	18.2	46	819	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 01:50	111	268	2	3.9	17.9	45	819.1	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 02:00	261	271	4.2	5.6	17.8	45	819.2	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 02:10	298	264	2	3.3	17.7	46	819.3	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 02:20	273	263	3	5.6	17.4	46	819.3	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 02:30	254	269	4.8	6.8	17.4	46	819.4	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 02:40	260	279	3.6	5.6	17.2	47	819.6	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 02:50	17	279	0	0	16.9	49	819.7	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 03:00	139	138	1.6	2.2	16.3	58	819.8	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 03:10	303	320	3.7	8.5	16	53	820	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 03:20	186	199	3.2	5.1	16	56	820	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 03:30	235	169	2.1	3.9	15.7	56	820.1	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 03:40	47	53	2.8	6.2	15.4	58	820.2	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 03:50	290	282	2.3	4.5	15.2	57	820.3	0
PRESA ALLENDE, GTO (ECHA)	01/01/2010 04:00	280	294	3.1	5.1	15	58	820.3	0

Figura 4.9 Formato 2 de la base de datos de CAEM; estación “Presa Allende”

fecha	DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativa	PresBarometri	Precipitacio	RadSolar	nombre_estacion
01/01/2011 00:00	218	204	15.2	22.9	20.3	13		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 00:10	216	202	16.1	27.5	20	13		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 00:20	224	219	14.8	26.4	19.8	13		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 00:30	221	205	11.9	20	19.4	14		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 00:40	221	213	12.5	20.6	19.1	15		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 00:50	219	219	12.1	18.9	18.8	16		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 01:00	230	220	10.6	18.3	18.5	16		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 01:10	233	235	9.3	14.9	18	17		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 01:20	221	220	10.5	14.3	17.7	18		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 01:30	224	225	8.9	14.3	17.6	18		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 01:40	225	215	8.7	12.5	17.4	19		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 01:50	226	221	7.3	9.7	17.2	19		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 02:00	231	231	7.4	11.4	16.8	20		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 02:10	241	249	8.1	12	16.5	20		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 02:20	260	213	3.5	6.2	16.2	21		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 02:30	291	283	3.2	6.2	15.5	22		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 02:40	202	264	2.4	5.6	14.8	24		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 02:50	250	223	4.7	12	14.1	25		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 03:00	229	231	10.5	17.7	14.9	26		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 03:10	223	210	8.8	13.7	15.6	25		0		PRESA ALLENDE, GTO (ECHA)
01/01/2011 03:20	223	204	10.4	17.7	15.7	25		0		PRESA ALLENDE, GTO (ECHA)

Figura 4.10 Formato 3 de la base de datos de CAEM; estación “Presa Allende”

Date	Time	Temp Out	Hi Temp	Low Temp	Out Hum	Dew Pt.	Wind Speed	Wind Dir	Wind Run
01/01/2013	12:30 a. m.	17.2	17.2	15.9	48	6.1	14.5	WSW	7.24
01/01/2013	01:00 a. m.	16.9	17.3	16.7	46	5.2	17.7	SW	8.85
01/01/2013	01:30 a. m.	16.8	17.1	16.7	46	5.1	19.3	WSW	9.66
01/01/2013	02:00 a. m.	17.1	17.1	16.8	45	5	16.1	WSW	8.05
01/01/2013	02:30 a. m.	17.8	17.8	16.9	43	5.1	16.1	SW	8.05
01/01/2013	03:00 a. m.	17.9	17.9	17.3	43	5.2	16.1	SW	8.05
01/01/2013	03:30 a. m.	18.1	18.2	17.8	44	5.6	12.9	SW	6.44
Continuación 1	Hi	Hi	Wind	Heat	THW		Rain	Heat	
	Speed	Dir	Chill	Index	Index	Bar	Rate	D-D	
	30.6	WSW	16.2	16	15	763.6	0	0.024	
	38.6	WSW	15.1	15.7	13.9	763.6	0	0.029	
	35.4	SW	14.6	15.6	13.3	763.1	0	0.032	
	32.2	NW	15.6	15.8	14.3	762.9	0	0.027	
	32.2	WSW	16.5	16.5	15.2	762.7	0	0.01	
37	WSW	16.7	16.6	15.3	762.3	0	0.008		
32.2	WNW	17.7	16.8	16.4	762	0	0.006		
Continuación 2	Cool	In	In	In	In	In	In Air	Wind	
	D-D	Temp	Hum	Dew	Heat	EMC	Density	Samp	
	0	15.2	76	11	15	14.97	1.2147	702	
	0	15.4	75	11	15.2	14.66	1.2139	702	
	0	15.6	74	10.9	15.3	14.25	1.2126	702	
	0	15.7	72	10.7	15.4	13.75	1.2118	702	
	0	15.9	71	10.7	15.7	13.46	1.2105	702	
0	16.3	69	10.6	15.9	12.98	1.2085	702		
0	16.6	69	10.9	16.3	12.99	1.2064	703		

Figura 4.11 Formato 4 de la base de datos de CAEM; estación “Presa Allende”

4.1.2.4 Bases de datos de PEMBU

La red de EMA's de PEMBU sólo cubre el área de la Ciudad de México. Tiene su base de datos en línea y se pueden descargar en la página <https://ruoa.unam.mx>. Cuenta con observaciones registradas desde el año 1996 hasta la fecha en algunas estaciones. Las observaciones son registradas con intervalos de 15 minutos para todas las variables meteorológicas. Esta red sólo maneja un formato único como se muestra en la figura 4.12.

Date	Time	Temp Index	Humidity	Outside Temperature	Wind Chill	Hi Temperature	Low Temperature	Outside Humidity
17/09/1996	6:30 p	13.0	13.0	11.2	13.0	12.9	53	5.4 N
17/09/1996	6:45 p	12.9	12.9	12.9	12.9	12.8	53	5.4 N
17/09/1996	7:00 p	13.0	13.0	13.0	13.1	12.9	53	4.0 N
17/09/1996	7:15 p	13.2	13.2	13.2	13.3	13.1	54	3.6 N
17/09/1996	7:30 p	13.4	13.4	12.4	13.5	13.3	54	5.4 WSW
17/09/1996	7:45 p	13.1	13.1	11.2	13.3	12.7	54	4.9 NW
17/09/1996	8:00 p	12.6	12.6	10.7	12.7	12.5	54	4.9 NNW
17/09/1996	8:15 p	12.4	12.4	9.8	12.5	12.4	54	5.8 NNW
17/09/1996	8:30 p	12.4	12.4	12.4	12.4	12.3	55	4.5 NW
17/09/1996	8:45 p	12.4	12.4	12.4	12.7	12.3	55	1.8 NNW
17/09/1996	9:00 p	12.4	12.4	12.4	12.7	12.2	55	2.7 NW
17/09/1996	9:15 p	12.0	12.0	12.0	12.2	11.8	55	2.2 WNW
17/09/1996	9:30 p	12.1	12.1	12.1	12.4	11.9	55	2.7 WNW
17/09/1996	9:45 p	12.7	12.7	12.7	12.8	12.4	55	2.7 NW

Figura 4.12 Formato de la base de datos de PEMBU

4.1.2.3 Bases de datos del IITCA

La red del IITCA, cuenta con sólo 3 estaciones hasta ahora. Una de ellas, se ubica en las instalaciones de este Instituto (estación uno) y contiene datos registrados desde el año 2001 a la fecha. Las otras dos se encuentran en San Felipe del Progreso y Rectoría de la UAMex (estación dos y tres respectivamente), pero sólo cuentan con registros diarios del 2018 hasta ahora. Para la estación uno, se encontraron dos tipos de formato uno del 2001 al 2007 con registros a cada 30 minutos y el otro del 2001 a la actualidad con registros a cada 10 minutos. Las estaciones dos y tres, sólo tienen un formato con registros diarios. Estos formatos se muestran en las figuras 4.13 a la 4.15, en ellas se puede observar que en los formatos cambia la forma de registrar la fecha y hora y el orden y cantidad de las variables meteorológicas.

Fecha	Hora	Temp.	P.Atmos.	D.Viento	V.Viento	Lluvia	Hr	Lluvia i
04/04/2002	12:05	52	921	240	0	0	8	0
04/04/2002	12:30	52	921	230	0	0	8	0
04/04/2002	13:00	52	921	230	3	0	8	0
04/04/2002	13:30	52	921	290	3	0	8	0
04/04/2002	14:00	52	921	280	0	0	8	0
04/04/2002	14:30	52	921	330	3	0	8	0
04/04/2002	15:00	52	921	250	1	0	8	0
04/04/2002	15:30	52	921	140	0	0	8	0
04/04/2002	16:00	52	921	30	9	0	8	0
04/04/2002	16:30	52	921	40	1	0	8	0
04/04/2002	17:00	52	921	40	6	0	8	0
04/04/2002	17:30	52	921	270	0	0	8	0

Figura 4.13 Formato 1 de la estación del IITCA (2001-2007)

Fecha	Hora	Temp.	P.Atmos.	D.Viento	DV STR	V.Viento	Lluvia	Hr
01/10/2001	10:30:00 a. m.	15	921	230	-99999	0	0	78
01/10/2001	11:00:00 a. m.	17	921	220	-99999	1	0	67
01/10/2001	11:30:00 a. m.	19	921	310	-99999	1	0	62
01/10/2001	12:00:00 p. m.	20	921	210	-99999	0	0	50
01/10/2001	12:30:00 p. m.	21	921	170	-99999	0	0	41
01/10/2001	01:00:00 p. m.	22	921	290	-99999	0	0	41
01/10/2001	01:30:00 p. m.	22	921	70	-99999	0	0	37
01/10/2001	02:00:00 p. m.	23	921	90	-99999	0	0	39
01/10/2001	02:30:00 p. m.	23	921	350	-99999	6	0	35
01/10/2001	03:00:00 p. m.	22	921	330	-99999	1	0	38
01/10/2001	03:30:00 p. m.	22	921	10	-99999	6	0	34
01/10/2001	04:00:00 p. m.	22	921	340	-99999	1	0	35
01/10/2001	04:30:00 p. m.	22	921	90	-99999	1	0	33
01/10/2001	05:00:00 p. m.	21	921	120	-99999	3	0	33
01/10/2001	05:30:00 p. m.	20	921	120	-99999	1	0	37

Figura 4.14 Formato 2 de la estación del IITCA (2001-actualidad)

id_ema	fecha	precip	temp	dv	vv	hum_a	pres_at
3	31/10/2018 00:00	0	14.1	67.5	1.6	79.1	96.3
3	01/11/2018 00:00	0	7.5	360	0.1	100	96.3
3	02/11/2018 00:00	0	9.4	157.5	0.3	100	96.3
3	03/11/2018 00:00	0	9.3	337.5	0	100	96.3
3	03/11/2018 00:00	0	8	135	0.5	100	96.3
3	04/11/2018 00:00	1	9.1	67.5	0.4	100	96.3
3	04/11/2018 00:00	0	8.9	0	0	100	96.3
3	05/11/2018 00:00	0	9.3	67.5	0	100	96.3
3	05/11/2018 00:00	0	8.4	0	0	100	96.3
3	05/11/2018 00:00	0	7.4	0	0	100	96.3
3	06/11/2018 00:00	0	7.5	112.5	0.3	100	96.3
3	06/11/2018 00:00	0	5.9	90	0.2	100	96.3
3	07/11/2018 00:00	0	10.7	22.5	0.3	97.8	96.3
3	08/11/2018 00:00	0	15.4	22.5	0.3	70.2	96.3
3	08/11/2018 00:00	0	11.3	112.5	0.3	80.5	96.3
3	09/11/2018 00:00	0	8.3	112.5	0.9	92.4	96.3

Figura 4.15 Formato 3 de las estaciones 2 y 3 del IITCA

4.1.2.4 ID oficial de las variables meteorológicas

Las base de datos de las EMA's no proporcionan un ID oficial como las variables que proporcionan las EMC's de CLICOM. por ello en la tabla 4.2 se tiene el listado oficial con las claves (ID) de las variables meteorológicas que se emplearon en este proyecto obtenidas de la tabla 0291 del Manual de claves internacionales (OMM-306, 2018) . Excepto para la

variable de radiación solar (Rs) ya que ésta no se encuentra reflejada dentro del manual, por lo cual se optó por utilizar el ID 1000.

Tabla 4.2 ID oficial de las variables metrológicas (OMM-306, 2018)

Variable	Abreviación	ID de la variable	Unidades
Temperatura ambiente	Tamb	004	°C
Temperatura máxima	Tmáx	005	°C
Temperatura mínima	Tmín	006	°C
Precipitación	P	050	mm
Evaporación	Ev	257	mm
Dirección del viento	Dv	020	°
Velocidad del viento	Vv	021	m/s
Humedad relativa	Hr	013	%
Radiación solar	Rs	1000	W/m ²
Presión atmosférica	Patm	001	hPa

4.1.3 Formatos de entrada la herramienta hidro-informática

Debido a la variedad de formatos que se tienen para las EMA's y ESIME's, se ha realizado un análisis de las características que comparten la mayoría de éstos. Siendo la común entre los formatos la forma de registrar la fecha y la forma en que se presentan los datos nulos. Con base en lo anterior, se procedió a estandarizar los formatos de las bases de datos de EMA's y ESIME's para que cumplan las siguientes características:

- I. La forma en que se presente la fecha y hora de registro será en dos opciones. El primero consta en una celda única y el segundo en dos celdas, en uno la fecha y en otra la hora
 1. [Día/mes/año hora] por ejemplo: [13/01/2017 14:00]
 2. [Día/mes/año] [14:00] por ejemplo: [13/01/2017] [14:00]

- II. El orden de registro de las variables no importa, ya que el usuario tendrá la opción de elegir la variable que desea analizar, esto seleccionando la columna en que se encuentre.
 - III. La presentación de los datos nulos es con “---” como lo maneja la mayoría de las estaciones.
 - IV. El tipo de archivo es en *.csv, ya que es el más usual para almacenar bases de datos.
- Los ejemplos de estos formatos se muestran en las figuras 4.16 y 4.17.

fecha	DirViento	DirRafaga	RapViento	RapRafaga	TempAire	HumRelativa	PresBarometric	Precipitacion	RadSolar
01/12/2016 00:00	305	281	0.5	1.7	11.2	88	725.8	0	0
01/12/2016 00:10	59	70	0.9	2.4	11	88	725.9	0	0
01/12/2016 00:20	52	21	0.3	1.1	10.4	90	727.1	0	0
01/12/2016 00:30	6	45	0.3	1.1	9.9	91	727.1	0	0
01/12/2016 00:40	97	203	0.5	1.5	9.5	92	727.2	0	0
01/12/2016 00:50	89	57	0.5	1.1	9.1	93	727.3	0	0
01/12/2016 01:00	58	79	0.4	1.1	8.7	93	727.4	0	0
01/12/2016 01:10	97	72	0.4	0.9	8.3	94	727.5	0	0
01/12/2016 01:20	67	64	0.6	0.9	8	95	727.6	0	0
01/12/2016 01:30	74	63	0.4	0.9	7.6	95	727.7	0	0

Figura 4.16 Formato 1 de la herramienta hidro informática

Date	Time	Temp Out	Hi Temp	Low Temp	Out Hum	Dew Pt.	Wind Speed	Wind Dir	Wind Run	Hi Speed
13/10/2005	14:30	22.7	22.7	19.7	46	10.5	0	SW	0	2.2
13/10/2005	15:00	24.9	24.9	22.7	40	10.4	0.4	NNE	0.8	2.7
13/10/2005	15:30	25.5	25.5	24.6	35	8.9	0	S	0	2.7
13/10/2005	16:00	25.1	26.6	25.1	38	9.8	0.9	ESE	1.61	3.1
13/10/2005	16:30	23.6	25.1	23.6	37	8	0.4	ESE	0.8	3.6
13/10/2005	17:00	22.9	23.7	22.9	41	8.9	0.4	ESE	0.8	3.1
13/10/2005	17:30	19.7	22.9	19.7	50	8.9	0	E	0	2.2
13/10/2005	18:00	18.6	19.7	18.6	53	8.8	0.4	E	0.8	2.7
13/10/2005	18:30	19.3	19.3	18.5	43	6.4	0.4	ESE	0.8	2.7
13/10/2005	19:00	17.7	19.3	17.7	57	9.1	0.9	W	1.61	3.1
13/10/2005	19:30	15.9	17.7	15.9	66	9.6	0.4	WNW	0.8	2.7
13/10/2005	20:00	14.8	15.8	14.8	67	8.7	0	---	0	0
13/10/2005	20:30	13.1	14.8	13.1	74	8.6	0	---	0	0

Figura 4.17 Formato 2 de la herramienta hidro informática

Para el caso de la base de datos CLICOM, como su formato es único, se maneja tal cual como es el archivo *.csv. En la herramienta el usuario sólo elegirá la calve de la estación y el tipo de variable que desea analizar.

4.2 Base de datos de huracanes

La temporada de huracanes es un periodo en el ciclo anual de formación de ciclones tropicales. En esa temporada, es monitoreada y al igual que los datos meteorológicos, se registran todos los fenómenos en una base de datos (NOAA, 2019). Esta base de datos contiene información que describe el fenómeno que se presenta, el lugar y la trayectoria que sigue. También describe el grado de afectación y el territorio impactado. Para el presente proyecto, sólo se extrajo la información de fecha, Océano, nombre y categoría el fenómeno, así como la lista de los estados afectados. Estos datos se ordenaron en un formato como se muestra en la figura 4.18.

AÑO	DÍA	MES	DIA JULIANO	OCEANO	NOMBRE	CATEGORIA	ESTADOS AFECTADOS DIRECTAMENTE	ID ESTADOS
1970	31	JULIO	212	ATLÁNTICO	CELIA	TT	COAH	05
1970	1	AGOSTO	213	ATLÁNTICO	CELIA	TT	COAH	05
1970	2	AGOSTO	214	ATLÁNTICO	CELIA	TT	COAH	05
1970	3	AGOSTO	215	ATLÁNTICO	CELIA	TT	COAH	05
1970	4	AGOSTO	216	ATLÁNTICO	CELIA	TT	COAH	05
1970	5	AGOSTO	217	ATLÁNTICO	CELIA	TT	COAH	05
1970	31	AGOSTO	243	PACÍFICO	NORMA	DT	BCS	03
1970	1	SEPTIEMBRE	244	PACÍFICO	NORMA	DT	BCS	03
1970	2	SEPTIEMBRE	245	PACÍFICO	NORMA	DT	BCS	03
1970	3	SEPTIEMBRE	246	PACÍFICO	NORMA	DT	BCS	03
1970	4	SEPTIEMBRE	247	PACÍFICO	NORMA	DT	BCS	03
1970	5	SEPTIEMBRE	248	PACÍFICO	NORMA	DT	BCS	03

Figura 4.18 Formato para la base de datos de huracanes (NOAA, 2019)

4.3 Desarrollo y programación de la herramienta hidro-informática

El objetivo principal del presente trabajo ha consistido en diseñar una herramienta hidro-informática con base en un conjunto de métodos robustos para el análisis, control de calidad y complemento de los registros meteorológicos. Estos métodos se encuentran descritos en el capítulo 3 y se codifican en la herramienta MatLab. Esta codificación de cada uno de los métodos se realiza siguiendo los procesos mostrados en los diagramas de flujo en el anexo 4.2, Cada uno de estos métodos, se encuentran integrados en la interfaz final de la herramienta.

4.3.1 Interfaz de la herramienta hidro-informática

El proceso de operación de la herramienta hidro- informática consta de 6 etapas principales. Estas etapas siguen un proceso como se muestra en el diagrama de flujo de la figura 4.19 donde se describe de manera general el funcionamiento de la herramienta hidro-informática. En la etapa 1, se realiza la bienvenida al usuario, así como la acción de abrir archivos o cerrar la aplicación. Posteriormente en la etapa 2, se selecciona el tipo de archivos a cargar, estos son: bases de datos de EMA's, CLICOM o archivos *.mat (bases de datos ya organizadas por la herramienta). También se ingresa la información general de la estación, las variables meteorológicas a analizar. Para el caso de los archivos *.mat, sólo se selecciona el folder de trabajo y conduce a la etapa 5. En la etapa 3 se realiza la selección del periodo de análisis y la organización de los datos en el formato de la herramienta. La etapa 4 consiste en guardar toda la información anterior y las bases de datos organizadas en el folder de trabajo. Básicamente en las etapas anteriores, se realiza la preparación de los datos para pasar al análisis de ellos.

La etapa 5, consiste en la selección de los distintos métodos de análisis de calidad, divididos en módulos de umbrales físicos, coherencia interna, coherencia espacial, homogeneidad, complemento de datos y resumen estadístico. Cada método tiene su proceso específico, que como ya se mencionó, se muestran en los diagramas de flujo del anexo 4.3.

Una vez realizado el proceso del método seleccionado, se puede obtener el reporte de resultados, guardar y continuar con un nuevo análisis o en su caso, salir del interfaz, lo cual se encuentra en la etapa 6.

Para el desarrollo de las interfaces que conforman la herramienta hidro-informática, los diagramas de flujo del anexo 4.4 describen paso a pasos los procesos específicos del análisis de calidad y complemento para cada una de ellas.

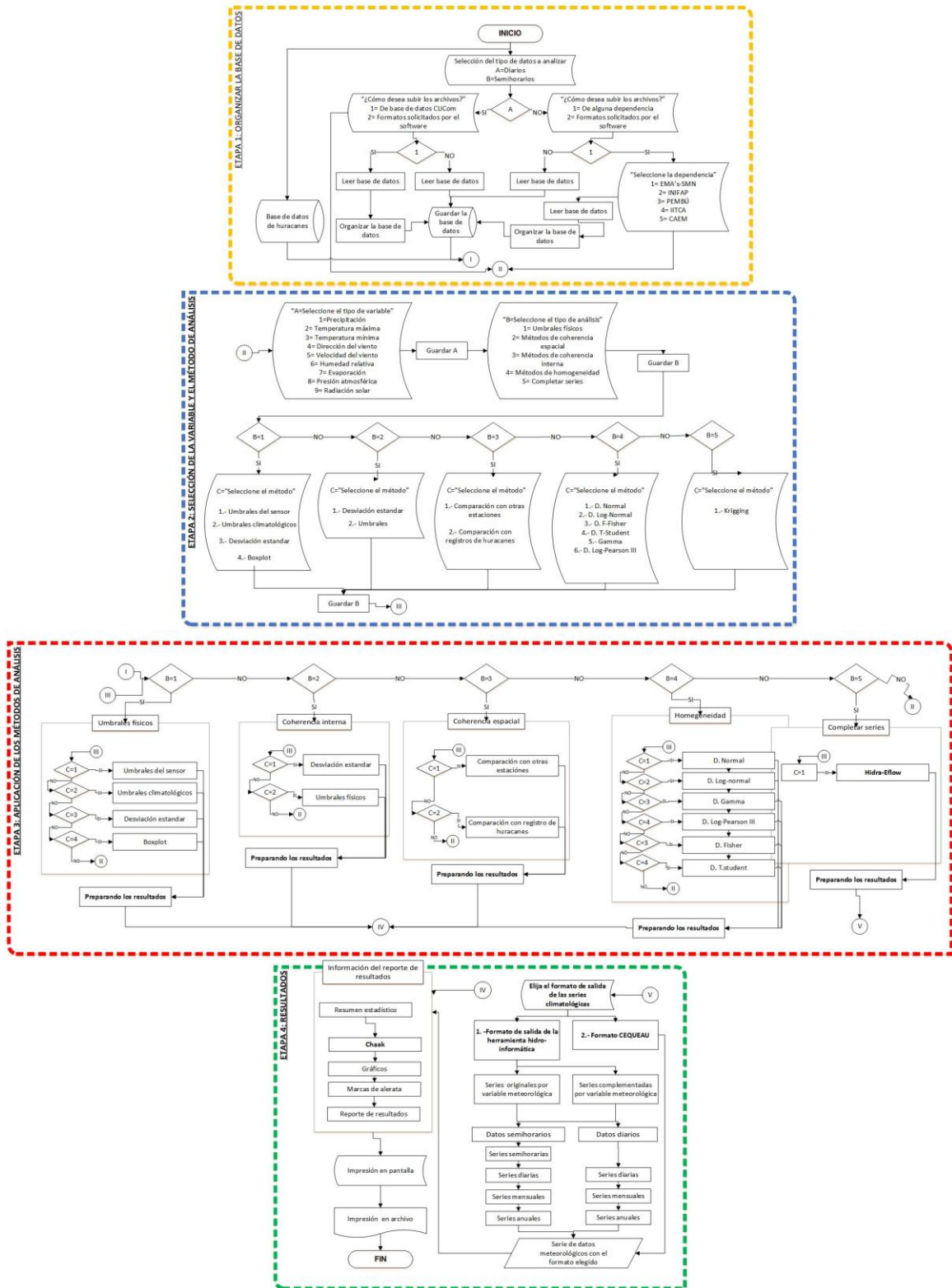


Figura 4.19 Diagrama de flujo de las etapas generales de la herramienta hidro-informática

Capítulo 5

Resultados

El analizar la calidad de datos meteorológicos es un proceso indispensable para que éstos sean fundamento en la toma de decisiones de proyectos de recursos hídricos. Cabe mencionar que siempre habrá un grado de incertidumbre, esto se debe a que algunas estaciones son difíciles de controlar ya que su consistencia y ubicación no son siempre las más adecuadas. O en su caso se presenta algún fenómeno meteorológico, o humano, que altere la medición de las variables meteorológicas.

Con la finalidad de facilitar el análisis de calidad de datos meteorológicos, se desarrolló la herramienta hidro-informática que contiene los métodos de análisis de calidad de datos meteorológicos propuestos en los capítulos anteriores. Esta propuesta metodológica de análisis ha sido aplicada a la base de datos de las EMA's, ESIME's y ECM's del mapex del Estado de México. De esta manera se efectúa un análisis de calidad y complemento de datos, así como caracterizar la calidad de información que la red de monitoreo brinda a los usuarios.

5.1 Herramienta hidro-informática "SICCDAMET"

Como resultado principal a esta investigación se tiene el desarrollo de la herramienta hidro-informática SICCDAMET (Sistema de Control de Calidad de Datos Meteorológicos) que lleva un logotipo original que distingue a la herramienta. Este logotipo se muestra en la figura 5.1.

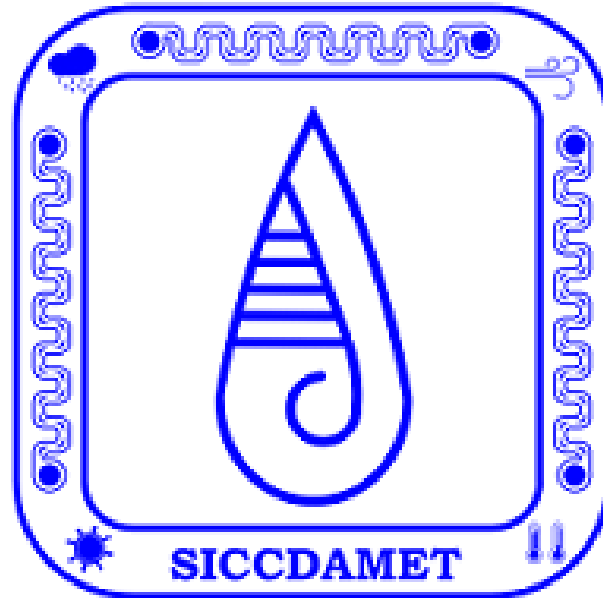


Figura 5.1 Logotipo y nombre de la herramienta hidro-informática

Esta herramienta se desarrolla en el software de lenguaje técnico de programación Matlab, el cual, facilita la programación de cada una de las funciones, métodos de análisis y procesos que realiza la herramienta. Teniendo una interfaz gráfica amigable y de fácil uso, dando la bienvenida al usuario con un interfaz como se muestra en la figura 5.2, en la cual se presenta el nombre y logotipo de la herramienta, el nombre de los colaboradores y de las instituciones donde se desarrolló.

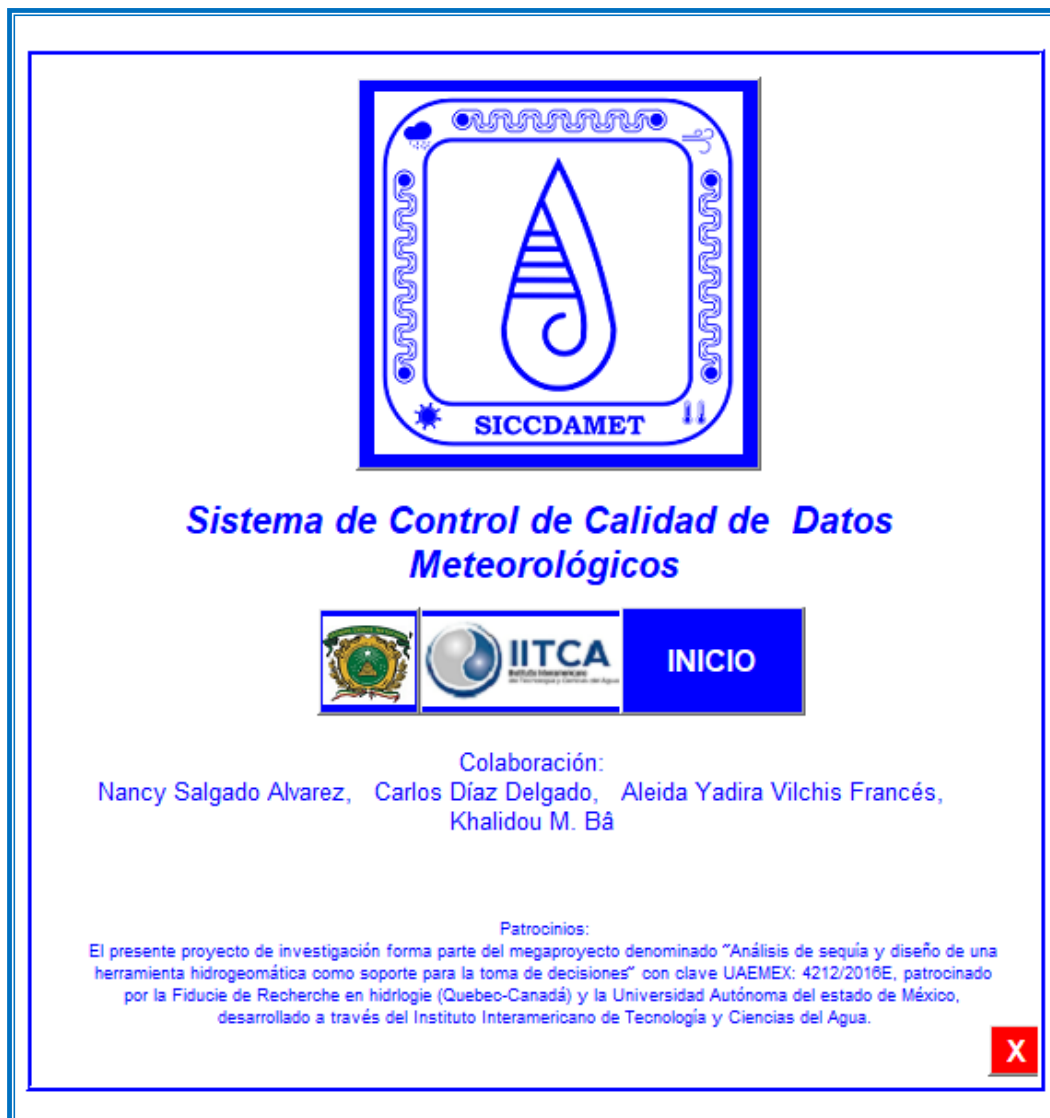


Figura 5.2 Interfaz de bienvenida al usuario de la herramienta hidro-informática

Al dar clic en inicio, se dirige a una interfaz para comenzar con el proceso de organización de la base de datos. Estos procesos se realizan a partir de bases de datos de EMC's o EMA's que cumplan con los formatos de entrada. Esta interfaz de organización muestra la información general de la estación, por ejemplo en la figura 5.3 muestra la información general de la estación "Observatorio-UAEMEX" a partir de una base de datos de CLICOM. También dentro de esta interfaz, se seleccionan las variables a analizar para extraer los registros proporcionados por la estación.

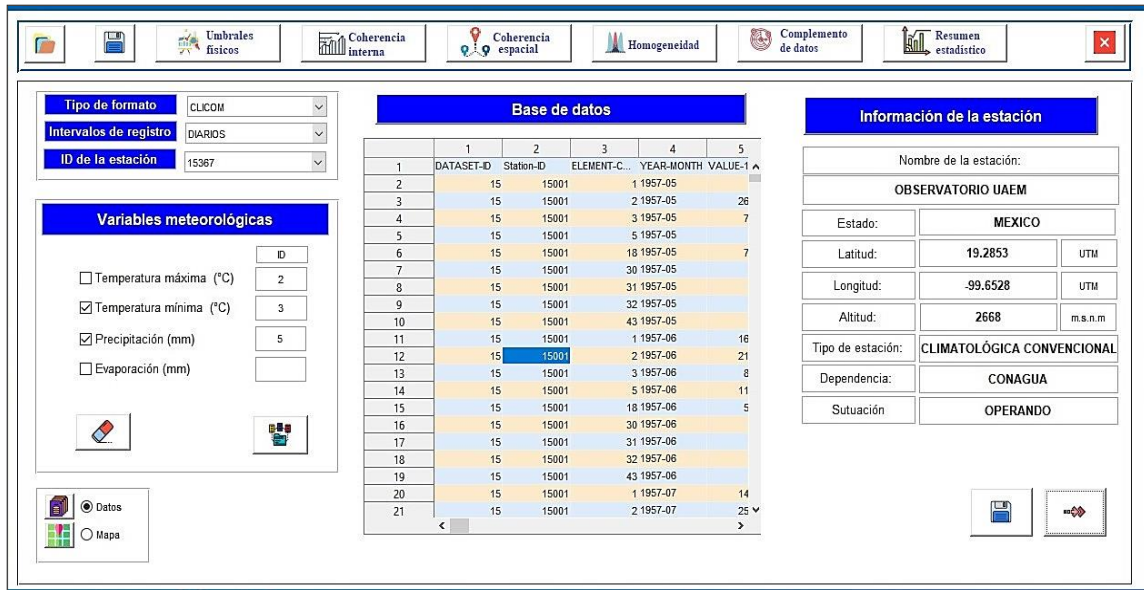


Figura 5.3 Selección de la estación a analizar

Una vez ordenada la base de datos, la herramienta dirige al usuario a una interfaz donde se muestran los diferentes métodos de análisis cómo son umbrales físicos, coherencia interna, coherencia espacial, homogeneidad, complemento de datos y resumen estadístico como se muestra en la figura 5.4. También se resalta en color azul las variables meteorológicas disponibles para aplicar los diversos métodos de análisis, teniendo la opción de seleccionar entre una y dos variables meteorológicas.

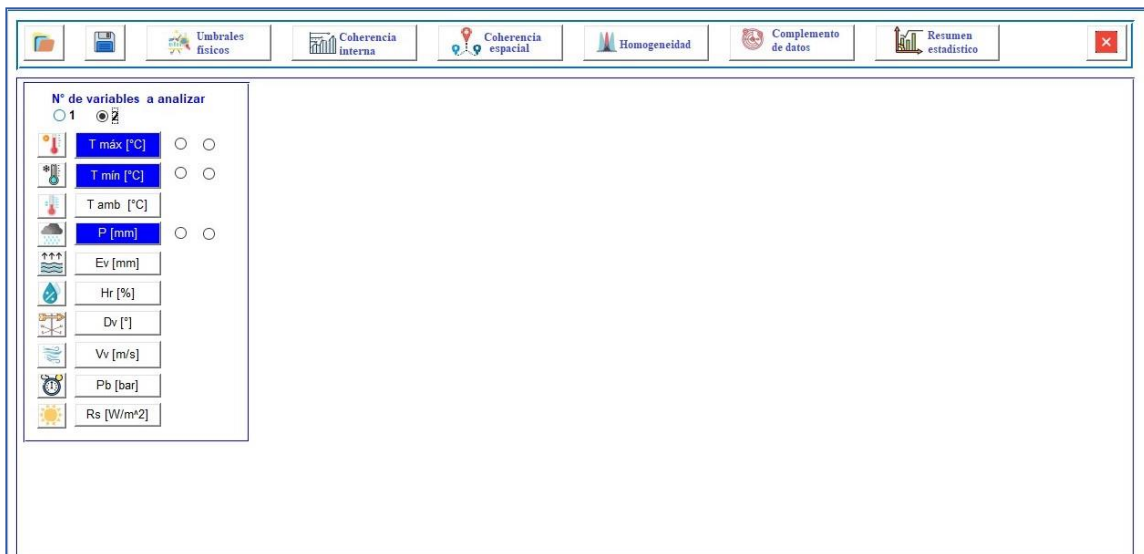


Figura 5.4 Interfaz básica de la herramienta hidro-informática SICCDAMET

Para cada uno de los métodos seleccionados, la herramienta proporciona los resultados del análisis, así como las gráficas que caracterizan a la variable meteorológica. Por ejemplo, en la en la figura 5.5 se muestran los resultados de la estación de Observatorio UAEM con el método desviación estándar para las variables de T_{máx} y T_{mín}.

Se aprecia como resultados del análisis, el número de registros encontrados y el porcentaje de ellos que aprueban el método. También se muestran los límites que se tomaron en cuenta en el método. Estos resultados se grafican para 5 años consecutivos, mostrando en color rojo y azul los límites del método para detectar visualmente la fecha y año en que las variables no cumplen con el método.

También se puede ingresar el valor de la variable meteorológica registrada en tiempo real, con la finalidad de visualizar y comparar su comportamiento respecto a los registros históricos de la estación meteorológica.

Al aplicar cada uno de estos métodos, es posible generar los reportes de resultados. Esto mediante los botones que se activan en la interfaz. Estos reportes consisten en archivos *.xlsx, los cuales descargan un reporte general de la estación y un reporte por método aplicado.

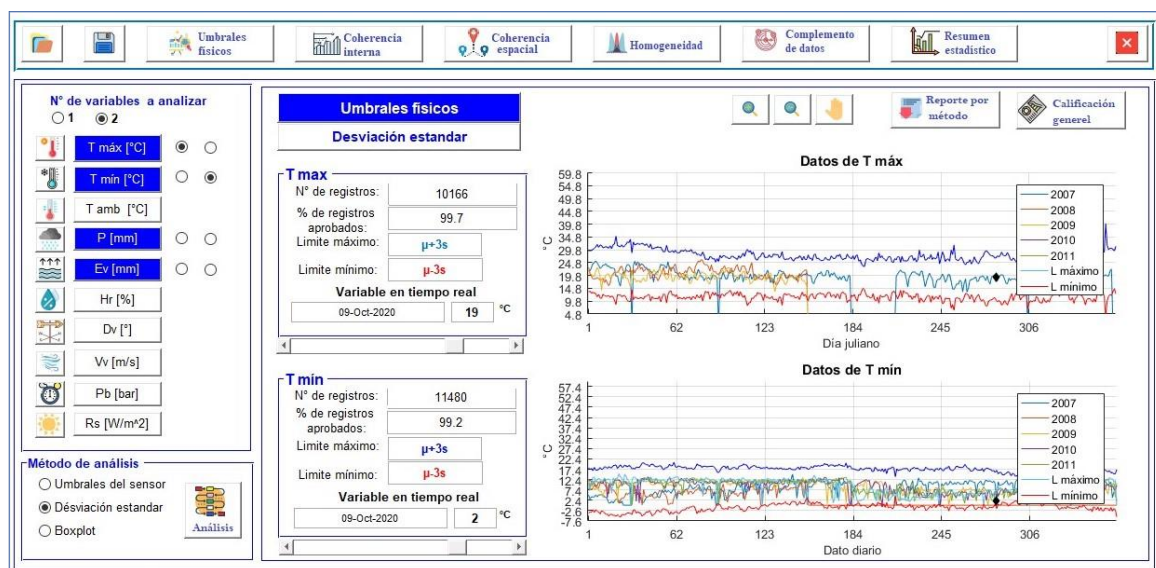


Figura 5.5 Interfaz gráfica de los métodos de análisis

Siendo la figura 5.5 la interfaz gráfica básica de la herramienta hidro-informática, en el anexo 5.1 “Manual de usuario” se muestra detalladamente como utilizar la herramienta.

Cabe resaltar que para llevar a cabo los métodos de interpolación, se agregó un módulo desarrollado en la herramienta hidro informática “HIDRA-Eflow” (Arévalo. M, 2019). Adicional a esto también se cuenta con un módulo para la generación de curvas I-D-Tr de la herramienta “CHAAK” (Becerril, 2020). Estos módulos también se presentan con el detalle suficiente para su uso en el anexo 5.1. Obteniendo de la herramienta SICCDAMET los archivos de entrada para ellos, generados a partir de la base de datos de la estación meteorológica.

Estos formatos de entrada consisten en archivos *.txt con las características específicas del módulo para los registros diarios de $T_{\text{máx}}$, $T_{\text{mín}}$ y P, para el módulo de interpolación. Y para el módulo de curvas I-D-Tr, únicamente los registros de P, con datos máximos en 24 horas, registros anuales y diarios.

5.1.2 Reportes de resultados

Una vez realizados los métodos de análisis con la herramienta hidro informática, el usuario puede generar los reportes de resultados. Esto consiste en proporcionar al usuario la evaluación obtenida por los registros de las variables por cada método aplicado. Colocando un 1 si el dato aprobó el análisis, de lo contrario un 0. Posterior a ello se genera el porcentaje de datos que aprobaron el análisis del total. La forma de proporcionar esta información a los usuarios es por medio de un reporte de resultados generados en archivos *.xlsx para cada uno de los métodos aplicados y un reporte general, el cual resume los resultados obtenidos por la estación.

5.1.2.1 Reporte por método

En este tipo de reportes se proporciona la calificación obtenida por cada método, así como datos relevantes que éstos proporcionen. Siguiendo el ejemplo de la estación “Observatorio UAEM”, la figura 5.6, muestra el reporte obtenido al aplicar el método de desviación estándar para la variable de $T_{\text{máx}}$.

Se puede apreciar en el reporte la información general de la estación, así como el total de registros encontrados y la calificación obtenida con el método de análisis. Posteriormente, se tienen datos de la variable analizada, el total de registros que se tienen y el porcentaje que aprueban el método. Enseguida se observan los registros y su calificación. Estos últimos ordenados por año y día juliano. Marcando en color rojo y azul los límites mínimo y máximo del método respectivamente


RESULTADOS DEL ANÁLISIS DE CALIDAD METEOROLÓGICO																																																				
Tipo de estación:	CLIMATOLÓGICA CONVENCIONAL					Dependencia																																														
Nombre de la estación:	OBSERVATORIO UAEM					CLICOM																																														
Estado:	MEXICO					Tipo de registros:																																														
Latitud:	19.28527778	UTM			Registros anuales																																															
Longitud:	-99.65277778	UTM			Periodo analizado																																															
Altitud:	2668	m.s.n.m			1980	2016																																														
<table border="1"> <tr> <td>Total de registros:</td> <td>10166</td> <td colspan="2">Porcentaje de registros que aprobaron el método:</td> <td>99.7%</td> <td>%</td> <td colspan="4">Calificación</td> </tr> <tr> <td>Variable analizada:</td> <td colspan="5">TEMPERATURA MÁXIMA [°C]</td> <td>1</td> <td colspan="4">Correcto</td> </tr> <tr> <td>Método:</td> <td colspan="5">Umbral físico // Desviación estándar</td> <td>0</td> <td colspan="4">Dudoso</td> </tr> <tr> <td>Variable en tiempo real:</td> <td>19</td> <td>°C</td> <td colspan="2">Fecha de registro:</td> <td>09-oct-20</td> <td>-99999</td> <td colspan="4">Dato faltante</td> </tr> </table>										Total de registros:	10166	Porcentaje de registros que aprobaron el método:		99.7%	%	Calificación				Variable analizada:	TEMPERATURA MÁXIMA [°C]					1	Correcto				Método:	Umbral físico // Desviación estándar					0	Dudoso				Variable en tiempo real:	19	°C	Fecha de registro:		09-oct-20	-99999	Dato faltante			
Total de registros:	10166	Porcentaje de registros que aprobaron el método:		99.7%	%	Calificación																																														
Variable analizada:	TEMPERATURA MÁXIMA [°C]					1	Correcto																																													
Método:	Umbral físico // Desviación estándar					0	Dudoso																																													
Variable en tiempo real:	19	°C	Fecha de registro:		09-oct-20	-99999	Dato faltante																																													
Limite máximo	μ	+	3 s																																																	
Limite máximo	29.40		29.49	29.56	29.95	30.39	30.96	31.89	30.71	31.69																																										
Limite mínimo	μ	-	3 s																																																	
Limite mínimo	13.42		13.72	12.92	13.22	13.00	12.75	11.48	12.44	11.06																																										
1980	22.5		21	21	21.5	21.2	22.5	22	21	21																																										
	1		1	1	1	1	1	1	1	1																																										
1981	21.5		20.2	21.5	21.5	20.2	20.5	21.8	20.5	21																																										
	1		1	1	1	1	1	1	1	1																																										
1982	19.2		19.5	19	19	19.2	17.8	19	19	19.3																																										
	1		1	1	1	1	1	1	1	1																																										
1983	16.5		17.2	18.8	19	17	16.2	16	16.1	15.2																																										
	1		1	1	1	1	1	1	1	1																																										
1984	24.2		22.5	22.5	22.5	22.7	22.8	22.7	21.3	19.2																																										
	1		1	1	1	1	1	1	1	1																																										

Figura 5. 6 Ejemplo de reporte de resultados por método

En el anexo 5.2 se muestran todos los reportes obtenidos para las variables de Tmáx, Tmín, P y Ev de la estación Observatorio UAEM para el periodo de 1980 a 2016.

Para los métodos de coherencia espacial, los reportes contienen características específicas. En efecto, para el caso de coherencia con estaciones cercanas, se muestran los datos de las estaciones compradas con todos os años de registros, así como la correlación que hay entre las estaciones. Para el caso del método huracanes, se muestra la lista con la fecha en que la

ubicación de la estación se encuentra afectada por dicho fenómeno. En la figura 5.7 se presenta un ejemplo del reporte de huracanes para la estación de Observatorio UAEM y la variable de precipitación.


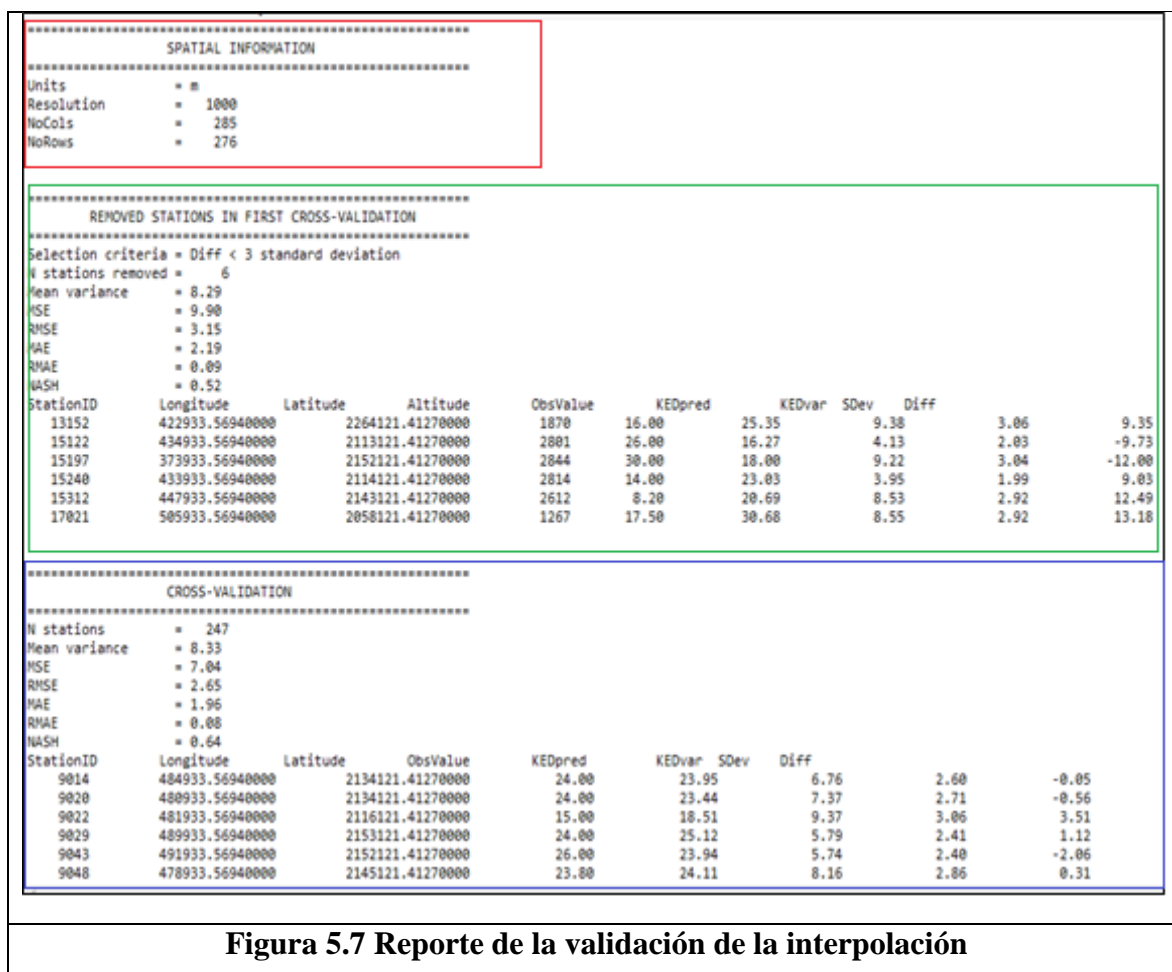
RESULTADOS DEL ANÁLISIS DE CALIDAD METEOROLÓGICO					
Tipo de estación:	CLIMATOLÓGICA CONVENCIONAL			Dependencia	
Nombre de la	OBSERVATORIO UAEM			CLICOM	
Estado:	MEXICO			Tipo de registros:	
Latitud:	19.28527778	UTM		*DIARIOS	
Longitud:	-99.65277778	UTM		Periodo analizado	
Altitud:	2668	m.s.n.m		1900 2020	
Total de registros:	44286	Porcentaje de registros que coinciden:	65.2%	%	Calificación
Variable analizada:	PRECIPITACIÓN [mm]				1 Correcto
Método:	C.ESPACIAL-C.HURACANES				0 Dudoso
N° de datos que coinciden:	66				-99999 Dato faltante
ACRÓNIMOS					
TT = Tormenta tropical					
TS = Tormenta de temporada					
H1 = Huracán categoría 1					
H2 = Huracán categoría 2					
H3 = Huracán categoría 3					
H4 = Huracán categoría 4					
H5 = Huracán categoría 5					
DT = Depresión tropical					
AÑO	DÍA JULIANO	OCEANO	NOMBRE DEL HURACAN	TIPO	P
2003	195	ATLÁNTICO	CLAUDETTE	TT	0
2005	234	ATLÁNTICO	JOSE	TT	12.4
2005	235	ATLÁNTICO	JOSE	TT	3.2
2012	172	PACÍFICO	CARLOTTA	H1	10.5
2012	173	PACÍFICO	CARLOTTA	H1	4.5
2012	174	PACÍFICO	CARLOTTA	H1	0.1
2012	175	PACÍFICO	CARLOTTA	H1	16.9
2014	244	ATLÁNTICO	DOLY	TT	25.5
2014	245	ATLÁNTICO	DOLY	TT	1.9
2014	246	ATLÁNTICO	DOLY	TT	20.1
2015	269	PACÍFICO	MARTY	H1	0
2015	270	PACÍFICO	MARTY	H1	23.5
2015	271	PACÍFICO	MARTY	H1	8.3
2015	272	PACÍFICO	MARTY	H1	1.6
2015	273	PACÍFICO	MARTY	H1	15.7
2015	274	PACÍFICO	MARTY	H1	1.3
2015	275	PACÍFICO	MARTY	H1	0

Figura 5. 7 Ejemplo de reporte de resultados para el método de coincidencia con huracanes

Los resultados que se obtienen a través de los módulos de I-D-Tr consisten en imágenes y datos para la construcción de las curvas a partir de los datos de precipitación. Y para el módulo de interpolación, los reportes de resultado consisten en dos archivos *.txt y una imagen tipo *.tiff. Uno de los archivos contiene los resultados de la interpolación como se muestra en la figura 5.8, se puede apreciar en este reporte primero la información espacial de la interpolación (recuadro color rojo), después el número de estaciones que se removieron por medio de los filtros de calidad de la estación del módulo (recuadro color verde). Y después de ello, los resultados de la validación cruzada de la interpolación (recuadro color azul).



El segundo reporte que proporciona el módulo corresponde a los resultados del variograma como se muestra en la figura 5.9. En dicha figura se presentan los resultados de las estaciones que no se tomaron en cuenta por su inconsistencia (recuadro color rojo), los resultados de la regresión multivariable (recuadro color verde) y finalmente los residuales para el cálculo del variograma (recuadro color azul).

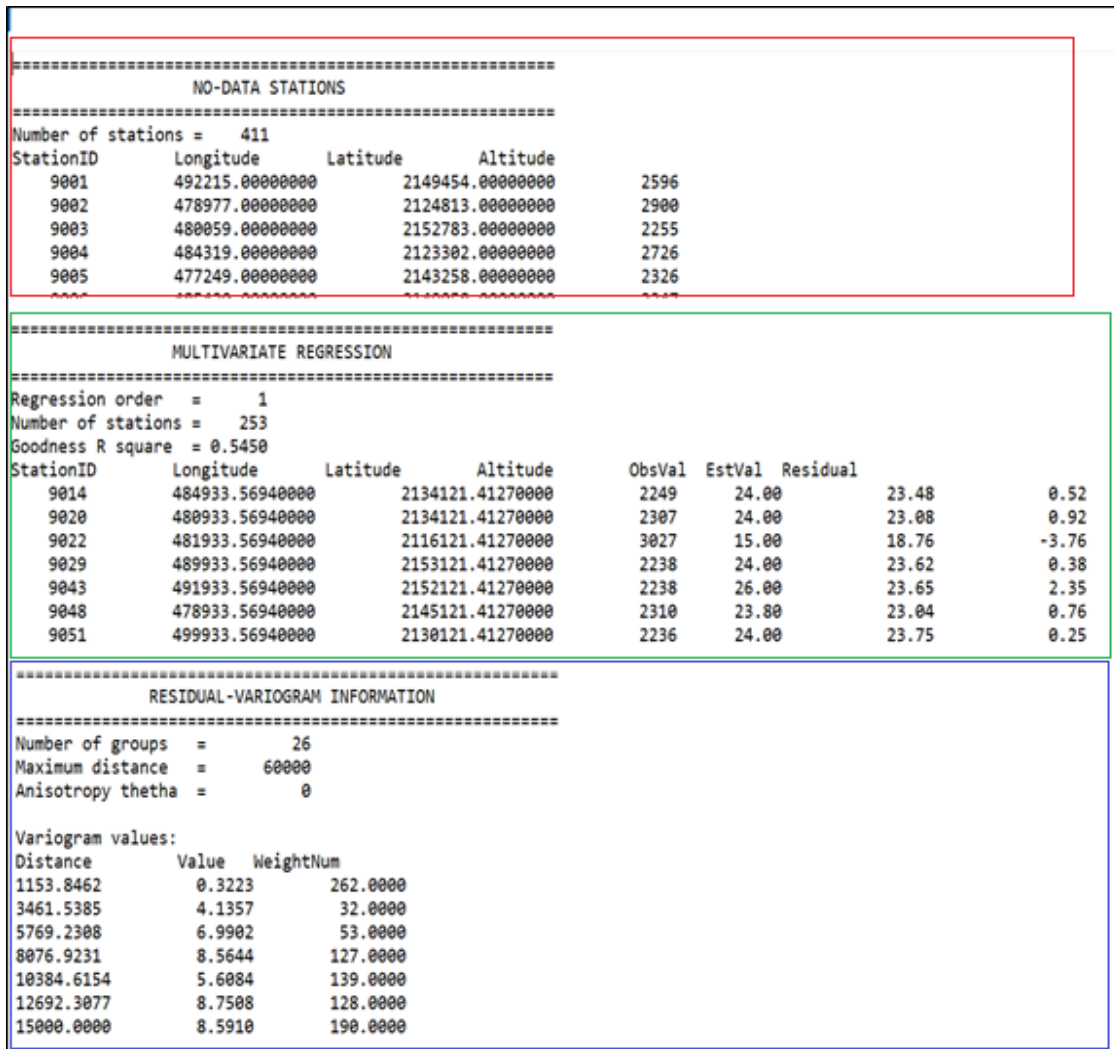


Figura 5.9 Reporte de los resultados del variograma de la interpolación

Finalmente, como resultado de la interpolación, se tiene una imagen tipo *.tiff como la que se muestra en la figura 5.10. En esa figura se muestra un ejemplo de la interpolación de T_{máx} diaria para las estaciones ubicadas en el mapex del Estado de México. Se puede apreciar la ubicación de las estaciones que se consideraron para la interpolación y los resultados del proceso.

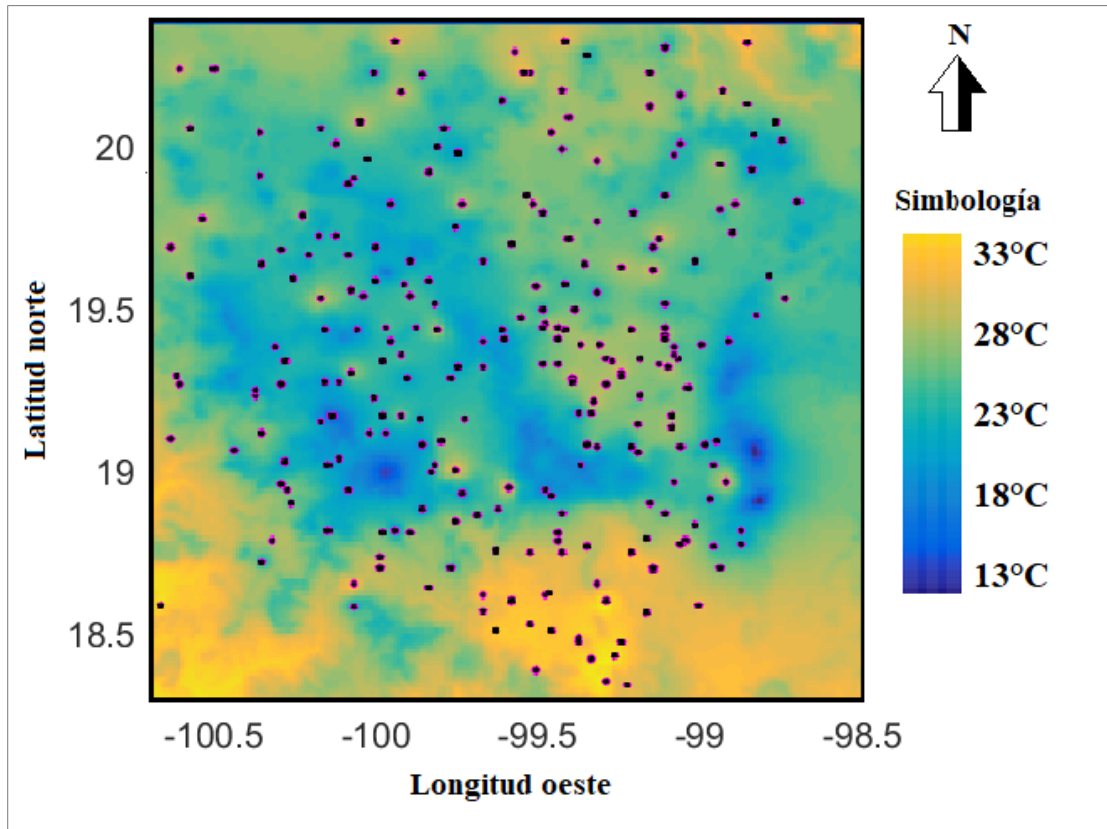


Figura 5.10 Ejemplo de Archivo tipo *.tiff de interpolación para Tmáx de un día en el mápex del Estado de México.

5.1.6.2 Reporte general

Como una forma de presentar los resultados del análisis de las variables meteorológicas por estación de forma general, la herramienta proporciona un el reporte resumen. Este reporte contiene la información general de la estación y después los porcentajes de datos que aprueban los registros de cada variable metrológica por método de análisis aplicado. Adicionalmente, en el reporte se genera una gráfica de barras donde se comparan estos porcentajes para todas las variables.

Un ejemplo de este reporte se encuentra en la figura 5.11, donde se muestran los resultados de la estación Observatorio UAEM para las variables de Tmáx, Tmín y P. Es posible observar en la gráfica del reporte que la estación tiene porcentajes superiores al 80% para la mayoría de los métodos aplicados en las 4 variables meteorológicas.

RESULTADOS DEL ANÁLISIS DE CALIDAD METEOROLÓGICO

CLIMATOLÓGICA CONVENCIONAL		Dependencia								
OBSERVATORIO UAEM		CLICOM								
MEXICO		Tipo de registros:								
19.28527778	UTM	DIARIOS								
-99.65277778	UTM	Periodo analizado								
2668	m.s.n.m	1980	2016							
Total de datos analizados:										
Métodos										
VARIABLES METEOROLÓGICAS										
	Tmáx °C	Tmín °C	T amb °C	P mm	Ev mm	Hr %	Rs W/m ²	Pb Bar	Dv °	Vv m/s
Umbrales físicos	100%	100%		100%	100%					
Desviación estándar	100%	99.2%		36%	100%					
Boxplot	98.5%	99.0%		22%	99%					
Umbrales físicos temporales		100%		92%	100%					
Desviación estándar temporal	100%	100%		100%	100%					
D. Normal	55%	78%		70%	88%					
D. Log-Normal	64%	0%		0%						
D. Gamma	89%	87%		94%						
D. Log-Pearson III	81%	78%		92%	81%					
D.T-Student	85%	50%		100%	87%					

Figura 5.11 Formato del reporte general del análisis de calidad de datos

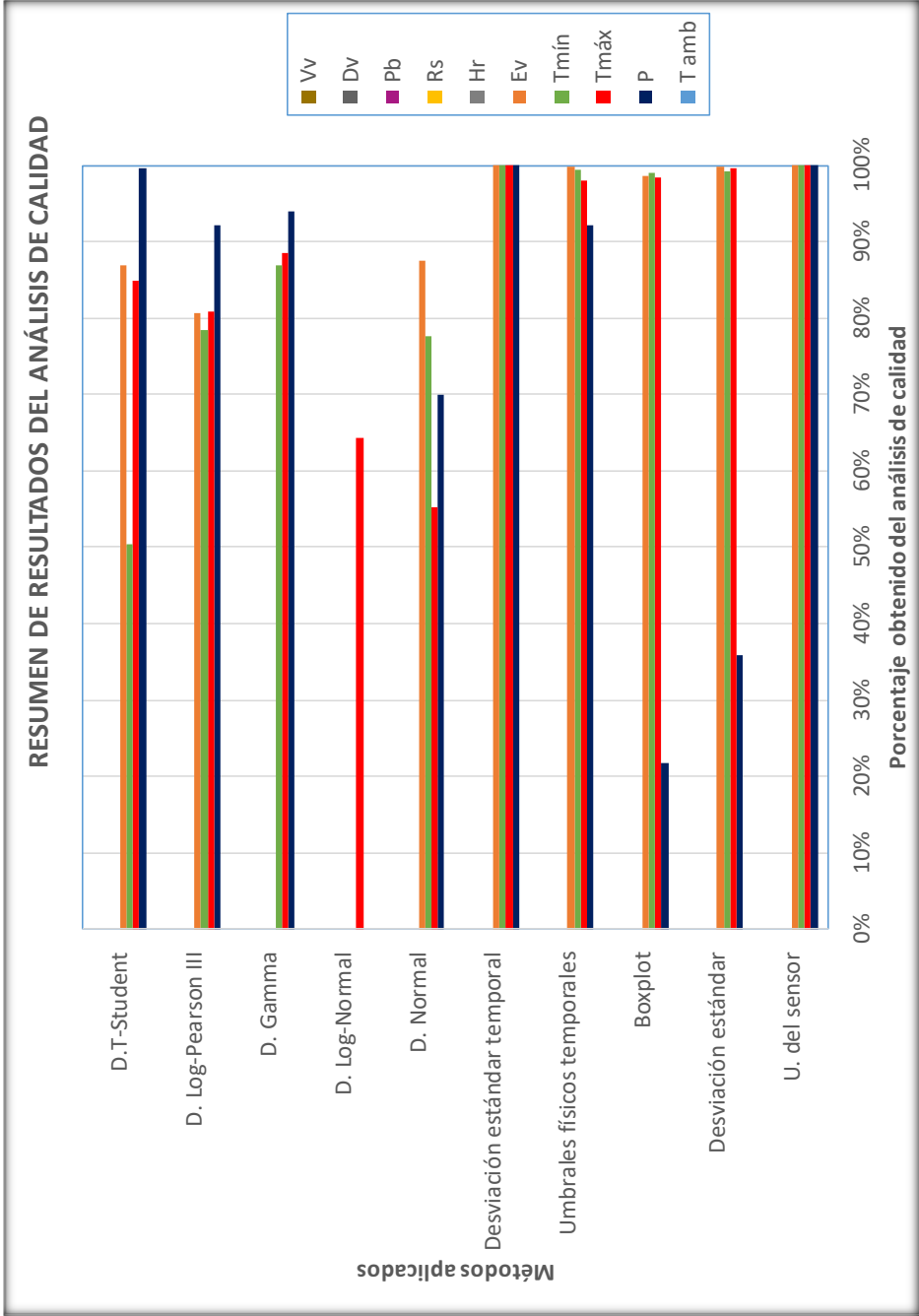


Figura 5.11 Formato del reporte general del análisis de calidad de datos (continuación)

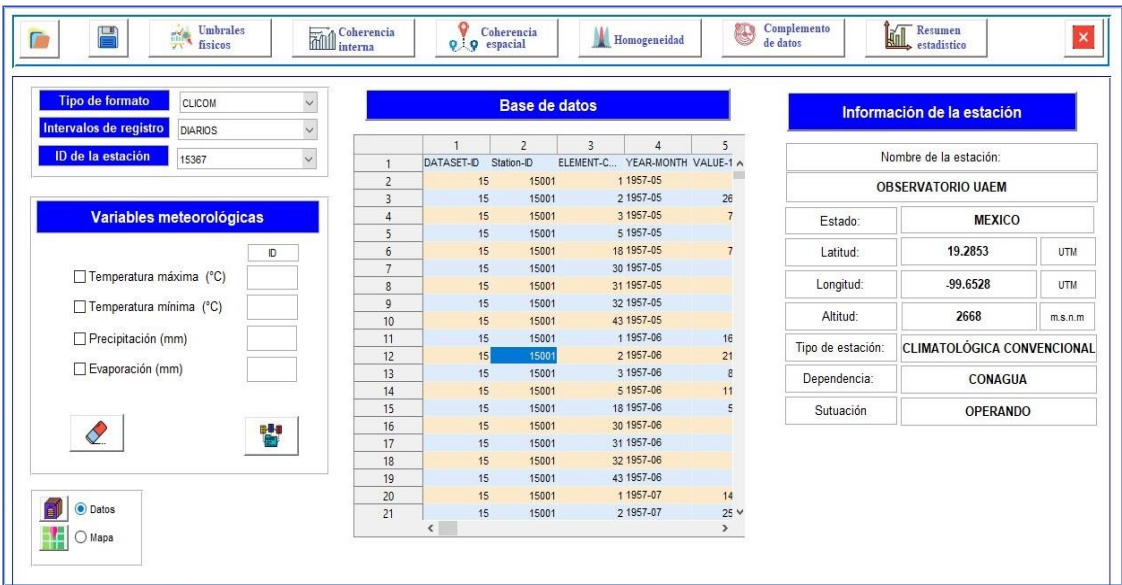
5.2 Validación de la herramienta hidro-informática SICCDAMET

Para comprobar que los diversos métodos de análisis aplicados brindan al usuario información verídica y resultados correctos, se realiza a continuación la validación de todos los métodos de análisis aplicados. Los resultados proporcionados por la herramienta hidro-informática fueron contrastados con los obtenidos con ayuda de una hoja electrónica, de Excel, mostrando los resultados obtenidos en el anexo 5.3.

Para llevar a cabo la validación, se considera la base de datos de la estación meteorológica convencional “OBSERVATORIO UAEM” con coordenadas 19.2853, -99.6528 y altitud 2668 (latitud, longitud, altitud), ID=15367 y con las variables meteorológicas de temperatura máxima, temperatura mínima y precipitación (T_{máx}, T_{mín} y P), con un periodo de análisis de 1980 al 2016.

5.2.1 Organización de los datos

De acuerdo con los formatos de entrada de la herramienta, en este caso se selecciona la base de datos CLICOM del Estado de México y la estación con ID 15367, la cual efectivamente corresponde al Observatorio de la UAEM, como se muestra en la figura 5.12.



The screenshot displays the SICCDAMET software interface. At the top, there is a menu bar with icons for 'Umbralés físicos', 'Coherencia interna', 'Coherencia espacial', 'Homogeneidad', 'Complemento de datos', and 'Resumen estadístico'. Below the menu, there are three main panels:

- Left Panel:** Contains dropdown menus for 'Tipo de formato' (set to CLICOM), 'Intervalos de registro' (set to DIARIOS), and 'ID de la estación' (set to 15367). Below these are checkboxes for 'Variables meteorológicas': 'Temperatura máxima (°C)', 'Temperatura mínima (°C)', 'Precipitación (mm)', and 'Evaporación (mm)'. There are also input fields for 'ID' and 'm.s.n.m'.
- Center Panel:** Titled 'Base de datos', it shows a table with columns: 1, 2, 3, 4, 5. The table lists 21 rows of data, including DATASET-ID, Station-ID, ELEMENT-C., YEAR-MONTH, and VALUE-1.
- Right Panel:** Titled 'Información de la estación', it displays details for 'OBSERVATORIO UAEM'. Fields include: 'Nombre de la estación: OBSERVATORIO UAEM', 'Estado: MEXICO', 'Latitud: 19.2853 UTM', 'Longitud: -99.6528 UTM', 'Altitud: 2668 m.s.n.m', 'Tipo de estación: CLIMATOLÓGICA CONVENCIONAL', 'Dependencia: CONAGUA', and 'Situación: OPERANDO'.

Figura 5.12 Selección de la estación a analizar

A partir de ahí, se selecciona el periodo de análisis 1980 a 2016 para ajustar la base de datos a ese periodo. Para corroborar que los datos que extrae y ajusta la herramienta son los de la base de datos, se realizó la extracción manual, mostrando una parte de ellos en la tabla 5.1. Donde se comparan los datos originales de la base de datos del Estado de México, los datos obtenidos con la herramienta hidro-informática y los datos manuales. Se puede apreciar la congruencia de los datos en las tres imágenes de la tabla. Mostrando los datos completos en el Anexo 5.3 “Validación”.

Tabla 5.1 Validación de la extracción de datos

		Base de datos							
Original CLICOM	ID de temperatura máxima:2								
	DATASET-	Station-ID-	ELEMENT-	YEAR-MO	VALUE-1	FLAG1-1	VALUE-2	FLAG1-2	VALUE-3
	15	15367	1	2006-01	5.7		6.5		6.4
	15	15367	2	2006-01	19.5		20.8		22.5
	15	15367	3	2006-01	6		6		7
	15	15367	5	2006-01	0		0		0
15	15367	18	2006-01	4.5		3.8		3.8	
Herramienta SICCDAMET	Variable de Temperatura máxima [T _{máx}]								
	Año	Día Juliano							
		1	2	3	4	5	6		
	2006	19.5	20.8	22.5	21.2	21.7	22		
	2007	21.6	21.6	22.9	24.4	24.8	25.5		
	2008	23	20.6	19.4	18.5	18.5	18.6		
	2009	20.4	20.7	20.1	19.5	19.5	21.4		
	2010	-99999	-99999	-99999	-99999	-99999	-99999		
2011	-99999	-99999	-99999	-99999	-99999	-99999			
Extracción manual	AÑO	DÍA JULIANO	TEMPERATURA MÍNIMA	TEMPERATURA MÁXIMA	PRECIPITACIÓN	EVAPORACIÓN			
			°C	°C	mm	mm			
	2006	1	6	19.5	0	4.5			
	2006	2	6	20.8	0	3.8			
	2006	3	7	22.5	0	3.8			
	2006	4	5	21.2	0	4.5			
	2006	5	6	21.7	0	3.2			
	2006	6	6	22	0	4.2			
2006	7	2	23.5	0	4.9				

5.2.2 Métodos de análisis

Los métodos y graficas de la herramienta ayudan al usuario a verificar la calidad de la información mediante los límites y características que establece cada método. Describiendo el comportamiento de las variables y compararlo con los datos históricos de la estación. Para corroborar los resultados que proporciona la herramienta, se retoma la estación de Observatorio UAEM, a la cual se aplicaron de forma manual todos los métodos para el periodo de 1980 al 2016 (ver detalles en anexo 5.3 “Validación”). Estos resultados son contrastados con los obtenidos con la herramienta hidro informática.

Los resultados detallados de la herramienta hidro-informática se encuentran en los reportes del análisis del anexo 5.2. Cabe resaltar que en la validación manual se calcularon las probabilidades de los métodos de homogeneidad, por medio de las funciones de distribuciones disponibles y programadas en Matlab. A continuación, en la tabla 5.2 se tiene resumida la validación. Mostrando los resultados encontrados manualmente y con la herramienta para los métodos de umbrales físicos, coherencia interna, homogeneidad.

Tabla 5.2 Validación de la estación Observatorio UAEM

Tipo de estación:	CLIMATOLÓGICA CONVENCIONAL				Dependencia			
Nombre de la estación:	OBSERVATORIO UAEM				CLICOM			
Estado:	MEXICO				Tipo de registros:			
Latitud:	19.28527778	UTM		DIARIOS				
Longitud:	-99.65277778	UTM		Periodo analizado				
Altitud:	2668	m.s.n.m		1980	2016			
Total de registros	13542							
	Manual				SICCDAMET			
	Tmáx [°C]	Tmín [°C]	P [mm]	Ev [mm]	Tmáx [°C]	Tmín [°C]	P [mm]	Ev [mm]
N° de registros efectivos	10166	11480	8714	5745	10166	11480	8714	5745
Desviación estándar	2.94	4.15	5.71	2.68	2.94	4.15	5.71	2.68
Media	19.71	6.95	1.88	2.01	19.71	6.95	1.88	2.01
Máximo	47.8	16	103	15.41	47.8	16	103	15.41
Mínimo	6.8	-5.6	0	0	6.8	-5.6	0	0

Tabla 5.2 Validación de la estación Observatorio UAEM (continuación)

Método de análisis	Manual				SICCDAMET			
	Tmáx	Tmín	P	Ev	Tmáx	Tmín	P	Ev
Umbrales del sensor	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Desviación estándar	100.0	99.2	36.0	100.0	100.0	99.2	36.0	100.0
Boxplot	98.5	99.0	22.0	99.0	98.5	99.0	22.0	99.0
U. F. temporales	98.0	100.0	92.0	100.0	98.0	100.0	92.0	100.0
D. E temporal	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
D. Normal 99.7%	0.0	69.3	36.8	96.7	0.0	69.3	36.8	96.7
D. Normal 95%	70.7	67.3	84.1	69.6	70.7	67.3	84.1	69.6
D. Normal 68%	95.2	96.6	89.3	96.5	95.2	96.6	89.3	96.5
D. L-Normal 99.7%	99.9	0.0	0.0	99.1	99.9	0.0	0.0	99.1
D. L-Normal 95%	93.8	0.0	0.0	96.3	93.8	0.0	0.0	96.3
D. L-Normal 68%	0.0	0.0	0.0	47.2	0.0	0.0	0.0	47.2
D. T-student 99.7%	66.2	19.7	99.6	75.9	66.2	19.7	99.6	75.9
D. T-student 95%	90.5	47.2	99.6	88.4	90.5	47.2	99.6	88.4
D. T-student 68%	97.8	84.4	99.7	96.6	97.8	84.4	99.7	96.6
D. Gamma 99.7%	99.1	100.0	97.4	99.5	99.1	100.0	97.4	99.5
D. Gamma 95%	99.7	96.7	94.7	95.1	99.7	96.7	94.7	95.1
D. Gamma 68%	72.0	64.4	99.0	70.8	72.0	64.4	99.0	70.8
D. Log-Pearson III 99.7%	96.8	97.2	95.7	97.2	96.8	97.2	95.7	97.2
D. Log-Pearson III 95%	88.6	88.1	92.8	85.5	88.6	88.1	92.8	85.5
D. Log-Pearson III 68%	57.2	49.7	88.3	57.6	57.2	49.7	88.3	57.6

Con respecto a los métodos de coherencia espacial, el método de estaciones cercanas relaciona mediante un resumen estadístico y una correlación las estaciones con las que se comparan los datos, haciendo una extracción de los datos coincidentes con el periodo de

registro. En la tabla 5.3 se tiene un ejemplo de los resultados manuales de algunos días del año 2006 de la variable de T_{máx}, utilizando para la estación de Observatorio UAEM Obteniendo los estadísticos característicos por día de los datos de 3 estaciones más cercanas.

Tabla 5.3 Validación manual de coherencia con estaciones cercanas

Estaciones comparadas:

N°	Estación	Estado	Latitud	Longitud					
1	Nueva Oxtotitlán	México	19.2769	-99.6858					
2	Escuela Sec. 8 Manuel Bernal	México	19.2831	-99.6683					
3	Toluca (DGE)	México	19.2917	-99.6306					
Comparación manual									
Año	Día juliano	Observatorio UAEM	1	2	3	Prom.	Des. est	Máx	Mín
2006	1	19.5	18.5	18.5	19.5	15.40	0.58	19.50	18.50
2006	2	20.8	19.2	19.8	19.2	16.20	0.75	20.80	19.20
2006	3	22.5	18.7	22.5	15.7	16.48	3.30	22.50	15.70
2006	4	21.2	20.4	22.2	20.4	17.64	0.85	22.20	20.40
2006	5	21.7	18.7	22.7	18.7	17.36	2.06	22.70	18.70
2006	6	22	18.4	22	15.4	16.76	3.19	22.00	15.40
2006	7	23.5	12.2	23.5	12.2	15.68	6.52	23.50	12.20
2006	8	22.4	18	22.4	17	17.56	2.86	22.40	17.00
2006	9	23.6	18.4	23.6	17.4	18.40	3.32	23.60	17.40
2006	10	23.6	16.6	23.6	18.6	18.48	3.56	23.60	16.60
2006	11	19.1	16.7	18.1	18.7	16.72	1.05	19.10	16.70
2006	12	19.6	16.7	18.6	18.7	17.12	1.22	19.60	16.70
2006	13	20.2	18.7	19.2	17.7	17.76	1.04	20.20	17.70

Para el método de coherencia con huracanes, se busca en la base de datos de huracanes de la herramienta, la coincidencia de ubicación y fecha de registro de los huracanes con los datos de la estación a analizar. Para verificar estos datos de búsqueda se realizó manualmente, encontrando los datos ubicados en la validación, mostrando algunos de ellos en la tabla 5.4.

Tabla 5.4 Validación manual de los resultados de coherencia con huracanes para la estación Observatorio UAEM

Fecha		Resultados manuales				Resultados con la herramienta			
Año	Día juliano	Océano	Nombre del huracán	Tipo	Variable meteorológica	Océano	Nombre del huracán	Tipo	Variable meteorológica
1988	243	Atlántico	Debby	H1	-99999	Atlántico	Debby	H1	-99999
1988	244	Atlántico	Debby	H1	-99999	Atlántico	Debby	H1	-99999
1988	245	Atlántico	Debby	H1	-99999	Atlántico	Debby	H1	-99999
1988	246	Atlántico	Debby	H1	17.8	Atlántico	Debby	H1	17.8
1988	247	Atlántico	Debby	H1	18.2	Atlántico	Debby	H1	18.2
1988	248	Atlántico	Debby	H1	19	Atlántico	Debby	H1	19
1988	249	Atlántico	Debby	H1	19.5	Atlántico	Debby	H1	19.5
1988	250	Atlántico	Debby	H1	23	Atlántico	Debby	H1	23
1988	251	Atlántico	Debby	H1	22.5	Atlántico	Debby	H1	22.5
1989	169	Pacífico	Cosme	H1	18.5	Pacífico	Cosme	H1	18.5
1989	170	Pacífico	Cosme	H1	18	Pacífico	Cosme	H1	18
1989	171	Pacífico	Cosme	H1	16.9	Pacífico	Cosme	H1	16.9
1989	172	Pacífico	Cosme	H1	18.7	Pacífico	Cosme	H1	18.7
1989	173	Pacífico	Cosme	H1	20.1	Pacífico	Cosme	H1	20.1
1989	174	Pacífico	Cosme	H1	20	Pacífico	Cosme	H1	20
2014	246	Atlántico	Doly	TT	-99999	Atlántico	Doly	TT	-99999

La validación de los datos interpolados se realiza internamente mediante el método de validación cruzada y los resultados se encuentran en el archivo *.txt que proporciona. Esta validación consiste en realizar la interpolación quitando una a una la observación de las estaciones involucradas y estimar su valor por interpolación. E último con la finalidad de observar las posibles incertidumbres que se tienen en la interpolación al faltar uno de los datos.

A partir de ello, se obtiene los valores observados y calculados por la validación cruzada, así como los errores y diferencias de cada uno de los resultados. Un ejemplo de esta validación

se muestra en la tabla 5.5, donde se tiene una parte de los resultados de la interpolación del día 15 de agosto del 2016 para la variable T_{máx}, resaltando en color amarillo los resultados obtenidos para esa fecha en la estación Observatorio UAEM. (Los datos de coordenadas están proyectadas en coordenadas cónicas conforme a Lambert).

Tabla 5.5 Validación cruzada de la interpolación del día 15/08/2016 de la variable T_{máx}

Validación cruzada 15/08/2016							
N° de estaciones = 256							
Media de la varianza= 7.92							
MSE = 6.04							
RMSE = 2.46							
MAE = 1.88							
RMAE = 0.08							
NASH = 0.46							
ID	Longitud	Latitud	Valor observado	Valor predicho	Varianza de la interpolación	Desviación estándar	Diferencia
15367	430933.569	2132121.41	21.3	22.8	8.4	2.9	1.5
15368	387933.569	2117121.41	20	23.89	8.91	2.98	3.89
15371	422933.569	2093121.41	20	23.53	7.08	2.66	3.53
15372	419933.569	2164121.41	24	22.12	8.16	2.86	-1.88
15373	435933.569	2116121.41	23	23.51	6.58	2.56	0.51
15374	399933.569	2131121.41	20	18.38	7.37	2.72	-1.62
15378	378933.569	2142121.41	20	22.86	5.39	2.32	2.86
15383	500933.569	2153121.41	25.5	25.58	8.24	2.87	0.08
15390	416933.569	2180121.41	21	22.13	8.43	2.9	1.13
15391	387933.569	2152121.41	22	22.62	3.24	1.8	0.62
15392	386933.569	2109121.41	24	23.33	6.5	2.55	-0.67
15393	450933.569	2187121.41	21	23.51	9.08	3.01	2.51

Con los resultados de esta validación se corrobora que los datos proporcionados con la herramienta SICCDAMET son adecuados para la variable meteorológica considerada en este proyecto.

5.3 Análisis a la base de datos

Con base en la estructuración de los datos de la herramienta hidro-informática, se organizó la base de datos de las estaciones que se encuentran en el mapex del Estado de México para poder realizar el análisis de calidad de la información. El total de estaciones y datos que se organizaron se muestran en la tabla 5.6. En dicha tabla se indica la dependencia que opera la estación, tipo de estación y el periodo de registros con los que cuenta la base de datos de cada dependencia y las variables meteorológicas que se consideraron. Es importante destacar que más de 547 millones de datos fueron revisados con el análisis de calidad que se implementó en la herramienta hidro-informática.

Tabla 5.6 Total de registros de la base de datos del mapex del Estado de México

Dependencia	Tipo de estación	N° de estaciones	Intervalos de registro	Registros por día	Variables meteorológicas	Periodo de análisis promedio		Aproximado de datos analizados
CONAGUA	EMC	723	24 h	1	Tmáx, Tmín, Ev, P	1920	2019	104788728
SMN	EMA	8	10 min	144	Tamb, Hr, P, Dv, Vv, Pb, Rs	2000	2018	53125632
IITCA	EMA	3	10 min	144	Tamb, Hr, P, Dv, Vv, Pb	2001	2018	16127424
INIFAP	EMA	55	15 y 30 min	96-48	Tamb, Hr, P, Dv, Vv, Pb	2000	2018	208707840
PEMBU	EMA	13	15 min	96	Tamb, Hr, P, Dv, Vv, Pb, Rs	1996	2019	73539648
CAEM	EMA	58	30 min	48	Tamb, Hr, P, Dv, Vv, Pb	2000	2015	91704960
Total de estaciones:		860				Total de datos:		547994232

El listado de las estaciones consideradas se encuentra en el anexo 5.4 donde se tienen los datos relevantes como el nombre, tipo, ubicación (coordenadas UTM y estado). Los resultados globales obtenidos para cada una de las variables y estaciones se muestran a continuación.

5.3.1 Base de datos de las EMA's

Las estaciones correspondientes a las EMA's del Estado de México son administradas por el SMN, IITCA, INIFAP, PEMBU y CAEM. Al realizar el análisis de calidad de datos para cada variable meteorológica, se encontró que la mayoría de las estaciones genera registros de calidad con porcentajes superiores al 90% para los métodos de umbrales físicos, coherencia interna y espacial. Siendo el método de umbrales del sensor el que todas las estaciones aprueban con un porcentaje del 100%. A diferencia de los métodos de homogeneidad donde se presentan porcentajes que van del 0% hasta el 97%. Principalmente las distribuciones de Log-Pearson III y Log-Normal son las que presentan menores porcentajes para las variables de P. Estos resultados se muestran en el anexo 5.5 donde se tiene el listado de todas las estaciones EMA's con el porcentaje de datos que aprobaron cada uno de los métodos y variables analizadas.

En resumen, estos resultados se pueden apreciar en las páginas siguientes, donde se presentan distintas tablas y gráficas en las que se muestra el porcentaje promedio que las estaciones aprobaron por dependencia que las administra y para cada una de las variables meteorológicas que se analizaron. Por ejemplo, la tabla 5.7 muestra los resultados de precipitación y en ella se presentan los porcentajes promedio que obtuvieron las estaciones de cada una de las dependencias que administra las EMA's, así como los porcentajes máximos mínimos y la desviación estándar de ellos. Gráficamente en la figura 5.5 se aprecian los porcentajes promedios obtenidos por cada método. Cabe resaltar que según la mayoría de las gráficas, son los métodos de homogeneidad los que tiene menor porcentaje de aprobación principalmente para las variables de P y Hr. Estos resultados se pueden apreciar con detalle en las tablas 5.8 a la 5.13, y en las figuras 5.14 a la 5.19 para las variables de Hr, Pb, Dv, Vv, Rs y Tamb.

Tabla 5.7 Porcentaje de datos aprobados de las EMA's para la variable de precipitación [P]

Dependencia	Estadístico	Umbrales Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umbrales físicos	D. Estandar	Boxplot	D. Estandar temporal	Umbrales físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
SMN	Promedio	100	95	93	96	95	95	94	88	88	86	84	85
	Máximo	100	99	99	99	99	99	97	93	92	90	90	92
	Mínimo	100	90	90	91	91	90	90	82	82	81	80	80
	D. estd	0	3	3	3	3	3	3	4	5	3	3	5
IITCA	Promedio	100	95	96	95	95	98	95	85	84	83	85	91
	Máximo	100	98	98	99	98	99	98	89	88	84	86	93
	Mínimo	100	91	94	91	92	97	91	80	80	81	84	88
	D. estd	0	5	3	6	4	1	5	6	6	2	1	4
PEMBU	Promedio	100	94	95	94	94	93	94	87	88	86	85	85
	Máximo	100	98	99	99	99	99	99	94	93	93	93	93
	Mínimo	100	90	86	87	81	29	58	80	80	80	45	45
	D. estd	0	3	3	3	4	12	7	4	4	4	8	8
INIFAP	Promedio	100	94	95	94	94	93	94	87	88	86	85	85
	Máximo	100	98	99	99	99	99	99	94	93	93	93	93
	Mínimo	100	90	86	87	81	29	58	80	80	80	45	45
	D. estd	0	3	3	3	4	12	7	4	4	4	8	8
CAEM	Promedio	100	94	95	94	94	95	95	87	86	87	86	86
	Máximo	100	99	99	99	99	99	99	93	93	93	93	93
	Mínimo	100	90	90	90	90	90	90	80	80	80	80	80
	D. estd	0	3	3	3	3	3	3	4	4	4	5	4

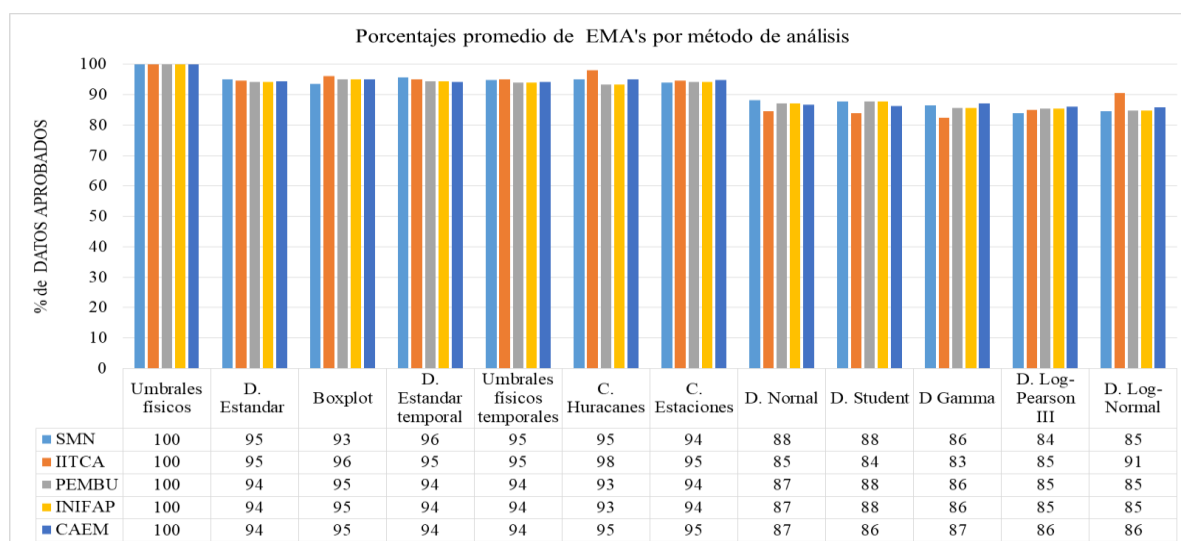


Figura 5.13 Porcentaje promedio de resultados de precipitación [P] para las EMA's

Tabla 5.8 Porcentaje de datos aprobados de las EMA's para la variable de temperatura ambiente [Tamb]

Dependencia	Estadístico	Umbralés Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umbralés físicos	D. Estandar	Boxplot	D. Estandar temporal	Umbralés físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
SMN	Promedio	100	97	96	96	97	77	94	88	88	86	83	83
	Máximo	100	100	100	99	100	94	97	93	92	90	90	92
	Mínimo	100	90	90	91	93	50	90	82	82	81	76	70
	D. estd	0	4	4	3	3	16	3	4	5	3	5	7
IITCA	Promedio	100	95	96	95	99	59	95	85	84	83	85	91
	Máximo	100	98	98	99	100	67	98	89	88	84	86	93
	Mínimo	100	91	94	91	98	50	91	80	80	81	84	88
	D. estd	0	5	3	6	1	12	5	6	6	2	1	4
PEMBU	Promedio	100	95	97	96	97	82	94	87	85	88	85	84
	Máximo	100	99	100	99	100	95	99	93	91	93	92	93
	Mínimo	100	90	90	90	90	50	90	80	80	81	76	70
	D. estd	0	3	4	4	3	17	4	5	4	4	4	8
INIFAP	Promedio	100	94	95	94	96	79	94	87	88	86	85	83
	Máximo	100	100	100	99	100	95	99	94	93	93	92	93
	Mínimo	100	90	86	87	81	29	58	80	80	80	45	45
	D. estd	0	3	4	3	4	18	7	4	4	4	8	9
CAEM	Promedio	100	95	95	94	97	78	95	87	86	87	84	85
	Máximo	100	100	100	99	100	95	99	93	93	93	92	93
	Mínimo	100	90	90	90	90	50	90	80	80	80	76	70
	D. estd	0	4	3	3	3	16	3	4	4	4	5	5

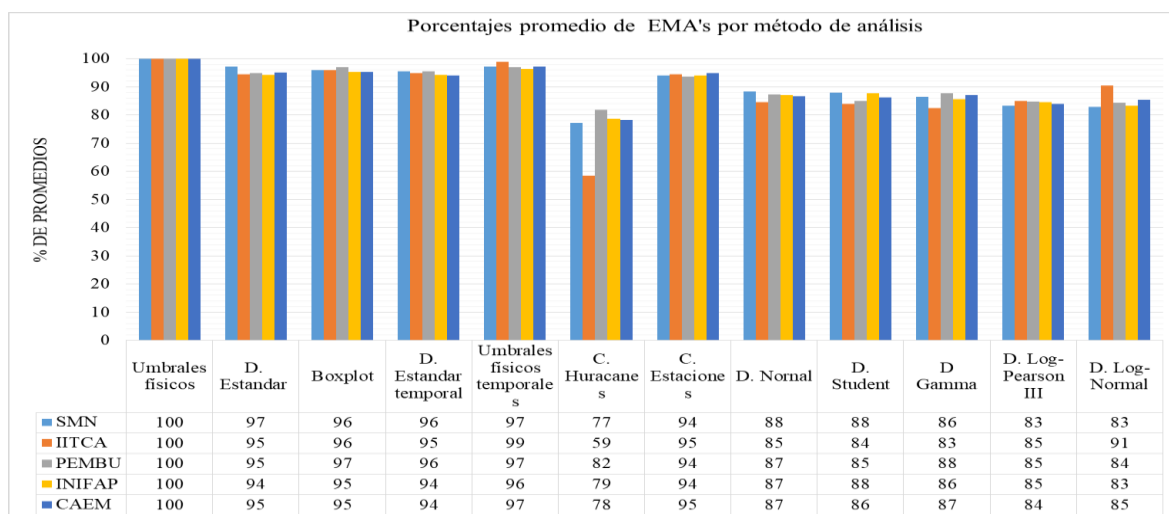


Figura. 5.14 Porcentaje promedio de resultados de temperatura ambiente [Tamb] para las EMA's

Tabla 5.9 Porcentaje de datos aprobados de las EMA's para la variable de humedad relativa [Hr]

Dependencia	Estadístico	Umbralés Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umbralés físicos	D. Estandar	Boxplot	D. Estandar temporal	Umbralés físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
SMN	Promedio	100	95	93	98	99	78	74	88	88	85	81	79
	Máximo	100	99	99	100	100	94	92	93	92	88	87	91
	Mínimo	100	90	90	95	98	60	56	82	82	81	72	50
	D. estd	0	3	3	2	1	14	16	4	5	3	5	13
IITCA	Promedio	100	95	96	100	99	69	82	85	84	83	85	44
	Máximo	100	98	98	100	100	72	91	89	88	84	86	88
	Mínimo	100	91	94	99	98	65	72	80	80	81	84	0
	D. estd	0	5	3	1	1	5	13	6	6	2	1	62
PEMBU	Promedio	100	95	94	98	98	80	82	87	85	88	77	69
	Máximo	100	99	99	100	100	94	93	93	91	94	89	91
	Mínimo	100	90	90	94	94	56	56	80	80	81	40	0
	D. estd	0	3	3	2	2	15	13	5	4	4	18	32
INIFAP	Promedio	100	94	95	96	98	74	75	87	89	84	82	78
	Máximo	100	98	99	100	100	94	94	94	99	94	91	91
	Mínimo	100	90	86	87	81	29	56	80	80	76	40	0
	D. estd	0	3	3	3	3	16	15	4	6	4	13	17
CAEM	Promedio	100	94	95	97	98	79	77	87	86	86	77	75
	Máximo	100	99	99	100	100	94	94	93	99	94	91	91
	Mínimo	100	90	90	93	94	56	56	80	80	76	40	0
	D. estd	0	3	3	3	2	13	15	4	4	4	15	28

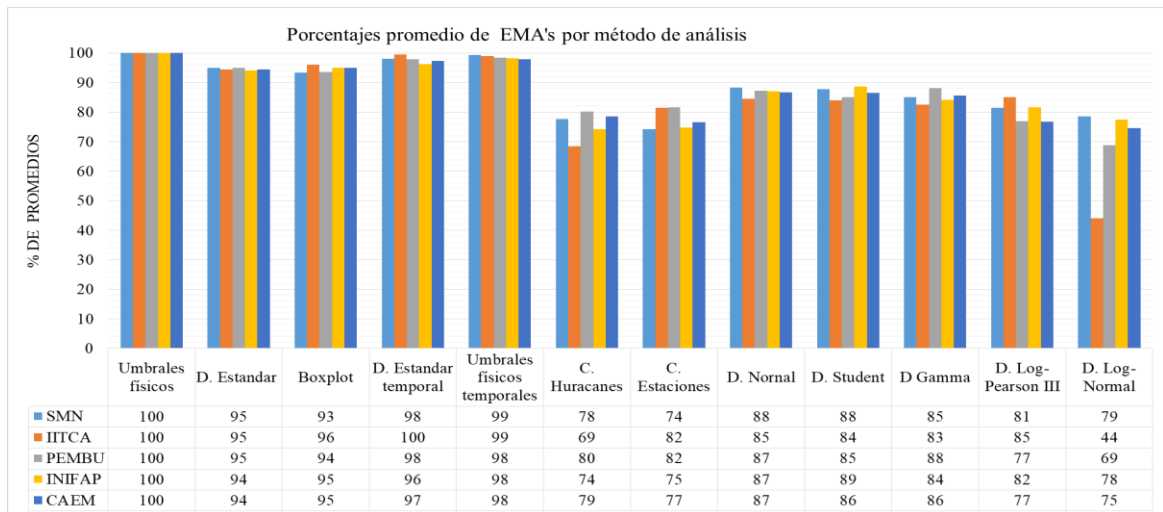


Figura 5.15 Porcentaje promedio de resultados de humedad relativa [Hr] para las EMA's

Tabla 5.10 Porcentaje de datos aprobados de las EMA's para la variable de dirección del viento [Dv]

Dependencia	Estadístico	Umbralés Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umbralés físicos	D. Estandar	Boxplot	D. Estandar temporal	Umbralés físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
SMN	Promedio	100	95	94	96	95	95	94	88	91	86	86	86
	Máximo	100	99	99	100	99	99	97	93	99	90	92	92
	Mínimo	100	90	90	91	91	90	90	82	82	81	81	81
	D. estd	0	3	3	3	3	3	3	4	8	3	4	5
IITCA	Promedio	100	95	96	95	95	98	95	91	90	83	85	91
	Máximo	100	98	98	99	98	99	98	92	92	84	86	93
	Mínimo	100	91	94	91	92	97	91	89	88	81	84	88
	D. estd	0	5	3	6	4	1	5	2	3	2	1	4
PEMBU	Promedio	100	95	94	96	95	95	94	88	87	88	87	88
	Máximo	100	99	99	100	100	99	99	93	92	93	93	93
	Mínimo	100	90	90	90	90	92	90	81	81	81	81	82
	D. estd	0	3	3	4	3	2	4	5	4	4	4	4
INIFAP	Promedio	100	95	95	96	94	93	94	87	90	88	86	86
	Máximo	100	98	99	100	100	99	99	94	99	93	93	93
	Mínimo	100	90	86	87	81	29	58	81	81	81	45	45
	D. estd	0	3	3	4	4	12	7	4	6	4	8	8
CAEM	Promedio	100	95	95	95	95	95	95	87	87	87	88	87
	Máximo	100	99	99	100	100	99	99	93	99	93	93	93
	Mínimo	100	90	90	90	90	90	90	81	81	81	81	81
	D. estd	0	3	3	3	4	3	3	4	5	3	4	4

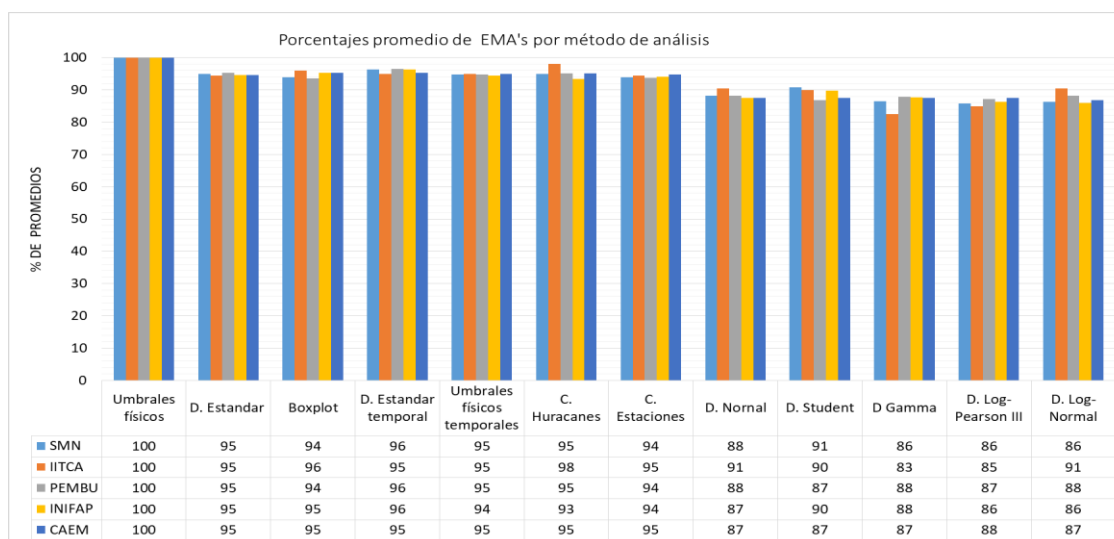


Figura 5.16 Porcentaje promedio de resultados de dirección del viento [Dv] para las EMA's

Tabla 5.11 Porcentaje de datos aprobados de las EMA's para la variable de velocidad del viento [Vv]

Dependencia	Estadístico	Umrales Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umrales físicos	D. Estandar	Boxplot	D. Estandar temporal	Umrales físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
SMN	Promedio	100	95	93	96	95	95	94	90	88	86	84	71
	Máximo	100	99	99	99	99	99	97	99	92	90	90	91
	Mínimo	100	90	90	91	91	90	90	82	82	81	80	0
	D. estd	0	3	3	3	3	3	3	7	5	3	3	32
IITCA	Promedio	100	95	96	95	95	98	95	85	84	83	85	44
	Máximo	100	98	98	99	98	99	98	89	88	84	86	88
	Mínimo	100	91	94	91	92	97	91	80	80	81	84	0
	D. estd	0	5	3	6	4	1	5	6	6	2	1	62
PEMBU	Promedio	100	95	94	96	95	95	94	88	85	89	88	65
	Máximo	100	99	99	99	99	99	99	99	91	99	99	91
	Mínimo	100	90	90	90	90	92	90	80	80	81	81	0
	D. estd	0	3	3	4	3	2	4	6	4	6	6	37
INIFAP	Promedio	100	94	95	94	94	93	94	88	89	86	86	77
	Máximo	100	98	99	99	99	99	99	99	99	99	99	91
	Mínimo	100	90	86	87	81	29	58	80	80	80	45	0
	D. estd	0	3	3	3	4	12	7	5	6	5	8	25
CAEM	Promedio	100	94	95	94	94	95	95	87	86	88	87	75
	Máximo	100	99	99	99	99	99	99	99	99	99	99	91
	Mínimo	100	90	90	90	90	90	90	80	80	80	80	0
	D. estd	0	3	3	3	3	3	3	5	4	5	6	28

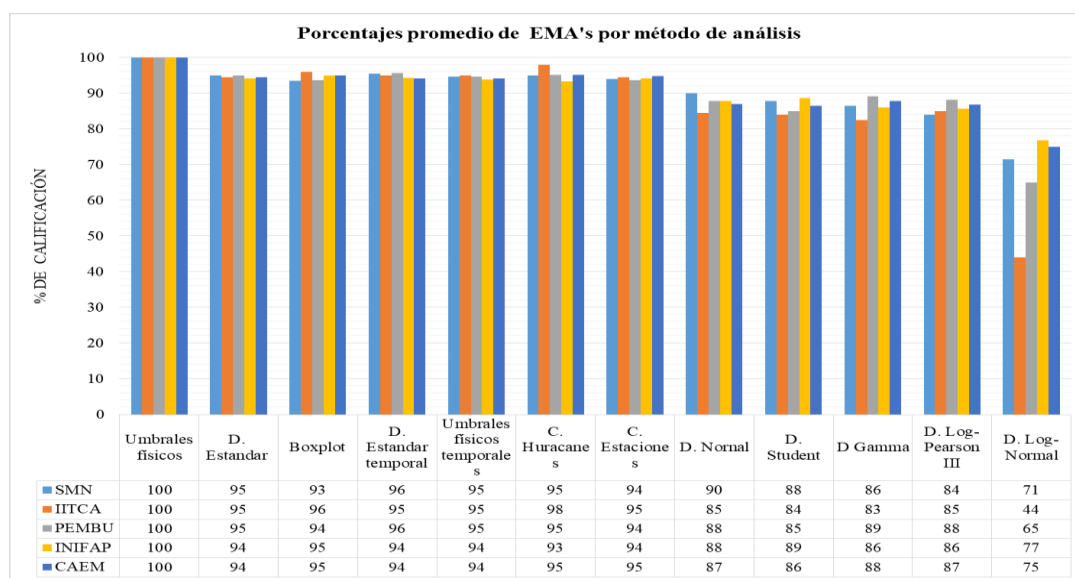


Figura. 5.17 Porcentaje promedio de resultados de velocidad del viento [Vv] para las EMA's

Tabla 5.12 Porcentaje de datos aprobados de las EMA's para la variable de presión atmosférica [Pb]

Dependencia	Estadístico	Umbrales Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umbrales físicos	D. Estandar	Boxplot	D. Estandar temporal	Umbrales físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
SMN	Promedio	100	94	95	95	92	94	94	84	87	84	88	88
	Máximo	100	99	98	99	96	97	98	92	92	87	90	93
	Mínimo	100	90	91	93	90	90	91	80	83	80	81	80
	D. estd	0	4	3	2	3	3	2	5	3	2	3	4
IITCA	Promedio	100	96	93	95	96	95	91	87	87	86	82	88
	Máximo	100	99	94	97	98	99	92	93	91	93	84	90
	Mínimo	100	95	92	91	95	91	90	83	84	82	81	85
	D. estd	0	2	1	3	2	4	1	5	4	6	2	3
PEMBU	Promedio	100	95	96	94	95	95	93	86	87	86	86	85
	Máximo	100	99	99	98	99	99	97	93	92	93	93	93
	Mínimo	100	90	91	90	90	92	90	80	80	80	80	80
	D. estd	0	3	2	3	4	2	2	5	4	4	4	5
INIFAP	Promedio	100	94	95	94	94	94	94	86	87	86	86	86
	Máximo	100	99	99	99	99	99	99	94	93	93	94	93
	Mínimo	100	90	86	82	81	29	39	58	80	80	45	45
	D. estd	0	3	3	3	3	9	9	6	4	4	7	8
CAEM	Promedio	100	94	95	94	94	95	95	91	91	91	92	91
	Máximo	100	99	99	99	99	99	99	99	99	99	99	99
	Mínimo	100	89	89	89	90	89	89	80	80	80	80	80
	D. estd	0	3	3	3	3	3	3	6	5	5	6	5

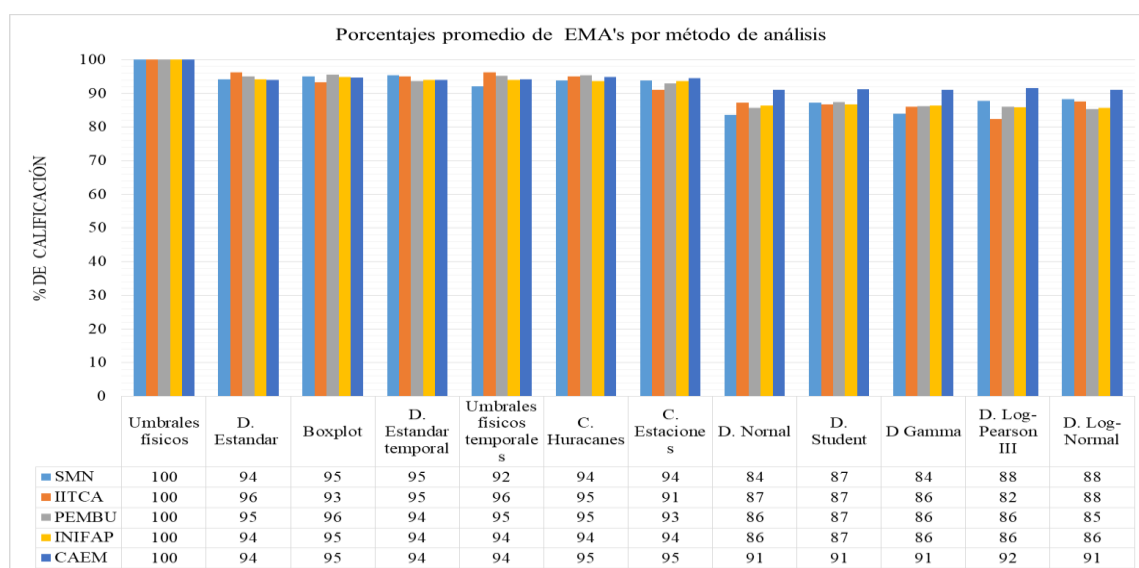


Figura 5.18 Porcentaje promedio de resultados de presión atmosférica [Pb] para las EMA's

Tabla 5.13 Porcentaje de datos aprobados de las EMA's para la variable de radiación solar [Rs]

Dependencia	Estadístico	Umbrales Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umbrales físicos	D. Estandar	Boxplot	D. Estandar temporal	Umbrales físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
SMN	Promedio	100	91	91	93	92	93	92	88	87	87	85	85
	Máximo	100	94	96	96	96	96	96	93	92	93	92	92
	Mínimo	100	89	89	89	89	90	89	82	82	81	81	81
	D. estd	0	3	2	3	2	2	3	4	5	4	4	4
PEMBU	Promedio	100	93	92	93	93	93	91	88	85	88	87	87
	Máximo	100	96	96	96	96	96	96	93	91	93	93	93
	Mínimo	100	90	90	89	90	89	89	81	81	81	81	82
	D. estd	0	2	2	3	2	2	2	5	4	4	4	4

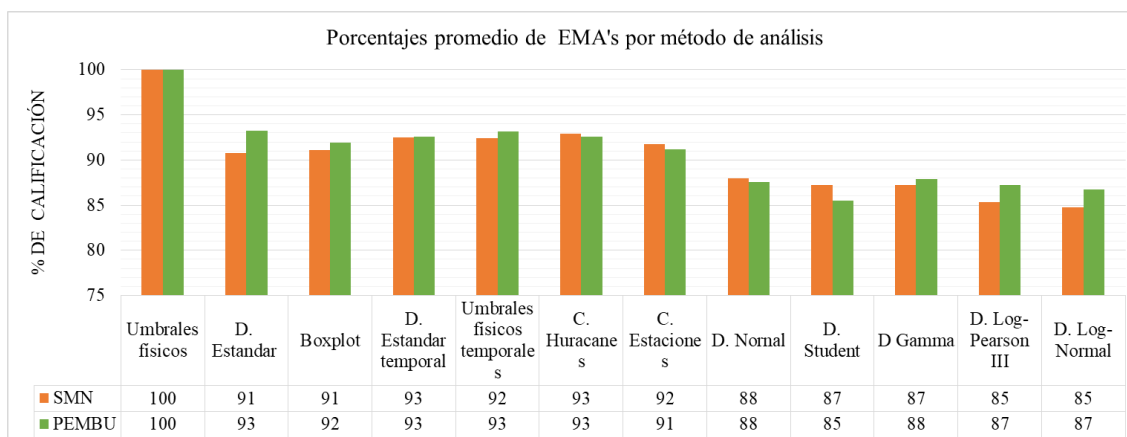


Figura 5.19 Porcentaje promedio de resultados de radiación solar [Rs] para las EMA's

Para este último caso, la variable de Radiación solar sólo algunas estaciones contienen registros. Estas estaciones son administradas por PEMBU y el SMN. Para todas estas estaciones se obtuvieron muy buenos porcentajes de datos aprobados al igual que las variables descritas anteriormente.

Como complemento a este análisis se crearon mapas con la ubicación de las EMA's analizadas y caracterizando cada una de las estaciones con el porcentaje obtenido para cada uno de los métodos de análisis. Un ejemplo de esta imagen se muestra en la figura 5.20, donde se observa el mapex del Estado de México con los porcentajes obtenidos para el método de desviación estándar aplicado a la variable de temperatura.

En el anexo 5.6 se encuentran todas las imágenes de los métodos aplicados a cada una de las variables de las EMA's para la distinción y ubicación de los resultados obtenidos en el análisis de calidad de datos de la red meteorológica de EMA's.

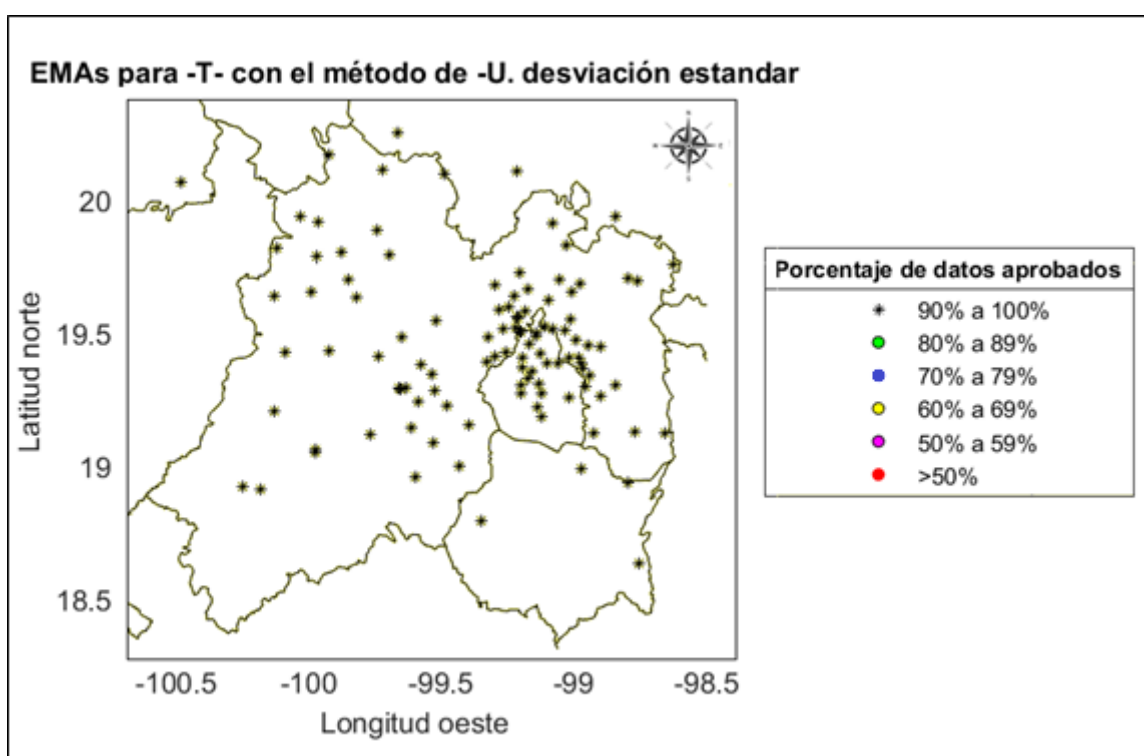


Figura 5.20 Mapex del Estado de México con porcentajes obtenidos para el método de desviación estándar aplicado a la variable de temperatura de las EMA's

5.3.2 Base de datos de las EMC's

Las EMC's del Estado de México, son administradas por CONAGUA y disponibles en la base de datos de CLICOM, de donde se obtuvo la base de datos que se analizó. Siendo un total de 643 estaciones con registros diarios para las variables de temperatura máxima [Tmáx], temperatura mínima [Tmín], precipitación [P] y evaporación [Ev]. Sin embargo, cabe destacar que un gran porcentaje (29.9%) de estas estaciones se encuentran actualmente suspendidas.

Con base en el análisis de calidad de datos, se encontró que la mayoría de las estaciones aún estando suspendidas cuentan con una base de datos con altos porcentajes de calificación. Estos resultados se muestran en el anexo 5.7, donde se aprecian los porcentajes de datos que aprobaron el método por cada estación y variable analizada.

Los resultados obtenidos son muy similares a los de las EMA's ya que presentan un descenso en los datos aprobados para los métodos de homogeneidad. Estos resultados se pueden apreciar en la tabla 5.14 y figura 5.21, donde se muestran los porcentajes máximos, mínimos y promedio por cada método aplicado.

Tabla 5.14 Porcentaje de datos aprobado de la base de datos CLICOM

Dependencia	Estadístico	Umbralés Físicos			Coherencia interna		Coherencia espacial		Homogeneidad				
		Umbralés físicos	D. Estandar	Boxplot	D. Estandar temporal	Umbralés físicos temporales	C. Huracanes	C. Estaciones	D. Normal	D. Student	D Gamma	D. Log-Pearson III	D. Log-Normal
TMÁX	Promedio	100	96	95	95	99	95	94	84	84	85	85	84
	Máximo	100	100	100	100	100	100	100	98	91	98	93	97
	Mínimo	100	91	81	73	76	36	58	55	22	45	50	33
	D. estd	0	3	2	2	2	4	4	4	4	5	6	10
TMÍN	Promedio	100	94	94	96	94	94	94	86	86	84	81	86
	Máximo	100	99	99	100	100	99	99	96	97	92	91	99
	Mínimo	100	85	81	67	15	70	63	21	22	20	31	0
	D. estd	0	3	3	3	5	3	4	5	6	7	11	7
P	Promedio	100	52	53	94	94	93	94	85	85	82	85	61
	Máximo	100	107	97	100	100	99	99	98	99	97	100	99
	Mínimo	100	19	19	16	19	29	58	0	18	0	0	0
	D. estd	0	24	25	11	32	12	7	10	11	13	9	38
EV	Promedio		95	95	97	96	93	94	87	87	88	87	86
	Máximo		99	99	100	100	99	99	100	98	98	98	99
	Mínimo		87	81	49	51	29	58	19	22	59	53	0
	D. estd		2	2	3	3	12	7	32	6	5	5	11

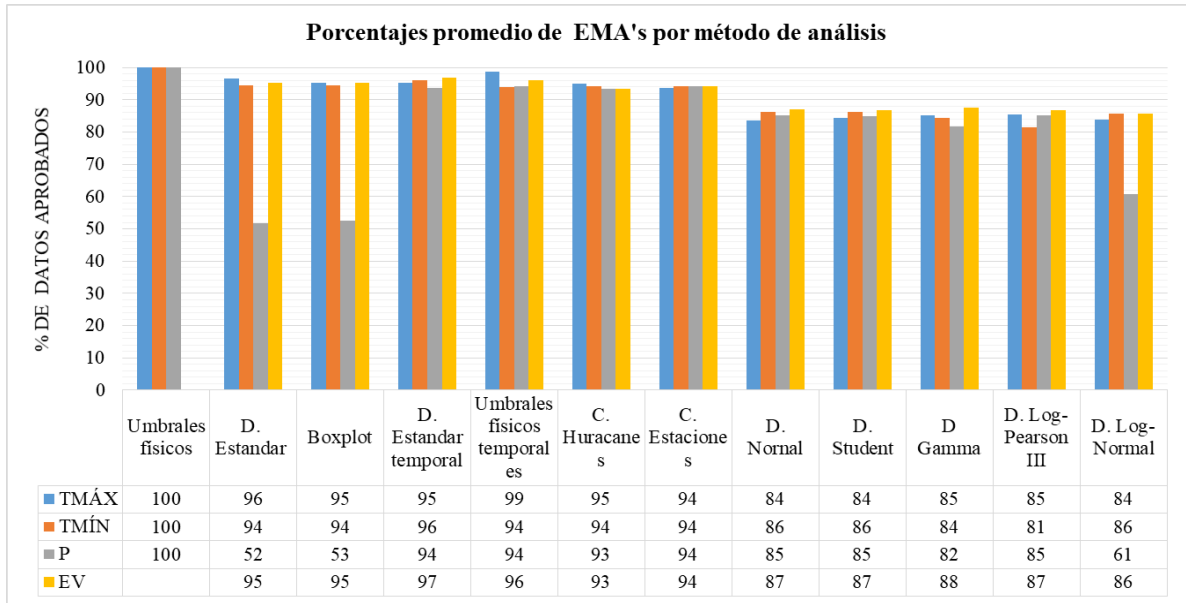


Figura 5.21 Gráfica de los porcentajes promedio de los valores aprobados en la base de datos CLICOM

Al igual que la red meteorológica de las EMA's, se generaron mapas de caracterización de las EMC's del mapex del Estado de México. Donde se muestra el mapex del Estado de México con los porcentajes obtenidos para cada uno de los métodos aplicados y estaciones de la red de EMC's. Por ejemplo, en la figura 5.22 se muestra la imagen de la caracterización de las estaciones para el método de desviación estándar aplicado a la variable de Tmáx. En el anexo 5.8 se encuentran todas las imágenes de los métodos aplicados a cada una de las variables de las EMC's del mapex del Estado de México para las variables de Tmáx, Tmín, P y Ev.

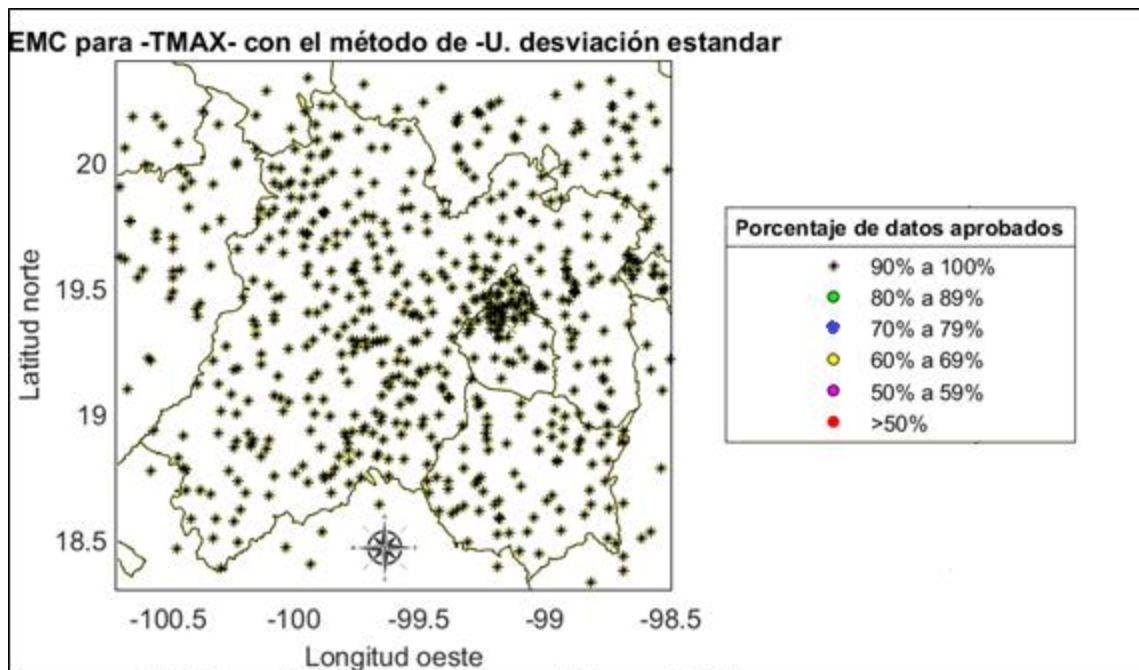


Figura 5.22 Mapex del Estado de México con porcentajes obtenidos para el método de desviación estándar aplicado a la variable de T_{máx} de las EMC's.

5.3.3 Interpolación de datos diarios.

Para realizar la interpolación de los datos se tomaron en cuenta los registros diarios de T_{máx}, T_{mín} y P. Obteniendo los máximos y mínimos diarios de las de las EMA's para las variables de T_{máx} y T_{mín} y la sumatoria de los registros semi horarios de un día para la variable de P. A partir de ello, se realizó la interpolación de los datos, preparando los datos de entrada para el módulo de interpolación Hidra (Arévalo. M, 2019). Clasificando estos registros en archivos *.txt para los datos históricos diarios como se muestra en la figura 5.23. Donde se muestran los archivos de la variable de T_{máx} para el día juliano 228, correspondiente al 15 de agosto. Apreciando al lado derecho de la imagen la estructura de los archivos.

Nombre	StationID	Longitud	Latitud	Altitud	TMAX	
	9001	492215.00000000	2149454.00000000	2596	NaN	
	9002	478977.00000000	2124813.00000000	2900	NaN	
2000-228-TMAX	9003	480059.00000000	2152783.00000000	2255	NaN	
2001-228-TMAX	9004	484319.00000000	2123302.00000000	2726	NaN	
	9005	477249.00000000	2143258.00000000	2326	NaN	
2002-228-TMAX	9006	485420.00000000	2149058.00000000	2247	NaN	
	9007	487752.00000000	2146936.00000000	2326	NaN	
2003-228-TMAX	9008	480742.00000000	2137168.00000000	2278	NaN	
	9009	492125.00000000	2145028.00000000	2235	NaN	
2004-228-TMAX	9010	478827.00000000	2146483.00000000	2271	26.00	

Figura 5.23 Ejemplo de archivos de entrada para la interpolación

Para realizar la interpolación se utilizó el método Kriging con cambio externo, el cual asume que existen condiciones externas que interfieren en la variable a estimar. Estas condiciones externas que afectan la variabilidad meteorológica es la altitud. Por lo cual, para la interpolación se requirió de un DEM (Digital elevation models) del área de estudio como se muestra en la figura 5.24, el cual muestra también la ubicación de las estaciones consideradas en la interpolación.

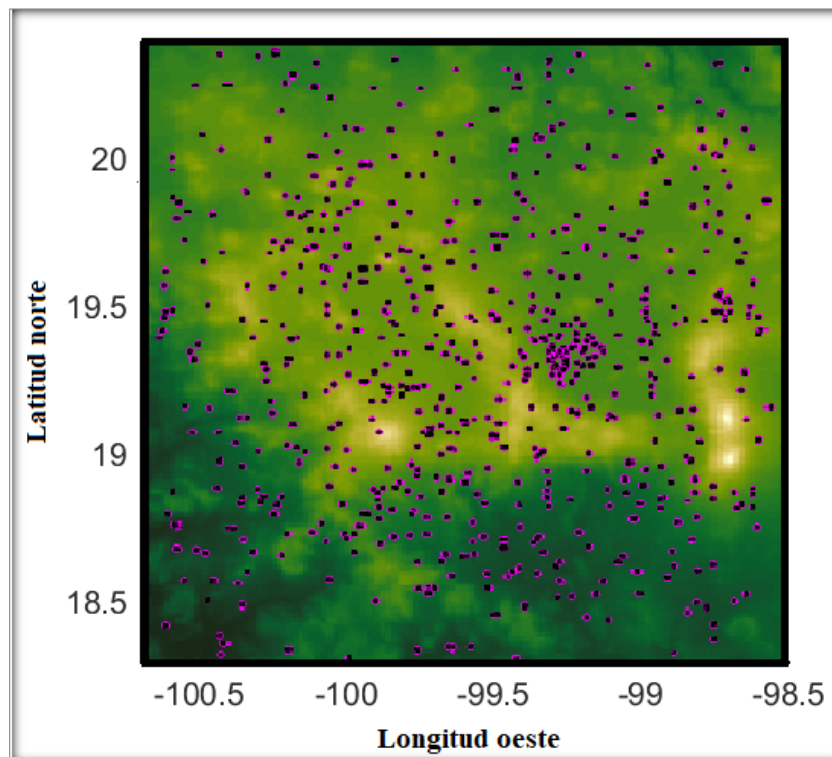


Figura 5.24 Mapa digital de elevaciones del mapex del Estado de México con la ubicación de las estaciones seleccionadas para la interpolación.

A partir de los datos de las estaciones meteorológicas y del DEM se realizó la interpolación de los datos. El módulo realiza internamente los procesos de interpolación y regresiones, generando un variograma de ajuste con los datos históricos del día. Este variograma permite que a pesar de tener pocas estaciones con registros para un día en particular, la interpolación se realice con los antecedentes de los datos históricos.

Los resultados de la interpolación de los datos diarios de $T_{\text{máx}}$ y $T_{\text{mín}}$ para el periodo del 2000 al 2016, se muestran en los archivos resultantes en el anexo 5.9. Donde se tienen dos carpetas nombradas con el día juliano correspondiente (por ejemplo 1 y 1-r). Una contiene los datos de entrada (la carpeta con sólo el número) y en la otra los resultados de la interpolación (la carpeta con el anexo “-r”). La carpeta de resultados contiene los datos de la regresión, del variograma y de la interpolación.

Como resultado principal de estas interpolaciones se obtienen imágenes espaciales de los resultados, por ejemplo en la figura 5.25 y 5.27, se tienen los resultados para las variables de $T_{\text{máx}}$ y $T_{\text{mín}}$ para el día 15 de agosto del 2016, así como las estaciones que se consideraron en el análisis. En las figuras 5.26 y 5.28 los variogramas correspondiente al día con el valor del coeficiente de ajuste R^2 el cual proporciona la medida que define el porcentaje de varianza explicado por la regresión.

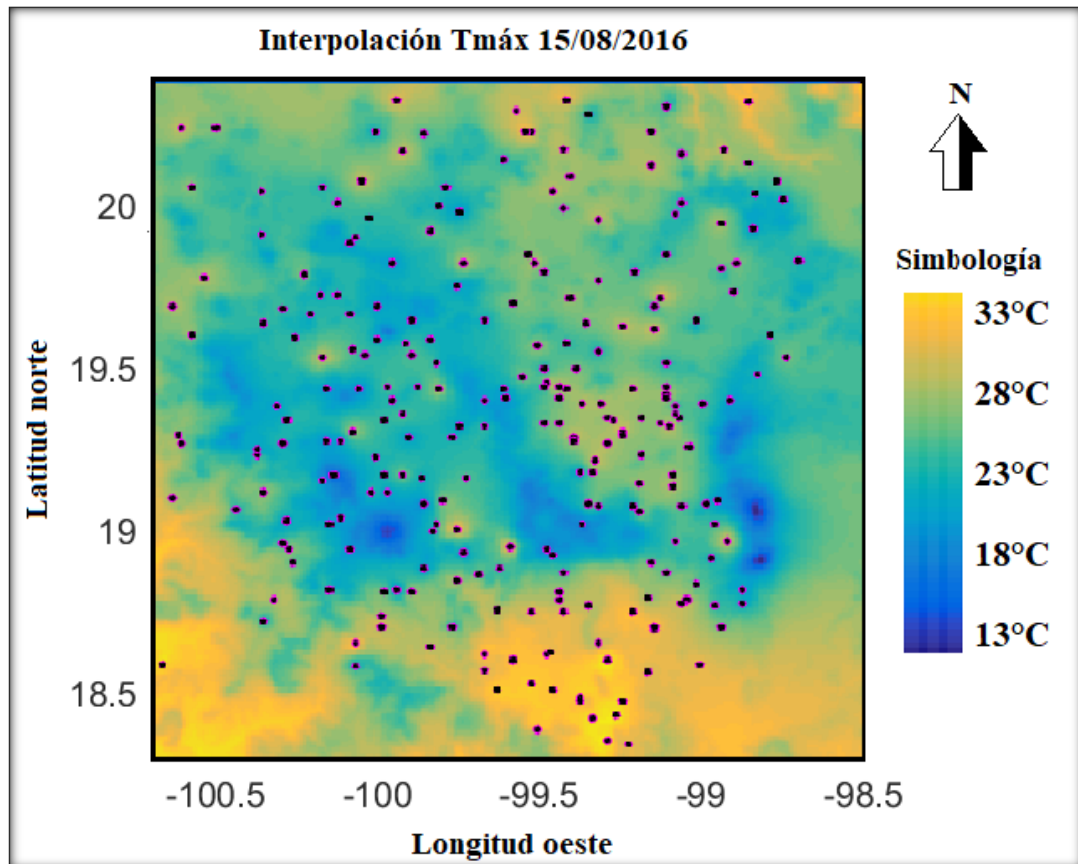


Figura 5.25 Imagen de interpolación de Tmáx para el día 15 de agosto del 2016

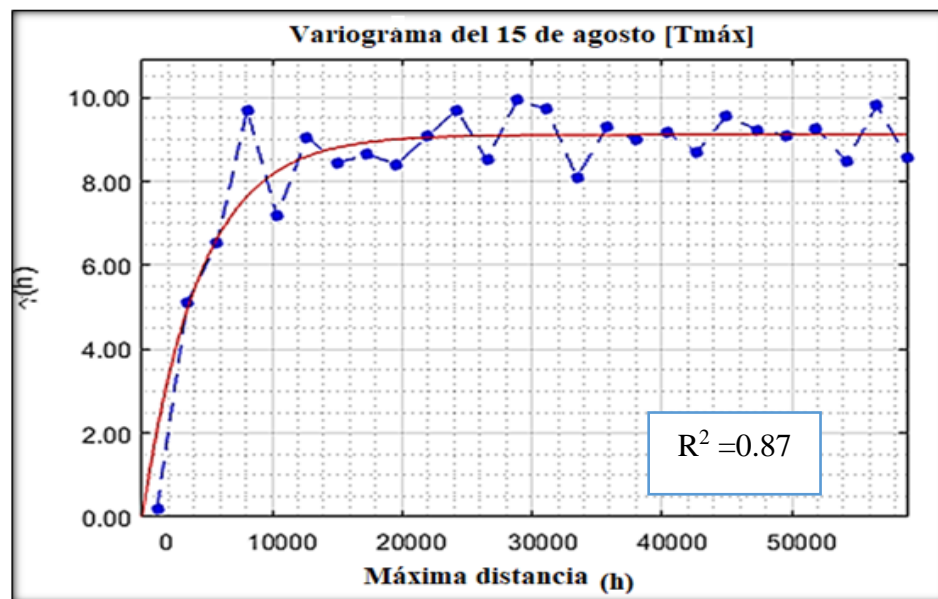


Figura 5.26 Variograma del día 15 de agosto para Tmáx

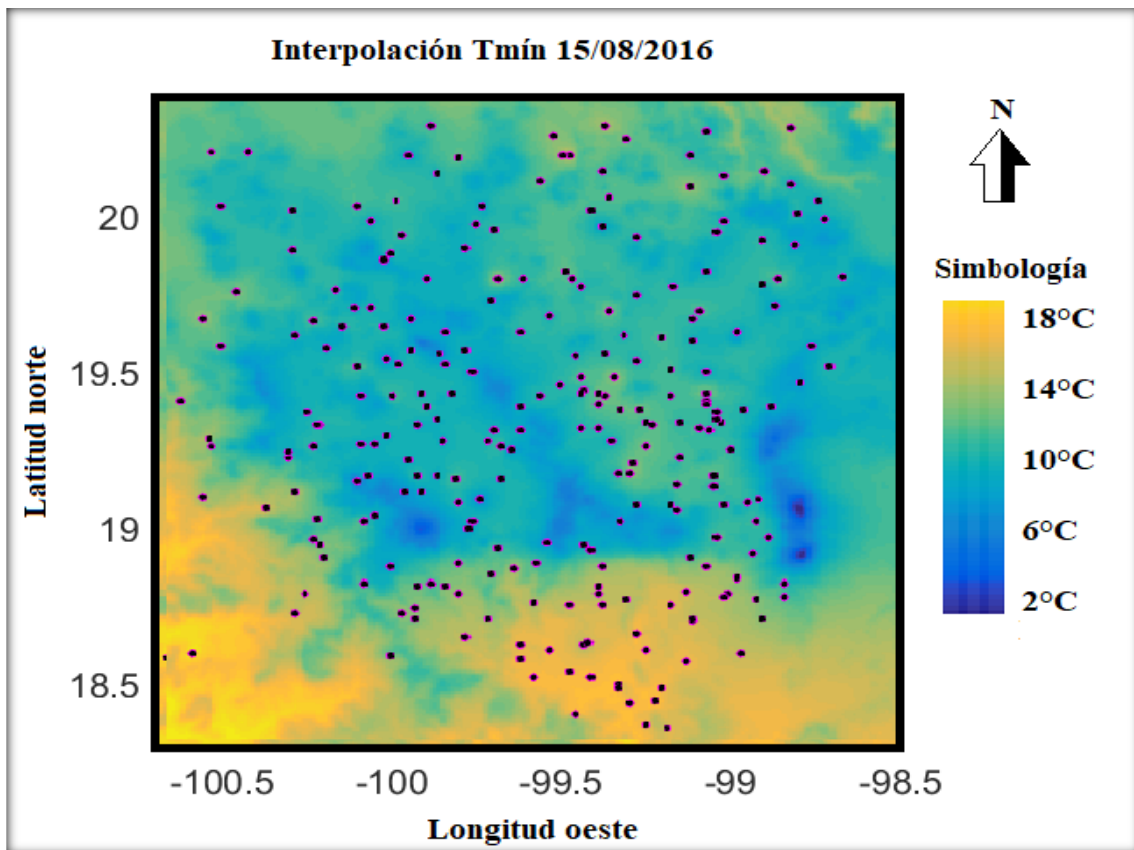


Figura 5.27 Imagen de interpolación de Tmín para el día 15 de agosto del 2016

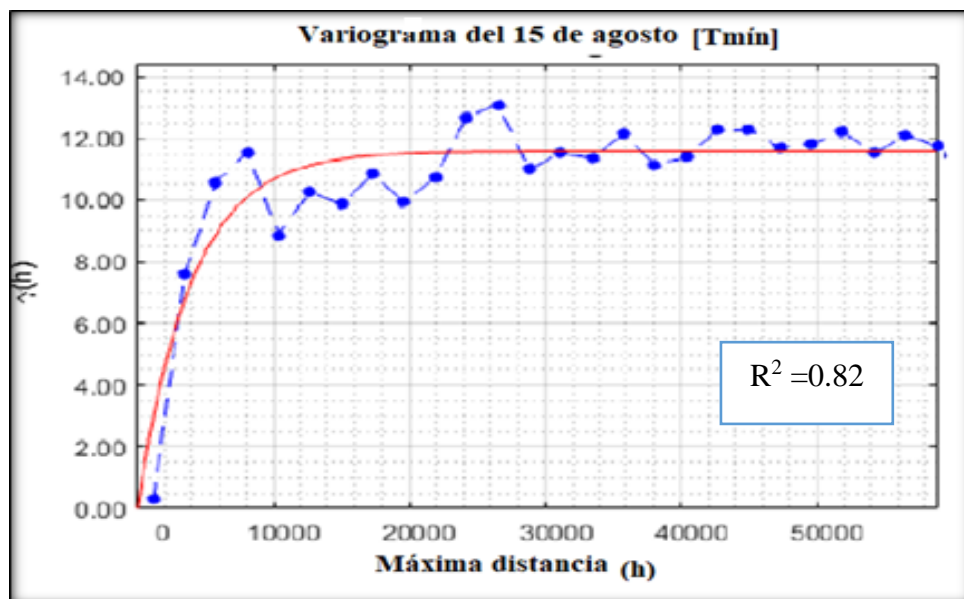


Figura 5.28 Variograma del día 15 de agosto Tmín

Al realizar todas las interpolaciones, se registraron R^2 de ajuste al variograma y de la regresión para todos los días del año y para cada una de las variables interpoladas. Estos resultados se muestran en la figura 5.29, Donde se comparan las gráficas de R^2 de $T_{\text{máx}}$ y $T_{\text{mín}}$. En ambas gráficas se aprecia mayor separación al inicio y final, lo cual corresponde a los días de los meses de enero, febrero, marzo, noviembre y diciembre. Esto se debe a que en esas fechas existe mayor variabilidad en la temperatura de las zonas.

Se puede apreciar que en su mayoría los ajustes a los variogramas tienen un R^2 con rango de 0.8 a 0.95 aproximadamente. Esto quiere decir que las interpolaciones son adecuadas y muy acercadas a la realidad. Ya que de acuerdo con el valor de R^2 entre más se acerque a 1, el ajuste será mejor lo cual garantiza la interpolación más real.

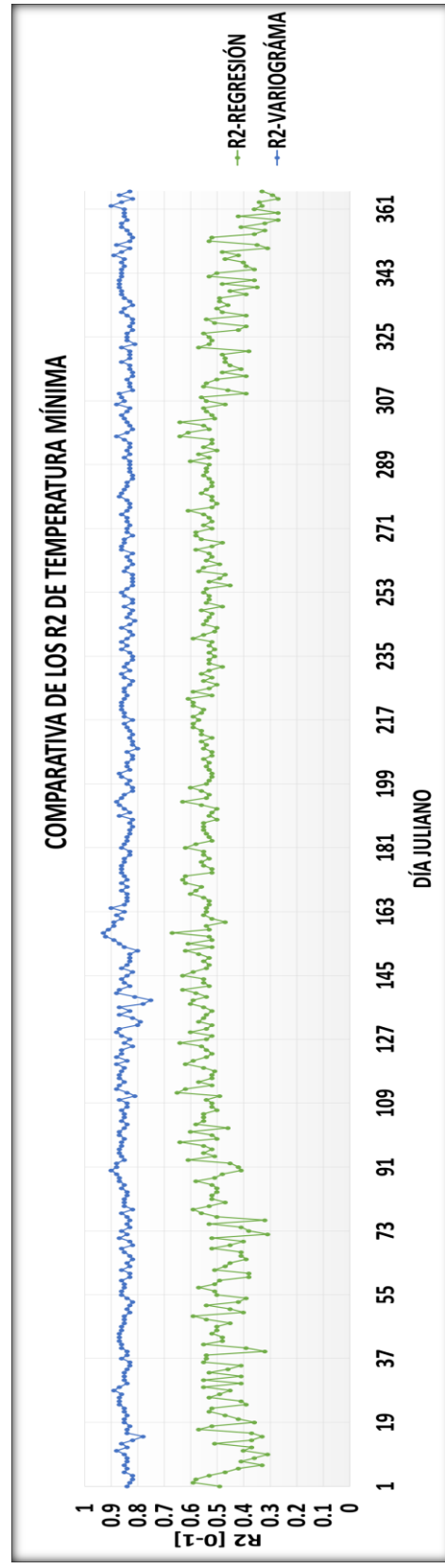
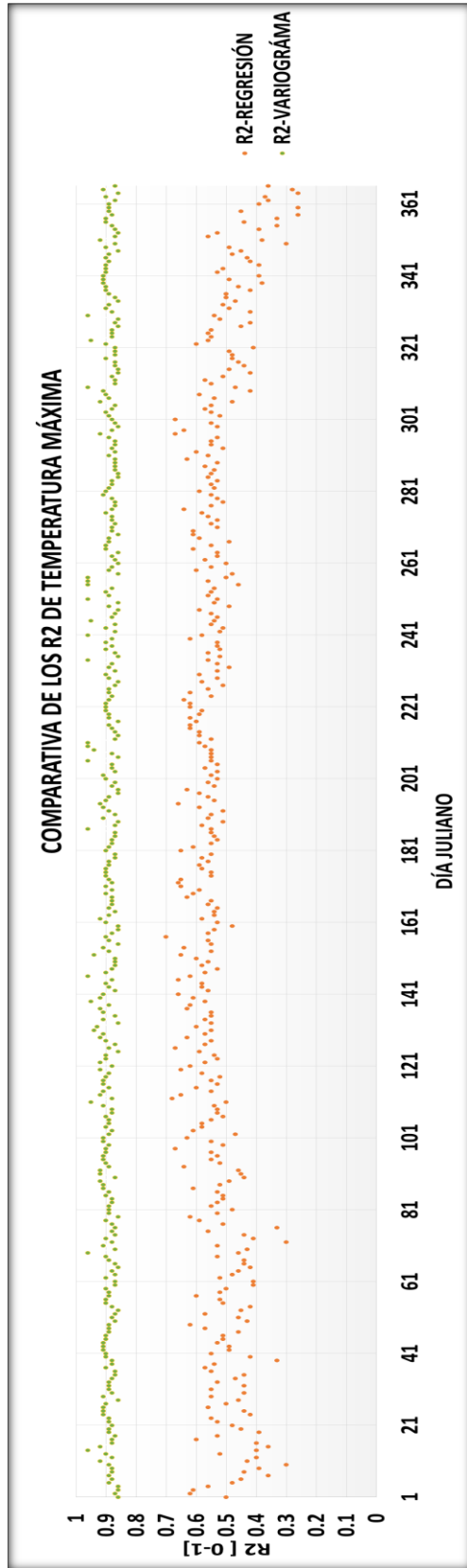


Figura 5.29 Gráficas de R^2 de ajuste al variograma y regresión de las variables de $T_{máx}$ y $T_{mín}$.

Para la interpolación de la variable de precipitación, se encontraron algunas particularidades. Debido a la topografía irregular de la zona de estudio, el tipo de precipitación que se presenta es de tipo de lluvias convectivas. Por lo cual, a realizar una interpolación para periodos de lluvias como los meses de junio, julio, agosto, septiembre, las interpolaciones muestran valores negativos. Esto quiere decir que el método no es adecuado para interpolar esta variable.

Un ejemplo de esta interpolación se muestra en la figura 5.30 y 5.31, donde se interpolaron los registros del 15 de agosto del 2016 y 1 de julio del 2016 respectivamente. En ambas imágenes se aprecian los valores negativos como valores resultantes de la interpolación.

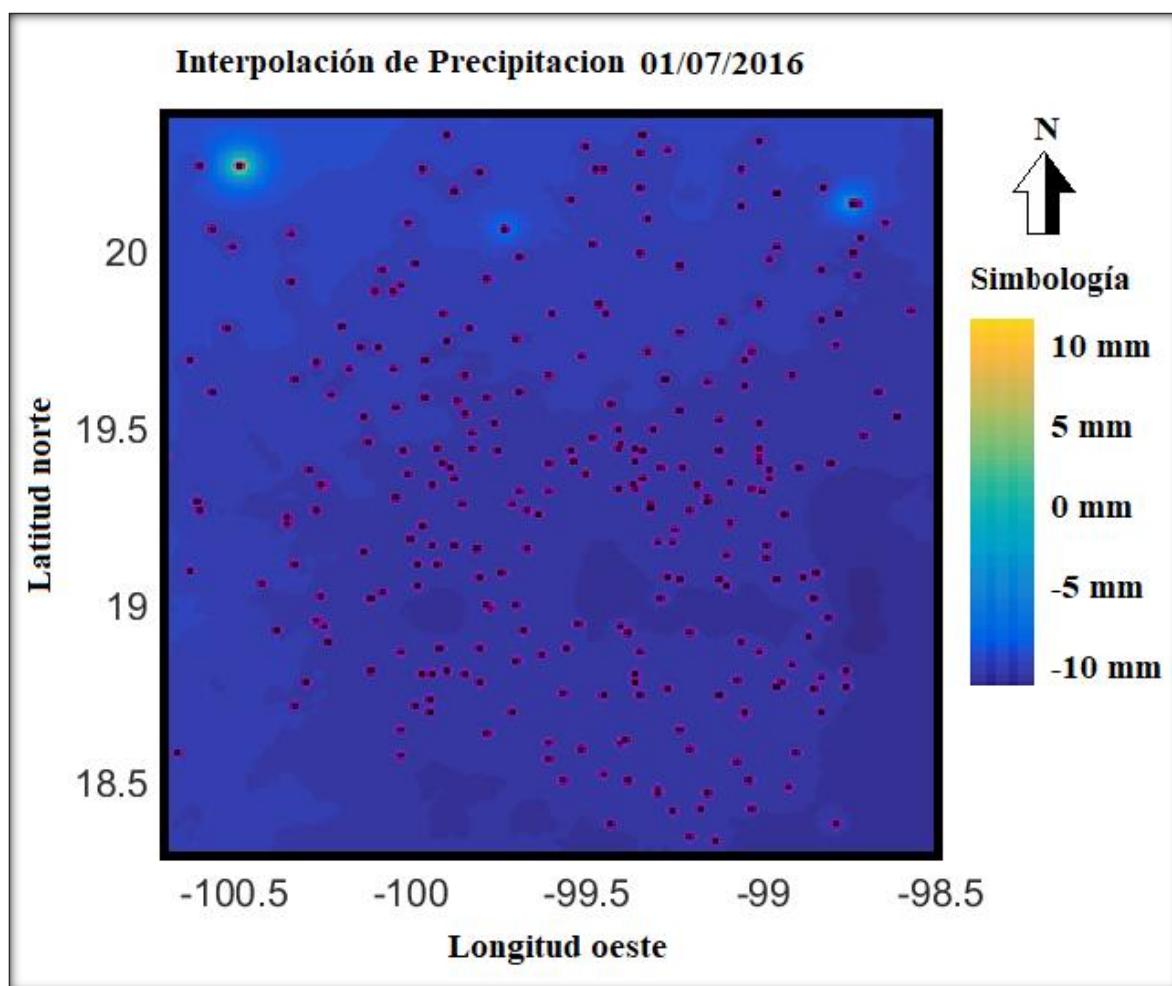


Figura 5.30 Ejemplo de interpolación de precipitación 15/08/2016

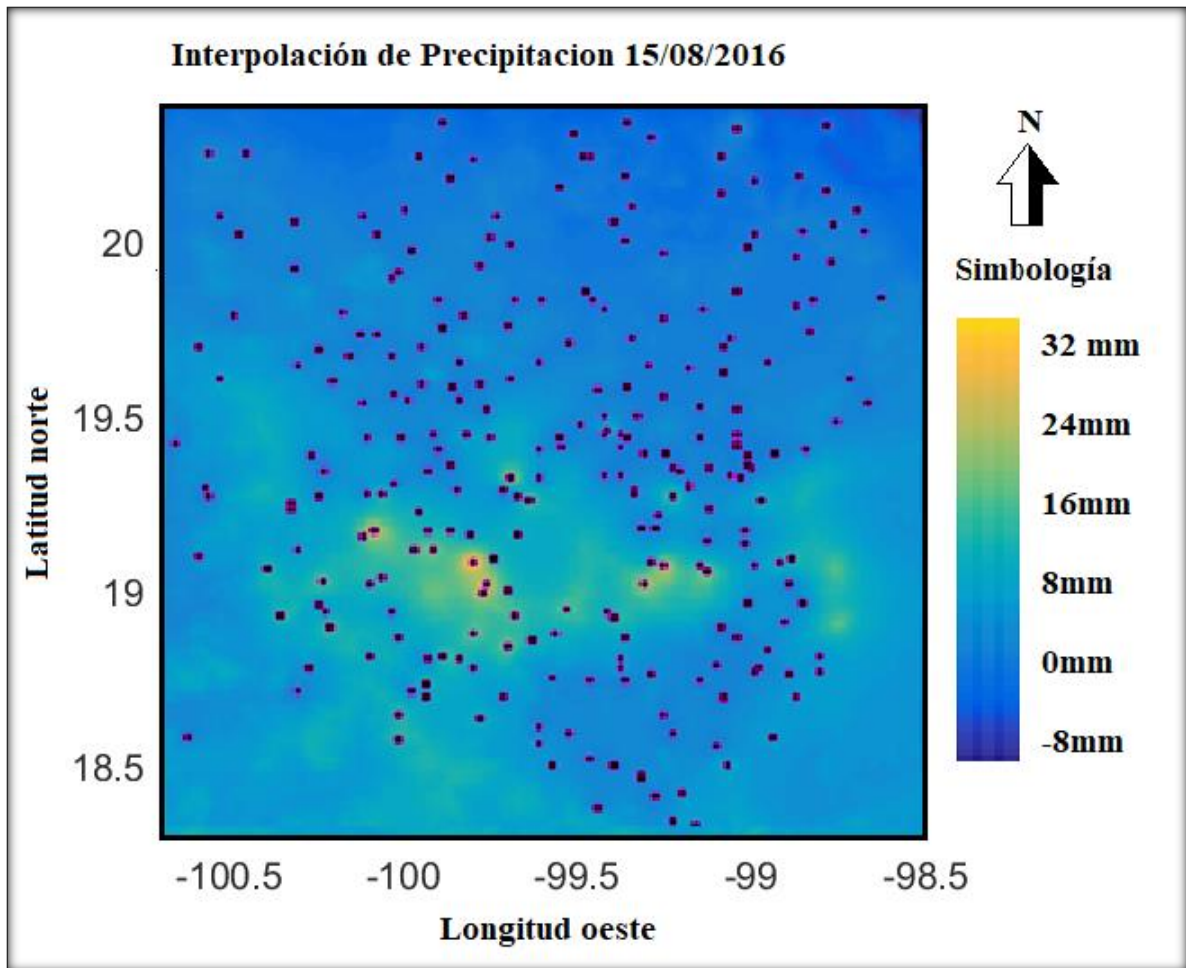


Figura 5.31 Ejemplo de interpolación de precipitación 01/07/2016

Por ello, para realizar una buena interpolación de precipitación diaria, es necesario recurrir a métodos más sofisticados como la simulación condicional. Esta simulación es una función aleatoria de distribución espacial condicionada por variables externas. Estas variables forman campos aleatorios que figuran la dispersión de los datos uniformes (Leblois & Creutin, 2013). Como ejemplo representativo de cómo realizar este proceso, se realizó una simulación en la herramienta hidro-informática “HIDRA-Eflow “ (Arévalo. M, 2019). El cual consiste en la generación de campos gaussianos en tres dimensiones para generar campos advectivos de precipitación intermitente con periodos de tiempo corto. Como resultado a esta simulación se obtienen los resultados del anexo 5.10 para el día 1 de Julio del 2016, obteniendo una imagen representativa mostrada en la figura 5.32.

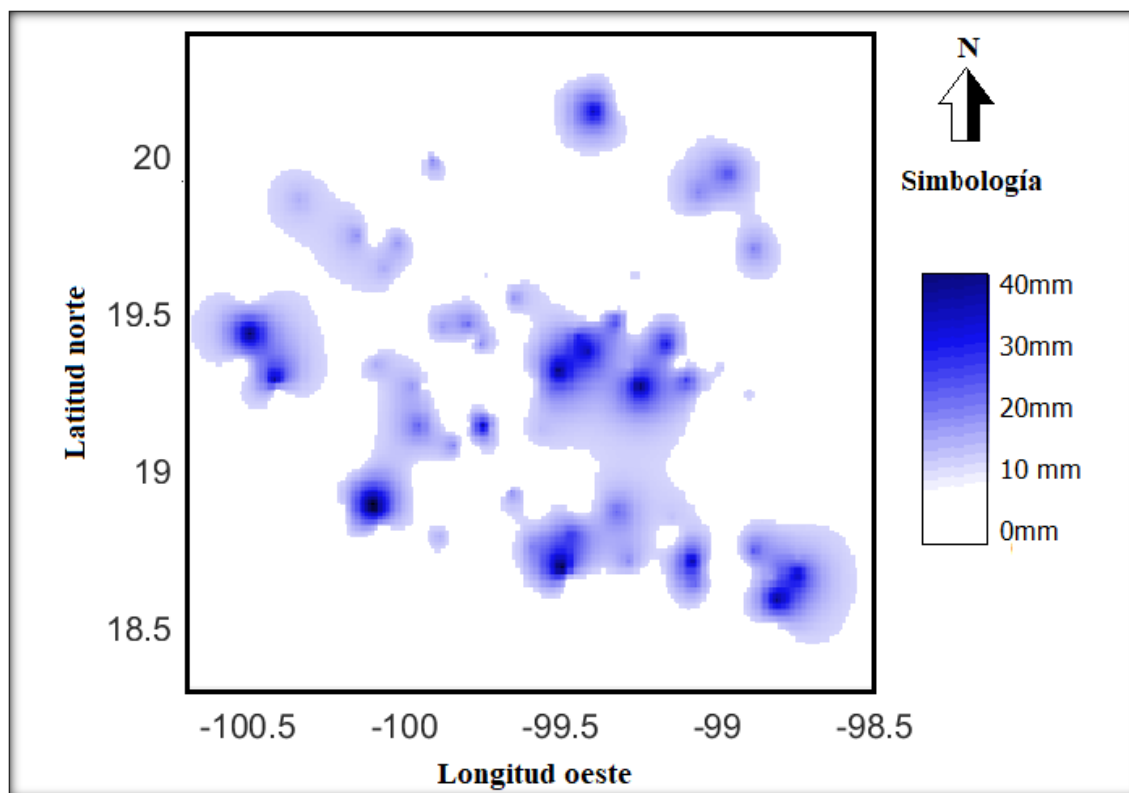


Figura 5.32 Resultados de la simulación del día 01/07/2016

5.4 Discusión de resultados

La herramienta hidro informática “SICCDAMET” tiene una interfaz amigable y de fácil uso. Llevando de la mano al usuario para generar correctamente el análisis de calidad de cada método y visualizando en gráficas básicas los resultados. Donde estas últimas ayudan al usuario a identificar rápidamente si la estación o los registros generan información de calidad. Dándole fundamentos para reducir la incertidumbre en la toma de decisiones al utilizar la información meteorológica.

En contraste con otras herramientas como RCLimTool (Llanos, 2015), SICCDAMET es una herramienta que recopila las metodologías de análisis de calidad más recientes y eficaces encontradas en bibliografía especializada, así como manuales y normas de la OMM proporcionando al usuario 16 métodos de análisis de calidad que le permiten identificar rápidamente los registros dudosos. SICCDAMET también cuenta con mayor accesibilidad

para los formatos de entrada, ya que pueden presentar distinto orden de las variables meteorológicas siempre y cuando cumplan con el formato de fecha específico que requiere la herramienta.

Con base en la aplicación de la metodología de análisis aquí propuesta a la base de datos de las estaciones que se encuentran en el mapex del Estado de México, se encontró que las estaciones generan información de calidad. Estos resultados concuerdan con estudios anteriores efectuados por Salgado (2018), donde las calificaciones más bajas se encontraban en la revisión de los métodos de homogeneidad.

La distribución Normal, resultó ser la más adecuada para el análisis de los registros diarios, principalmente para variables de temperatura y ello concuerda con los resultados de estudios similares tales como los de Aguilar (Aguilar, 2013). Mientras que la distribución Gamma, por sus características, resulta ser un modelo adecuado y genera mejores resultados para variables con registros no negativos como P, Hr, Ev, Rs y Pb (Štěpánek et al., 2013). Los métodos de umbrales físicos y coherencia interna se mantenían con calificaciones altas para todas las variables meteorológicas, ya que su revisión se mantiene en umbrales obtenidos de las variables, lo cual indica que los registros entre sí mantienen una coherencia (Dudás, 2007; Shulski et al., 2014)

Para los métodos de interpolación, el método geoestadístico de Kriging con covariable externa es adecuado para variables con un comportamiento con poca variabilidad. En este caso, las variables son T_{máx} y T_{mín.}, en las cuales puede influir la condición externa como la altitud (covariable). Estas variables meteorológicas interpoladas obtuvieron buenos resultados para el periodo del 2000 al 2016, generándose imágenes diarias de la variable por interpolación con alto grado de confiabilidad.

Para el caso de la variable de P, el método de Kriging no fue adecuado debido al comportamiento que presenta. Esto quiere decir que para variables meteorológicas con mayor inestabilidad espaciotemporal no es adecuado este tipo de interpolación, por lo cual se requieren métodos geoestadísticos más sofisticados como lo es la simulación condicional (Leblois & Creutin, 2013)

Capítulo 6

Conclusiones y recomendaciones

A nivel mundial la OMM proporciona guías, manuales y normas que en las cuales reflejan muy poca información para el análisis de calidad de los datos meteorológicos. Por ello, diversos países han venido realizando investigación sobre la búsqueda de los métodos óptimos para determinar la calidad de estos datos.

6.1 Conclusiones generales

Con base en los objetivos de este proyecto, se desarrolló una herramienta hidro-informática que realiza el análisis y control de calidad, validación y complemento de series de datos meteorológicos de las estaciones automáticas y convencionales, que genera reportes de resultados y gráficas que ayudan al usuario a determinar la confiabilidad de la información que éstas transmiten y almacenan, para obtener mayor certeza en la toma de decisiones.

Todos los métodos estadísticos encontrados en bibliografía especializada se evaluaron y seleccionaron para incluirlos en la herramienta. Este conjunto de métodos fue probado en la base de datos de la zona de estudio. Para lo cual se recopiló y estructuró las bases de datos en un formato único para llevar a cabo el análisis de la calidad de datos, cumpliendo con los objetivos establecidos en el presente proyecto.

Herramienta hidro-informática “SICCDAMET”

La herramienta hidro-informática “SICCDAMET”, genera rápidamente un análisis de calidad y complemento de datos. La herramienta proporciona resultados y reportes que ayudan al usuario en la toma de decisiones, principalmente para la aplicación de la variable meteorológica en actividades como la agricultura, meteorología, climatología, etc. También proporciona una base de datos organizada en un formato *.xlsx de fácil manipulación y en diferente temporalidad, ya sean datos semi horarios (sólo para EMA’s y ESIME’s), diarios, mensuales, o bien anuales.

Esta herramienta en comparación con otras en su tipo presenta una metodología que recopila los distintos métodos específicos en el tema encontrados en bibliografía especializada. SICCDAMET toma en cuenta 16 métodos para la revisión de calidad de la información, distribuidos para la revisión de umbrales físicos, coherencia interna, coherencia espacial, homogeneidad y complemento de datos. Los resultados son presentados al usuario a través de gráficas y reportes para facilitar la visualización del comportamiento de los fenómenos meteorológicos bajo estudio. Estas características de la herramienta brindan al usuario una un soporte eficaz para poder tomar decisiones sobre la información y los fenómenos meteorológicos correspondientes.

Una de las ventajas que presenta SICCDAMET, se encuentra en los formatos de entrada, ya que otras herramientas se limitan a su propio formato de entrada. Esta herramienta, permite la entrada de archivos *.csv o *.xlsx con un formato similar al que genera la estación meteorológica, sólo cumpliendo los requisitos para el formato de la fecha de registro. Esto ayuda al usuario a organizar rápidamente los archivos y se encuentre en medida de generar el análisis de calidad de datos. Además de ello, dentro de la herramienta se generan los distintos formatos de entrada para los módulos anexos que permiten interpolar los datos, así como la creación de curvas IDTr. Esto facilita la generación de información a los usuarios.

Métodos de análisis aplicados

Que una estación no apruebe los métodos aplicados sobre la calidad de datos registrados, se debe a varios factores. Uno de ellos es que los métodos de análisis tienen límites de aplicación que se encuentran por debajo del límite máximo del sensor. Lo cual, si la variable se encuentra fuera de ese rango lo marca como malo, aunque existe la probabilidad de que este dato ocurra. Otro de los factores es la presencia de microclimas que tiene el Estado de México y donde el comportamiento de las variables meteorológicas rompe con el patrón regional esperado. En general, los métodos aplicados resultaron adecuados para las variables meteorológicas, y esto está soportado por los resultados obtenidos al analizar la base de datos de la zona de estudio. El método de Kriging con cambios externos, resultó ser adecuado únicamente para variables de temperatura. En efecto, las interpolaciones realizadas representan adecuadamente la información y es lo más próximo a la realidad. Caso contrario el que ocurrió para la variable de precipitación, donde la interpolación realizada con este método genera información que no representa la realidad del comportamiento de esta variable. Este último hallazgo, inesperado y desafiante, sugiere y abre la puerta a explorar el empleo de métodos de simulación condicional para la modelación espacial de variables como la precipitación en el contexto particular de la zona en estudio.

Base de datos meteorológicos

De las estaciones que conforman el mapex del Estado de México contempladas en este proyecto, se logró detectar que cada una de las dependencias que las administra genera un formato de registro distinto. Incluso se encontró que, en algunos casos, para una misma estación meteorológica el registro de los datos cambia. A pesar de las distintas normas y guías que se tienen, no se respeta la estructura de almacenamiento de los datos para que el usuario manipule esta información fácilmente. Cabe resaltar que los formatos para las EMC's de CONAGUA sí cumplen con un formato idéntico para todas las estaciones. Este formato es conocido como formato CLICOM, el cual sólo proporciona un ID para distinguir las variables meteorológicas.

El no tener un formato único de registro, principalmente en las EMA's y ESIME's presenta retrasos importantes en el desarrollo de proyectos. Esto se debe a que primero se tiene que ordenar la base de datos y después manipularla para obtener la información deseada. El tener

los datos unificados bajo un sólo formato, fue uno de los procesos más tardados en este proyecto y es una de las etapas primordiales a realizar en la herramienta puesto que ello facilita la aplicación de los métodos de análisis.

Al aplicar la metodología de la herramienta a la base de datos de las estaciones del mapex del Estado de México, se encontró que todas las estaciones tienen un buen funcionamiento al aprobar todos los registros el método de umbrales del sensor. Este comportamiento fue similar para todas las variables en los métodos de umbrales físicos, coherencia interna y coherencia espacial. Para el caso de los métodos de homogeneidad, los resultados obtenidos variaron de acuerdo a la distribución de probabilidad y la variable meteorológica analizada. Esto se debe al comportamiento de las variables y su ajuste representativo con las distribuciones de probabilidad.

6.2 Recomendaciones

Durante el desarrollo del presente trabajo se identificaron diversas vías para investigaciones futuras con la finalidad de enriquecer el contenido de la base de datos en distintos aspectos como la generación de aplicaciones móviles que analicen las variables meteorológicas en tiempo real obteniendo información tanto satelital como de estaciones cercanas a la zona. Esto para que la metodología de análisis siga funcionando, sea más accesible para los usuarios y se aprovechen los avances tecnológicos desarrollados hasta ahora. Respecto a los métodos implementados en la investigación desarrollada, sin duda es posible mejorar el análisis de coherencia espacial comparando en tiempo real la información de las estaciones activas, alertando al usuario sobre algún fenómeno inusual o en su caso alertar a la dependencia si la estación está fallando y sirva como apoyo para un sistema gestor de mantenimiento preventivo y correctivo adecuado.

Uno de los problemas que retrasan el uso de información meteorológica, son los diversos formatos en los que las estaciones registran y almacenan los datos, por lo cual, es recomendable la obligatoriedad de la norma NMX-AA-166/1-SCFI-2013 (NMX-AA-166/1-SCFI, 2013) para el caso de México, la cual por su clasificación es señalada como deseable

únicamente pues no es una Norma oficial mexicana (NOM), ésto facilitará el uso de datos en la lectura, procesamiento y análisis de los mismos.

Al tener una base de datos confiable, puede ser aplicada en diversas actividades como: la pronósticos del clima, agricultura, infraestructura, alerta de fenomenos meteorológicos, etc. Tambien se pueden generar archivos o imágenes utilizables en software de Sistemas de Información Geográfica (SIG).

Como recomendacion general para a interpolacion de datos en zonas con topografía accidentada, el método de Kriging sólo es adecuado para variables meteorológicas con variabilidad gradual como lo es temperatura. Ya que en variables como precipitacion, es mas probable que se presenten lluvias convectivas (corta duración y espacialmente aisladas) y la interpolacion no refleje la realidad. Para ello se recomienda realizar una simulacion condicional y tomando en cuenta dicersas covariables como dirección del viento, velocidad del viento, entre otras.

- Abaurrea, J., Asín, J., Cebri, A., & Centelles, A. (2004). Metodología para el control de calidad y homogeneidad de una base de datos de precipitación diaria. *El Clima Entre El Mar y La Montaña*, (2001), 431–440.
- Aguilar, E. (2013). Aproximación a la homogeneización de una red regional de series climáticas a resolución diaria. *Reporte Técnico CRC-SAS-2013-003.*, 24. Retrieved from http://www.crc-sas.org/es/pdf/reporte_tecnico_CRC-SAS-2013-003.pdf
- Aguilar, López, J. M., Brunet, M., Saladié, O., Sigró, X., & López, D. (2013). Control de calidad y proceso de homogenización de series térmicas catalanas. *Cambio Climático, Unidad de Geografía (U.R.V. Tarragona)*, 15–23.
- Alvarez, H. O., Montaña, T., Quentin, E., Maldonado, J., & Solano, J. (2013). Homogeneización de series de velocidad del viento mensuales en las estaciones meteorológicas del INAMHI en Loja, Ecuador. *Revista de Climatología*, 13(unknown), 35–44.
- Araya, J., & Alfaro, E. (2009). Algoritmos de control de calidad para detección de secuencias atípicas de contactos en pluviómetros de balancín. *Tecnología En Marcha*, 22, 63–78. Retrieved from http://www.tecdigital.itcr.ac.cr/servicios/ojs/index.php/tec_marcha/article/view/117
- Arévalo, M, R. (2019). *Desarrollo de herramienta hidro-informática para estimar caudal ecológico en cuencas no aforadas* (UAEM, ed.). México: Universidad Autónoma del Estado de México.
- Balairón, R. L. (2014). Estadística para el análisis y la investigación del clima: Retos y soluciones estadísticas en clima y cambio climático. *Facultad de Estudios Estadísticos*.
- Basualdo, A. (2015). Manual de buenas prácticas para la generación, el almacenamiento y la difusión de informática climática en instituciones y organismos del MERCOSUR. In

- Instituto Interamericano de Cooperación para la Agricultura (IICA)*. Costa Rica.
- Becerril, J. (2020). *Diseño de una herramienta hidro informática de sistemas urbanos de drenaje sostenible* (IITCA-UAEM). Estado de México.
- Bidegain, M., Alvaro, F. C., & Diaz, F. (2011). Curso análisis estadístico de datos climáticos distribuciones de probabilidad. *Universidad de La República Montevideo, Uruguay*.
- Boshell, J. F., & León, G. E. (2013). *Metodologías para generar y utilizar información meteorológica a nivel subnacional y local frente al cambio climático* (GIZ, ed.). Perú.
- Castro, F. E. (2008). Manual de procedimientos para las estaciones meteorológicas. *Organización Para Estudios Tropicales*, 1–91.
- CLICOM. (2019). CLICOM. Retrieved from Características de CLICOM website: <https://clicom-mex.cicese.mx>
- COMET. (2019). Sistemas de gestión de la calidad: implementación en los servicios meteorológicos. Retrieved from <https://www.tiempo.com/ram/24047/sistemas-de-gestion-de-la-calidad-implementacion-en-servicios-meteorologicos/>
- CONAGUA. (2019). Comición Nacional del Agua.
- Cruz-Saco, V. R., Alvarado, J. N., & Espinoza, R. (2010). Software para el control de calidad y validacion de datos meteorológicos y generacion de rosa de vientos. *Centro de Energías Renovables de La Universidad Nacional de Ingeniería (CER - UNI)*, 8.
- Dubrulle. (1983). Cross validation of kriging in a unique neighborhood. *Mathematical Geology*, 15(0020-5958/83/1200-0687).
- Dudás, I. (2007). Quality control of worms. *The Theory and Practice of Worm Gear Drives*, 200–221. <https://doi.org/10.1016/b978-190399661-4/50009-5>
- Flemming, V., Cajé, J., Ulf, F., Margareth, M., Lars, A., Eino, H., ... Bordur, A. (2007). Quality control of meteorological observations. Automatic methods used in the Nordic Countries. *KLIMA-NORDKLIM*. <https://doi.org/10.1016/b978-190399661-4/50009-5>
- Gandin, L. (1988). Complex quality control of meteorological observations. *USSR*.
- GEM, G. del E. de M., & SMA, S. del M. A. (2009). *Iniciativa ante el cambio climático en el Estado de México*. 19–43. Retrieved from http://portal2.edomex.gob.mx/sma/cuida_medioambiente/cambio_climatico/groups/public/documents/edomex_archivo/sma_pdf_iniciativa_cam_clima.pdf
- Hans, A., & Moberg, A. (2002). Homogenization of Swedish Temperature Data. Part I:

- Homogeneity Test for Linear Trends. *International Journal of Climatology*, 17(1), 25–34. [https://doi.org/10.1002/\(sici\)1097-0088\(199701\)17:1<25::aid-joc103>3.3.co;2-a](https://doi.org/10.1002/(sici)1097-0088(199701)17:1<25::aid-joc103>3.3.co;2-a)
- Hansen, D. ., & Poulain, P. . (1995). *Quality Control and interpolations of WOCE-TOGA Drifter Data*. USA: Journal of atmospheric and oceanic technology.
- Hubbard, K., Guttman, N., Jinsheng, Y., & Chen, Z. (2006). An improved QC process for temperature in the daily cooperative. *Atmospheric and Oceanic Technology*, 24, 206–213. <https://doi.org/10.1175/JTECH1963.1>
- INIFAP. (2019). *INIFAP*. Retrieved from www.inifap.gob.mx/
- ISO-9001. (1947). *4. Normas Nacionales E Internacionales De Calidad Y Premios 4.1*. 1–18.
- ISO-9001. (2008). Norma internacional: Sistemas de gestión de calidad. Retrieved from www.iso.org
- Leblois, E., & Creutin, J. (2013). *Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method*. 49(December 2011), 3375–3387. <https://doi.org/10.1002/wrcr.20190>
- Llanos, H. L. (2015). RCLimTool Manual de usuario. *Clima y Sector Agropecuario Colombiano, 1*, 1–17. Retrieved from www.aclimatesectoragropecuariocolombiano.org
- Machiwal, D. (2009). Time Series Analysis of Hydrologic Data for Water Resources Planning and Management: A review. *J. Hydrol. Hydromech*, 54(3), 237–257. Retrieved from <http://dlib.lib.cas.cz:8080/472/>
- Maldonado, & Aravena. (2006). Redes de estaciones meteorológicas automáticas y sus aplicaciones productivas. *Chile, INIA*.
- Martinez, P., & Amoreo, E. (2015). Análisis de los datos aportados por la estación meteorológica automática del departamento de aeronautica, Facultad de ingeniería UNLP. *ITE*.
- Matlab. (2019). *Manual de funciones MATLAB* (MATworks, Ed.).
- Medina, G. G., Grageda, G. J., Corral, R. A., & Báez, G. D. A. (2008). *Uso de estaciones meteorológicas en la agricultura* (INIFAP, Ed.). México.
- Mercado-Mancera, Ramírez-Rodríguez, M., Vizcarra-Hernández, I. F., López-Antonio, H., López-Vargas, D. M., Granados-Mayorga, A. K., ... Chaires-Montecinos, B. E. (2014). Distribución Y Probabilidad De La Lluvia En Cuautitlán Izcalli, Estado De México.

<http://www.ommac.org/memoria2014/extensos/E2014092601.pdf>

- Miró, J. J., Estrela, M. J., & Barberá, M. J. (2014). Análisi de tendencias de series diarias de temperatura a partir de un downscaling estadístico con datos de reanálisi y redes neuronales. Aplicación a la comunidad valenciana. *Laboratorio de Meteorología-Climatología, Unidad Mixta CEAM-UVEG*.
- Montgomery, D. C., & Runger, G. C. (1994). Applied Statistics and Probability for Engineers. In *European Journal of Engineering Education* (Vol. 19). <https://doi.org/10.1080/03043799408928333>
- N., V., P., D., & Pomares P. (2015). Control de calidad de las mediciones de temperatura seca y humedad relativa en la estación meteorológica Sancti Spíritus. *Ciencia de La Tierra y El Espacio, 1*, 29–39.
- NMX-AA-166/1-SCFI. (2013). *Estaciones Meteorológicas, Climatológicas e Hidrológicas. Parte 1: Especificaciones técnicas que deben cumplir los materiales e instrumentos de medición de las EMA's* (2013th ed.; S. de Economía, Ed.). Méico.
- NOAA. (2019). National Hurricane Center. Retrieved from <https://www.nhc.noaa.gov/>
- Olivares, B., Cortez, A., Parra, R., Rodríguez, M. ., & Guevara, E. (2013). Application of statistics procedures for the quality control of the monthly rainfall series in the Venezuelan. *Revista de La Facultad de Agronomía, 30*(3), 367–391.
- OMM-100. (2011). Guía de prácticas climatológicas. In *Organización Meteorológica Mundial* (Vol. 100). <https://doi.org/OMM-Nº 168>
- OMM-168. (2011). Guía de prácticas hidrológicas: Hidrología de la medición a la información hidrológica. In *Manual*. <https://doi.org/OMM-Nº 168>
- OMM-305. (2000). Guía del sistema mundial de proceso de datos. *Organización Meteorológica Mundial - , 1*.
- OMM-306. (2018). *Manual de claves internacionales* (Organizaci).
- OMM-485. (2010). Manual del Sistema Mundial de Proceso de Datos y de Predicción Volumen I-Aspectos mundiales. *Organización Meteorológica Mundial, 200*.
- OMM-788. (1993). *Guía sobre la gestión de la vigilancia meteorológica mundial* (WORLD METEOROLOGICAL ORGANIZATION, Ed.).
- OMM. (2012). *Manual del Sistema Mundial de Proceso de Datos y de Predicción. 200*.

- Ortiz, H., Escobar, C. D., & Sepúlveda, S. B. (2018). Análisis estadístico de variables climatológicas en la ciudad de Cúcuta. *Respuestas Engineering Journal*, 23(1), 39. <https://doi.org/10.22463/0122820x.1328>
- Ouarda et. al. (2006). On the critical values of the standard normal homogeneity test (SNHT). *International Journal of Climatology*.
- Oviedo, B., & Rojas, N. (2012). Applied method for evaluating data series in automatic weather station. *CT-Colombia*.
- Pérez-Planells. (2015). Análisis de métodos de validación cruzada para la obtención robusta de parámetros biofísicos. *Revista de Teledetección*, (1133-0953). Retrieved from <http://dx.doi.org/10.4995/raet.2015>
- Prieto. (2008). *Diagnóstico de las capacidades, fortalezas y necesidades para la observación, monitoreo, pronóstico y prevención del tiempo y el clima ante la variabilidad del cambio climático en México*. México.
- PROY-NMX-AA-166. (2015). Estaciones meteorológicas , climatológicas e hidrológicas . Parte 2 : Especificaciones técnicas que deben cumplir el emplazamiento y exposición de los instrumentos de medición de las estaciones meteorológicas automáticas y convencionales. *PROYECTO DE NORMA MEXICANA PROY-NMX-AA-166/2-SCFI-2015, 1*.
- Ramchandra, K. (2010). Status of automatic weather stations in nepal and comparison of air temperature and precipitation data between automatic weather station and manual observation. *Department of Hydrology and Meteorology, (DHM)*. Retrieved from www.wmo.int/pages/prog/www/IMOP/publications/
- Rodríguez, R., García, V., Hernández, L. A., Taibo, P., Seoane, A., Jaspe, A., & Valera, I. (2008). ANÁLISIS Y UTILIZACIÓN DE LA INFORMACIÓN TRÁFICO MEDIANTE SISTEMAS INTEGRADOS DE INFORMACIÓN GEOGRÁFICA (SIG) Y SISTEMAS DE NAVEGACIÓN 3D. *Universidad La Coruña*.
- Salgado, N. (2018). Diseño de una herramienta hidro-informática para el análisis de calidad de datos de las EMA's del Estado de México. *UAEMex*.
- Sanchez, J. (2008). Distribuciones Estadísticas. *Universidad de Salamanca*, 1–12.
- Serrano, V., Morata, A., González-Hidalgo, J., Rico, A. M., Tomas-Burguera, M., Moltó, E., ... Olcina, J. (2016). Control de calidad de siete variables del banco nacional de datos

- de AEMET. *Clima, Sociedad, Riesgos y Ordenación Del Territorio*, 407–415. <https://doi.org/10.14198/xcongresoaealicante2016-38>
- Shulski, M. D., You, J., Krieger, J. R., Baule, W., Zhang, J., Zhang, X., & Horowitz, W. (2014). Quality assessment of meteorological data for the Beaufort and Chukchi sea coastal region using automated routines. *Arctic Institute of North America*, 67(1), 104–112. <https://doi.org/10.14430/arctic4367>
- Shulskii, You, J., Krieger, J. R., Baule, W., Zhang, J., Zhang, X., & Horowitz, W. (2014). Quality assessment of meteorological data for the Beaufort and Chukchi sea coastal region using automated routines. *Arctic*, 67(1), 104–112. <https://doi.org/10.14430/arctic4367>
- SMN. (2012). *Servicio Meteorológico Nacional: 135 años de historia en México* (CONAGUA, Ed.). México.
- Štěpánek, P., Zahradníček, P., & Farda, A. (2013). Experiences with data quality control and homogenization of daily records of various meteorological elements in the Czech Republic in the period 1961-2010. *Idojaras Hungarian Meteorological Service*, 117(1), 123–141.
- UNE-176101. (2010). *Redes de estaciones agrometeorológicas automáticas: Características, instrumentación y aspectos específicos*. España.
- UNE-500510-IN. (2005). *Redes de estaciones meteorológicas automáticas. Aspectos generales y nomenclatura*. España.
- UNE-500520. (2002). *Redes de estaciones meteorológicas automáticas: criterios de localización de emplazamientos e instalación de sensores característicos de adquisición y muestreo*. *Asociación Española de Normalización y Certificación*.
- UNE-500540. (2004). *Redes de estaciones meteorológicas automáticas: Directrices para la validación de registros meteorológicos procedentes de redes de estaciones automáticas*. España.
- Valleumbroso, V. (2014). *Métodos probabilísticos de hidrología general*.
- Villarreal, G. ., Reyes, J. L., Carrillo, J. J., Velázquez, J. E., & Paredes, V. (2014). Estudio para el proyecto hidrológico para proteger a la población de inundaciones y aprovechar mejor el agua (PROHTAB). *CONAGUA*, 4.
- Walpole, & Mye. (2007). *Probabilidad Y Estadística Basica Para Ingenieros*. Retrieved

from

https://archuto.files.wordpress.com/2011/02/probabilidad_y_estadistica_basica.pdf

Walpole, R., Myers, R., Myers, S., & Ye, K. (2012). Probabilidad y estadística para ingeniería y ciencias. In Pearson (Ed.), *Pearson Educación* (Novena). <https://doi.org/10.1192/bjp.111.479.1009-a>

Anexo 4.1 Clave de los estados de la República mexicana

Anexo 4.2” Diagramas de flujo por métodos de análisis de la herramienta hidro-informática “SICCDAMET

Anexo 4.3” Diagramas de flujo de la interfaz de la herramienta hidro-informática “SICCDAMET

Anexo 5.1 Manual de usuario de la herramienta hidro-informática SICCDAMET

Anexo 5.2 Ejemplos de reportes de resultados para la estación Observatorio UAEM

Anexo 5.3 Validación del análisis de calidad a la base de datos de la estación Observatorio UAEM

Anexo 5.4 Listado de estaciones consideradas en el proyecto

Anexo 5.5 Resultados del análisis de calidad a la base de datos de las EMA’s

Anexo 5.6 Imágenes de resultados del análisis de calidad a la base de datos de las EMA’s

Anexo 5.7 Resultados del análisis de calidad a la base de datos de las EMC’s

Anexo 5.8 Imágenes de resultados del análisis de calidad a la base de datos de las EMC's

Anexo 5.9 Resultados de la interpolación de T_{máx}, T_{mín} y P

Anexo 5.10 Resultados de la simulación de P para el día 01/07/2016