

**PROYECTO “BOLETINES
METEOROLÓGICOS SINTÉTICOS PARA
MUNICIPIOS DE MÉXICO”
TH2015.1
Informe final**

COORDINACIÓN DE HIDROLOGÍA
SUBCOORDINACIÓN DE HIDROMETEOROLOGÍA

MC. Roberto Ramírez Villa
Dr. Julio Sergio Santana Sepúlveda
Dr. René Lobato Sánchez
MC. Oscar Pita Díaz

México, 2020

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 1 de 24	México, 2020	Clave: F.C0.2.04.01

INDICE

1	INTRODUCCIÓN.....	2
2	DATOS EMPLEADOS.....	2
3	SISTEMA DE GENERACIÓN DE BOLETINES METEOROLÓGICOS SINTÉTICOS	4
3.1	Procesamiento	4
3.2	Boletín Meteorológico sintético.....	5
4	METODOLOGÍA PARA LA EVALUACIÓN DEL PRONÓSTICO.....	7
4.1	Homogeneización de datos	7
4.2	Métodos de verificación.....	8
4.3	Análisis estadístico por municipio	9
4.4	Análisis estadístico regional	11
5	EVALUACIÓN DEL PRONÓSTICO.....	14
5.1	Análisis estadístico por municipios	14
5.2	Análisis a nivel nacional.....	17
6	PUBLICACIÓN DE PRONÓSTICOS Y EVALUACIÓN	19
6.1	Módulos de visualización.....	20
6.2	Desarrollo de los módulos de visualización	22
7	CONCLUSIONES.....	23
8	BIBLIOGRAFÍA.....	23

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 2 de 24	México, 2020	Clave: F.C0.2.04.01

1 INTRODUCCIÓN

La extensa área que comprende el territorio mexicano implica un reto para el análisis operativo de eventos hidrometeorológicos que, de acuerdo con el Centro Nacional de Prevención de Desastres, conforman el principal fenómeno asociado a pérdidas y desastres en el país los últimos 15 años (Cenapred, 2018). Los sistemas atmosféricos de los que se desprenden presentan una amplia variedad y complejidad en su generación, evolución, así como a las distintas escalas espaciales y temporales que pueden presentarse. Por tal motivo, instituciones de gobierno, académicas, centros de investigación y el sector privado han implementado a través de los años, de manera conjunta o independiente, una amplia infraestructura en observatorios, redes de monitoreo, equipos tecnológicos de análisis, y recursos humanos especializados, con el propósito de atender la necesidad de información estratégica útil sobre el clima y el estado del tiempo.

Los pronósticos meteorológicos son informes operativos que describen el posible estado del tiempo a corto plazo, generalmente a nivel nacional o regional, sin embargo actualmente existe una limitada disponibilidad de reportes especializados a nivel local para municipios de México. Una alternativa para abordar esta problemática la ofrece el uso de sistemas y recursos libres como software de código abierto, datos satelitales, sistemas de información geográfica, observaciones, reanálisis climáticos, modelos numéricos de la atmósfera y del océano, entre otros, los cuales son compartidos para su aprovechamiento por instituciones de investigación y de gobierno con la comunidad científica.

En 2019 la Subcoordinación de Hidrometeorología, del Instituto Mexicano de Tecnología del Agua, implementó un sistema operacional para la generación de boletines meteorológicos a escala local, generados de forma automática (sintética) a partir de recursos libres y que podían ser consultados en línea (<http://galileo.imta.mx/>). La segunda fase de este proyecto inició a principios de 2020 con la integración de un pronóstico por ensambles, el cual permite la incorporación de datos de probabilidad ante la ocurrencia de eventos, además de llevar a cabo una evaluación con fines de correcciones de sesgo y posible mejora del pronóstico.

2 DATOS EMPLEADOS

Se empleó el Sistema GFS (Global Forecast System) del Centro Nacional de Predicción Ambiental de EE.UU. (NCEP) como modelo meteorológico numérico con perspectiva a 168 hrs, el cual es un modelo determinista acoplado, compuesto por cuatro modelos separados (un modelo de atmósfera, un modelo de océano, un modelo de tierra / suelo y un modelo de hielo marino), que trabajan juntos para proporcionar las condiciones atmosféricas en un instante dado (<https://www.ncdc.noaa.gov/>). Regularmente se realizan cambios en el modelo GFS para mejorar su desempeño y precisión de pronóstico, por lo que se encuentra en constante evolución y mejora. Los datos están disponibles para su descarga a través del Sistema Nacional de Archivo y Distribución de Modelos Operativos (NOMADS) de la NOAA.

No obstante, al considerar la atmósfera como un sistema caótico no lineal, pequeños errores en las condiciones iniciales pueden ocasionar diferencias significativas en los resultados de los modelos numéricos deterministas (Slingo y Palmer, 2011). Una manera de abordar esta situación puede ser a través del uso de modelos por ensambles donde, en lugar de ejecutar simulaciones únicas, se utilizan múltiples corridas (miembros del ensamble) empleando perturbaciones físicas para su procesamiento, lo que permite obtener una estimación de la incertidumbre actual (Stensrud et al., 2000), así como la evolución más probable mediante el análisis de la dispersión en sus resultados (Figura 2.1).

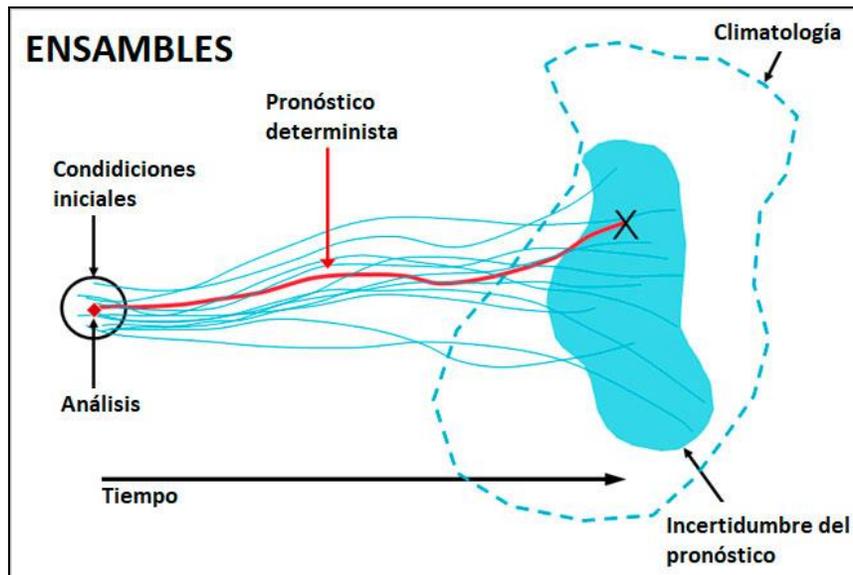


Figura 2.1. Evolución de pronósticos numéricos deterministas y construcción de un modelo por ensamble

El SREF (Short-Range Ensemble Forecast) es un sistema de ensamble del modelo WRF (Weather Research Forecast) que opera de forma operativa por NCEP desde mayo de 2001 (Du et. al., 2003), e integra 24 miembros de ensamble y 2 de control, configurado para un dominio que cubre Norteamérica, con 4 ejecuciones diarias, resolución espacial de 16 km y temporal de 3 hrs, para un periodo total de simulación de 84 hrs. Sus resultados se encuentran disponibles de forma libre en: <https://www.nco.ncep.noaa.gov/pmb/products/sref/>.

Por su parte, para el proceso de evaluación se utilizaron datos provenientes de la red de estaciones meteorológicas del Servicio Meteorológico Nacional, correspondiente a valores de temperatura máxima-mínima, precipitación y magnitud del viento en 24 horas, para el periodo de tiempo de realización contemplado en este proyecto (marzo-octubre 2020).

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 4 de 24	México, 2020	Clave: F.C0.2.04.01

3 SISTEMA DE GENERACIÓN DE BOLETINES METEOROLÓGICOS SINTÉTICOS

El proyecto de boletines meteorológicos sintéticos a nivel local se origina como una iniciativa del IMTA para contribuir a solventar la necesidad de reportes especializados del estado del tiempo para municipios del país, que puedan apoyar en procesos de planeación, prevención y atención de emergencias ante la probabilidad de eventos severos. Al término de este proyecto los boletines son generados diariamente de forma automática por un complejo conjunto de programas de cómputo que emplea software e información de carácter libre, y que a través de un sitio web publica sus resultados.

3.1 Procesamiento

El desarrollo e implementación del sistema de generación de boletines sintéticos se realizó mediante la utilización de un sistema operativo de código abierto (Linux/Rocks), en el cual se pueden instalar paquetes informáticos, configurar procesos, y elaborar códigos de programación a través de ficheros de sistema que realizan instrucciones específicas complejas, lo que reduce la demanda de recursos del sistema y que resulta favorable para equipos que no se consideran de alto rendimiento.

Otro recurso empleado fue el conjunto de paquetes informáticos libres (<https://www.python.org/>, <http://cola.gmu.edu/grads/>, <https://www.r-project.org/>) para procesamiento y lectura de datos meteorológicos globales y regionales que provienen de fuentes gubernamentales, académicas y científicas, y con ello finalmente elaborar los reportes a través de un sistema de composición de texto automatizado (<https://www.latex-project.org/get/>). Los códigos desarrollados en los distintos lenguajes de programación para la ejecución del sistema de generación de boletines se encuentran anexos a este documento en formato digital.

La instalación, configuración y ejecución del sistema de generación de boletines se implementó en un Clúster de alto rendimiento con las siguientes características:

Sistema de almacenamiento tipo NAS (Network Attached Storage) con sistema operativo (Linux/Rocks), compiladores PGI Cluster Development Kit, bibliotecas para manejo y visualización de netCDF (NCO, CDO, NCView, Ferret), intérpretes de los lenguajes de programación Python, R y GrADS. Cuenta con un nodo maestro de dos procesadores de doce núcleos (cores) de procesamiento de 2.6 GHz, configuración en memoria de 256 GB de RAM en dimms DDR3 1333/1066/1600MHz, 2 discos de 300GB tipo SAS a 15K RPM para Sistema Operativo, 2 HBA's FC de doble puerto a 8 Gb, e integra un GPU de 2496 cores con soporte para CUDA. También cuenta con 8 nodos esclavos de 64 núcleos (cores) de 2.8 GHz c/u, memoria de 256 GB en dimms DDR3 de frecuencia 1600MHz y disco duro con una capacidad de 200 GB SAS SSD para Sistema Operativo.

3.2 Boletín Meteorológico sintético

Los reportes se presentan en un esquema simple, práctico y legible que contiene el pronóstico de hasta 96 horas, redactado en párrafos resumidos con la siguiente información:

- Temperaturas máximas y mínimas así como condiciones descriptivas del ambiente (frío, templado, cálido, etc) en función de la climatología del lugar.
- Intensidad, periodo del día y probabilidad de ocurrencia de precipitación, utilizando como base la escala de clasificación de lluvias del Servicio Meteorológico Nacional para identificar los eventos intensos (SMN, Tabla 3.1).

Categoría	Lluvia acumulada en 24 hrs (mm)	Categoría	Lluvia acumulada en 24 hrs (mm)
Lluvia	0.1 a 25.0	Lluvia intensa	75.1 a 150
Lluvia fuerte	25.1 a 50.0	Lluvia torrencial	150.1 a 250.0
Lluvia muy fuerte	50.1 a 75.0	Lluvia extraordinaria	> 250.1

Tabla 3.1. Clasificación de la precipitación acumulada en 24 horas para pronósticos a corto y mediano plazo (24 a 96 horas)

- Condiciones de cielo.
- Ocurrencia de tormentas.
- Velocidad de vientos y rachas (intensidad y dirección).

Además el reporte incluye los siguientes elementos:

- Imagen satelital actualizada en canal visible (banda 2, 0.64 μm) con luz de día, e infrarrojo (banda 13, 10.35 μm) durante la madrugada y noche.
- Perspectiva a 7 días a través de series de tiempo de variables meteorológicas (temperaturas máximas y mínimas, viento y precipitación).
- Iconos del estado del tiempo.

El proceso de generación de boletines se ejecuta de forma diaria en un equipo que requiere de un tiempo aproximado de 4 horas, desde la descarga de insumos, procesamiento de la información y construcción de las series de tiempo hasta la elaboración de los párrafos de descripción meteorológica en documentos con formato “.pdf” (Figura 3.1), que se encuentran disponibles a partir de las 07:00 am para más de 1500 municipios del país.

Pronóstico Meteorológico para: Aldama, Tams.
3 de septiembre de 2020



Condiciones meteorológicas para hoy jueves:
Ambiente cálido con una Tmáx de 32°C y una Tmín de 24°C (templado a cálido), cielo despejado a medio nublado **con probabilidad de lluvia** - 60% (por la mañana, tarde, madrugada) y potencial de tormentas. Viento del sureste rolando al este de 15 a 20 km/h.

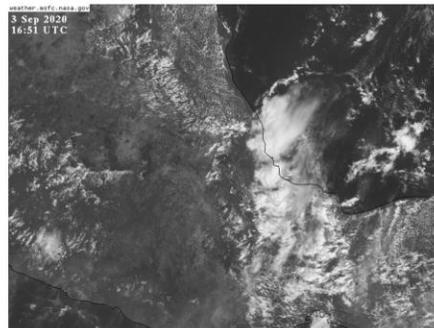


Imagen de satélite (weather.msfc.nasa.gov/)



Pronóstico para hoy

Pronóstico viernes 04 de septiembre: Ambiente templado a cálido con una Tmáx de 30°C y una Tmín de 23°C (templado a cálido), cielo parcialmente nublado a nublado **con probabilidad de lluvia** - 58% (a lo largo del día) y potencial de tormentas aisladas. Viento predominante del noreste de 15 a 25 km/h (rachas: 25 a 30 km/h).

Pronóstico sábado 05 de septiembre: Ambiente fresco a templado con una Tmáx de 26°C y una Tmín de 23°C (templado a cálido), cielo nublado **con probabilidad de lluvia** - 51% (a lo largo del día). Viento predominante del noreste de 15 a 20 km/h.

Pronóstico domingo 06 de septiembre: Ambiente templado a cálido con una Tmáx de 30°C y una Tmín de 22°C (templado), cielo medio nublado a nublado **con probabilidad de lluvia** (por la noche, madrugada) y potencial de tormentas aisladas. Viento del noreste rolando al norte de 15 a 20 km/h.



Pronóstico a 7 días para Aldama

No es un pronóstico oficial. Producto experimental para evaluación de modelos numéricos.
Subcoordinación de Hidrometeorología - Instituto Mexicano de Tecnología del Agua

Figura 3.1. Formato del boletín meteorológico sintético por municipios

La publicación de los resultados se realiza mediante la implementación de un sistema que despliega los elementos gráficos y dinámicos a través de un portal web, instalado en el servidor galileo.imta.mx. El servidor Galileo corresponde a un Dell Power Edge T610 con doble procesador Xeon a 2.8 Ghz, 32 GB RAM y 400 GB de disco duro, en el cual se instaló RedHat Enterprise Linux como sistema Operativo y se encuentran habilitados los servicios de Apache como servidor Web con soporte para PHP y MySQL como manejador de bases de datos, así como el conjunto de compiladores PGI y GrADS como herramienta para el análisis y despliegue de datos meteorológicos y climáticos, entre otras utilerías.

4 METODOLOGÍA PARA LA EVALUACIÓN DEL PRONÓSTICO

En este proyecto se llevó a cabo una evaluación de los resultados del pronóstico meteorológico a 24, 48 y 72 horas, con respecto a observaciones, empleando métodos de verificación como medidas de la relación entre lo pronosticado y lo ocurrido (Wilks, 2006), pues un mejor análisis mediante esquemas de verificación complejos resultará en una mejor apreciación sobre las deficiencias del pronóstico y provee la posibilidad de mejorar el entendimiento de los procesos físicos que están siendo pronosticados (Jollife, 2012).

4.1 Homogeneización de datos

Las observaciones empleadas para la evaluación fueron sometidas a un proceso de control de calidad y homogeneización utilizando el paquete del lenguaje R Climatol (<http://www.climatol.eu/>), en el que mediante pruebas estadísticas de homogeneidad, se ajustan los datos con el fin de descartar valores sospechosos en las series y completar datos ausentes. Esta herramienta funciona a través de series de referencia de estaciones cercanas, asimismo se puede controlar la calidad de las series y eliminar anomalías que superen un umbral prefijado (Gujarar, 2018).

Se utilizaron estaciones meteorológicas que tuvieran como mínimo 20 datos medidos por mes, desde el mes de marzo hasta el mes de agosto del 2020 como se muestra en la Tabla 4.1.

	Marzo	Abril	Mayo	Junio	Julio	Agosto
Temperatura máxima	1319	1347	1365	1335	1341	1282
Temperatura mínima	1317	1348	1364	1341	1338	1283
Precipitación	1976	1654	1635	1589	1580	1629
Magnitud del viento	87	130	123	120	118	119

Tabla 4.1. Número de estaciones analizadas por mes y por variable.

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 8 de 24	México, 2020	Clave: F.C0.2.04.01

Uno de los parámetros delimitados para la homogeneización de los datos de temperatura, precipitación y magnitud del viento fue el ajuste de los datos entre -5 y 5 desviaciones estándar. Para el análisis de calidad en Climatol, se optó por realizar un primer análisis exploratorio de los datos de cada variable, con el fin de realizar un diagnóstico y verificar errores que puedan contener los datos en cada una de las estaciones; con este análisis exploratorio se puede conocer la distribución global de los datos, ubicación de las estaciones, así como un agrupado de estaciones que presentan comportamientos similares, desviaciones estándar máximas y mínimas de los datos, valores de prueba de homogeneidad y gráficos con las series reconstruidas.

Una vez finalizado el control de calidad y homogeneización de los datos observados se generan gráficos exploratorios y valores de puntajes como métodos de verificación (diagramas de dispersión, boxplots, error y raíz del error cuadrático medio) para la validación del pronóstico, con el propósito de poder aplicar posteriormente esquemas de ajustes de sesgo para reducir la incertidumbre y mejorar los pronósticos meteorológicos (Cui et. al, 2012; Jolliffe, 2012).

4.2 Métodos de verificación

Son herramientas que indican la medida en la que difieren los valores de los pronósticos de los valores de las observaciones. Estos pueden incluir gráficos exploratorios tales como diagramas de dispersión, histogramas y diagramas de caja (box plots), así como también valores de puntajes como la raíz del error cuadrático medio.

Diagramas de dispersión: representa los valores pronosticados contra los observados mediante un gráfico de puntos en el cual las escalas en los ejes son iguales. Indica la medida de correspondencia entre los valores observados y los pronosticados. Un pronóstico perfecto estará representado por cualquier punto sobre la recta de 45 grados en el diagrama. Si se grafica en el diagrama de dispersión la línea de regresión ajustada a los datos, la correspondencia entre ésta y la recta de 45 grados representa una medida de la confiabilidad de los pronósticos. Otra medida de verificación que puede estimarse a partir de un diagrama de dispersión es el error promedio o sesgo (bias).

Box plots: muestran el rango de datos que cae entre los percentiles 25 y 75, representados por los extremos de la caja. La línea horizontal dentro de la caja muestra el valor de la mediana y los bigotes muestran el rango completo de los datos. También son útiles para identificar valores fuera de rango o outliers, los cuales se identifican mediante un símbolo especial.

Error cuadrático medio (Mean squared error, MSE): mide la diferencia cuadrática media entre los pronósticos y las observaciones. No indica la dirección de las desviaciones. Sus unidades son el cuadrado de las unidades básicas. Rango de valores: 0 a ∞ . Puntuación perfecta: 0.

$$MSE = \frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2$$

Raíz del error cuadrático medio (Root mean square error, RMSE): es una medida de la precisión total. Proporciona la magnitud promedio de los errores del pronóstico, ponderada de acuerdo al

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 9 de 24	México, 2020	Clave: F.C0.2.04.01

cuadrado del error. Es un puntaje simple y de uso común que se calcula como la raíz cuadrada del MSE.

$$RMSE = \left[\frac{1}{N} \sum_{i=1}^N (F_i - O_i)^2 \right]^{1/2}$$

Al igual que el MSE, el RMSE no indica la dirección de las desviaciones. Es un puntaje que se ve más influenciado por los errores grandes que por los pequeños, lo cual puede ser bueno si los errores grandes son especialmente indeseables. Su sensibilidad a errores grandes también significa que no puede proporcionar estimaciones estables del error si se utilizan muestras pequeñas. Esta medida, al igual que otras reglas de puntuación cuadrática, puede favorecer un pronóstico conservador (pronóstico cercano a la media climatológica). Rango de valores: 0 a ∞ . Puntuación perfecta: 0.

4.3 Análisis estadístico por municipio

En este proceso se realiza una evaluación de pronósticos, comparando datos de observaciones con datos pronóstico de 24, 48 y 72 horas.

- Obtención de datos de las mallas a través de las coordenadas de estaciones por municipio, en este caso se han utilizado 1570 municipios para ser analizados a nivel nacional. Mediante líneas de código ejecutadas en el lenguaje R, se extrajeron los datos diarios de las mallas de observaciones, así como de las mallas de los ensambles a 24, 48 y 72 horas para cada variable y por cada mes. Al extraer los datos de las mallas, los datos por municipio se han guardado en archivos .csv, utilizando una nomenclatura para el nombre de cada archivo iniciando con el mes y año, seguido de un guion bajo, la abreviación del estado correspondiente, guion bajo y el nombre del municipio y la variable climatológica, por ejemplo: 052020_Ags_Aguascalientes_tmax.csv; indicando en este ejemplo el mes de mayo del 2020, para el municipio de Aguascalientes, Ags. con la variable temperatura máxima.
- Obtención de diagramas de dispersión, comparando los datos de observaciones contra datos de 24, 48 y 72 horas obtenidos de los ensambles. En este diagrama, se busca, como principal objetivo, encontrar correlaciones entre los datos de observaciones y los datos de los ensambles. Siendo este diagrama una herramienta para verificar la calidad de los datos y que tan confiable es el pronóstico mediante ensambles. Para obtener estos diagrama, se realizaron líneas de código en el lenguaje R, en donde se importa el archivo .csv de cada municipio, posteriormente se realiza cálculo de la regresión lineal comparando observaciones vs 24 hrs, observaciones vs 48 hrs y observaciones vs 72 hrs.

En la Figura 4.1 se muestra un ejemplo de los diagramas de dispersión obtenidos para la variable temperatura máxima representado en el mes de mayo del 2020, indicando en el título el nombre de la variable, el municipio, las coordenadas en grados decimales de la ubicación del municipio y el mes representado; asimismo se muestran los tres gráficos de dispersión comparando las

observaciones (eje X) con 3 distintos intervalos de pronóstico (eje Y). Cada gráfico, en la parte superior, muestra una etiqueta con el intercepto de la pendiente, el valor de correlación de Pearson, y el RMSE:

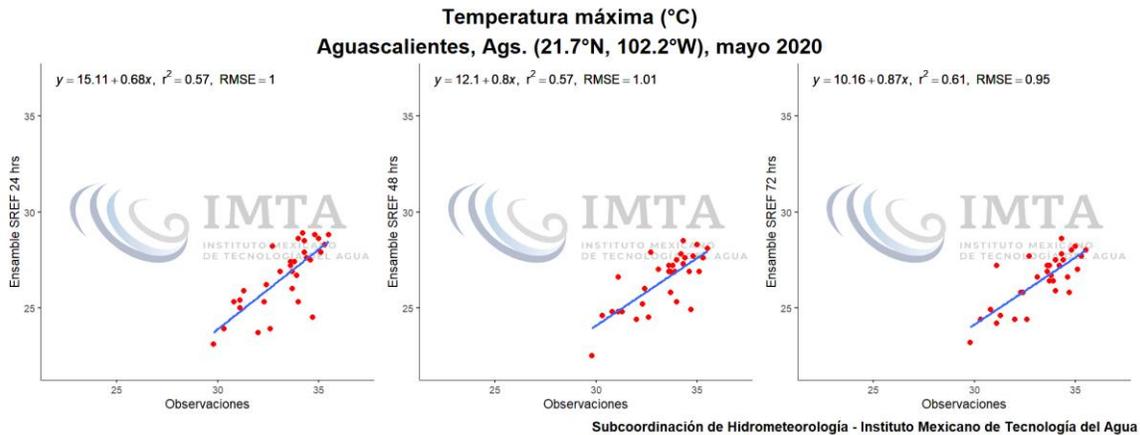


Figura 4.1. Ejemplos de diagramas de dispersión generados para el análisis

Además de los gráficos de dispersión, se realizaron diagramas de cajas mostrando el error, ya sea sobreestimado o subestimado, definido en unidades dependiendo de la variable utilizada (°C en temperaturas, mm en la precipitación y km/h en la magnitud del viento). Este diagrama es, al igual que el diagrama de dispersión, muy importante para poder analizar y detectar grandes diferencias a través de los límites máximos y mínimos del diagrama, así como los outliers. En la Figura 4.2 se muestra un diagrama de cajas de la variable precipitación para el mes de julio correspondiente al municipio de Bachiniva, Chihuahua:

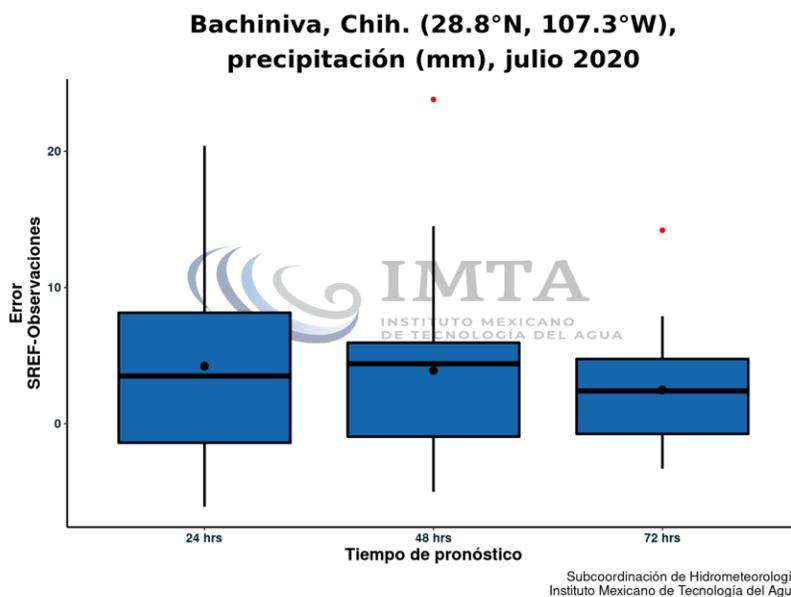


Figura 4.2. Ejemplo de diagramas de cajas generados para el análisis

Para complementar el análisis de evaluación de pronóstico, se crearon diagramas de caja por cada variable y por cada mes, calculando los valores de RMSE. Con este resultado se obtiene un análisis general de la diferencia absoluta entre observaciones y el pronóstico por ensambles, para el total de conjunto de datos como se muestra en la Figura 4.3.

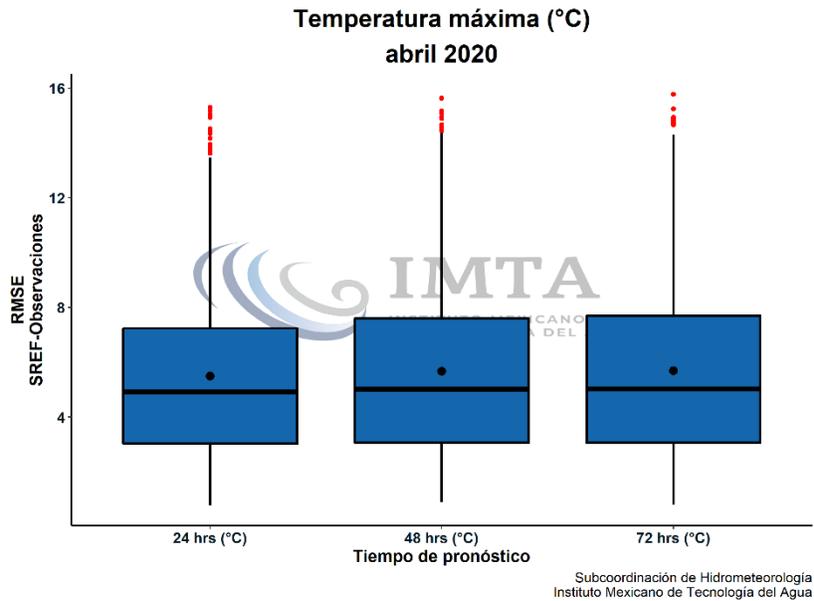


Figura 4.3. Ejemplos de diagramas de cajas generados para el total de conjunto de datos

4.4 Análisis estadístico regional

En la sección 4.2 se definió la fórmula para calcular el RMSE, que, para un conjunto de puntos en el espacio multidimensional, con el tiempo incluido, establece una medida del error o diferencia *absoluta* entre las observaciones y los valores estimados de una variable dada. El conjunto de puntos, dado un tiempo fijo, se puede acotar a una región espacial, como pudiera ser, por ejemplo, un estado, una cuenca o cualquier otra región de interés; o bien, dado un punto fijo en el espacio geográfico, se puede acotar a su vez un período de tiempo, como pudiera ser, por ejemplo, un mes, un año, una estación del año, o cualquier otro período de interés. En el presente proyecto se optó por el segundo enfoque, es decir, los valores de la variable, tanto las estimaciones como las observaciones se establecieron para cada punto particular de acuerdo con la escala del tiempo, tomando como período de interés el mensual. Así, cada punto de malla se le pudo asociar para cada mes, un único valor del RMSE resultante de aplicar la fórmula referida, a los valores estimados y de observaciones, correspondientes a los alrededor de 30 días del mes en cuestión. Estos valores en la malla, se utilizaron para crear los mapas de RMSE, por variable para cada uno de los meses del análisis. Finalmente, estos mapas brindan una visión regional del comportamiento del modelo a lo largo del tiempo.

Para llevar a cabo el procedimiento descrito en el párrafo anterior se desarrolló un sistema, escrito en el lenguaje de programación R, cuyas características principales se muestran en la Figura 4.4,

donde se presenta la arquitectura del sistema. En ese esquema se identifican varios módulos, así como las entradas que requieren y las salidas que generan. En seguida se describe cada uno de los módulos.

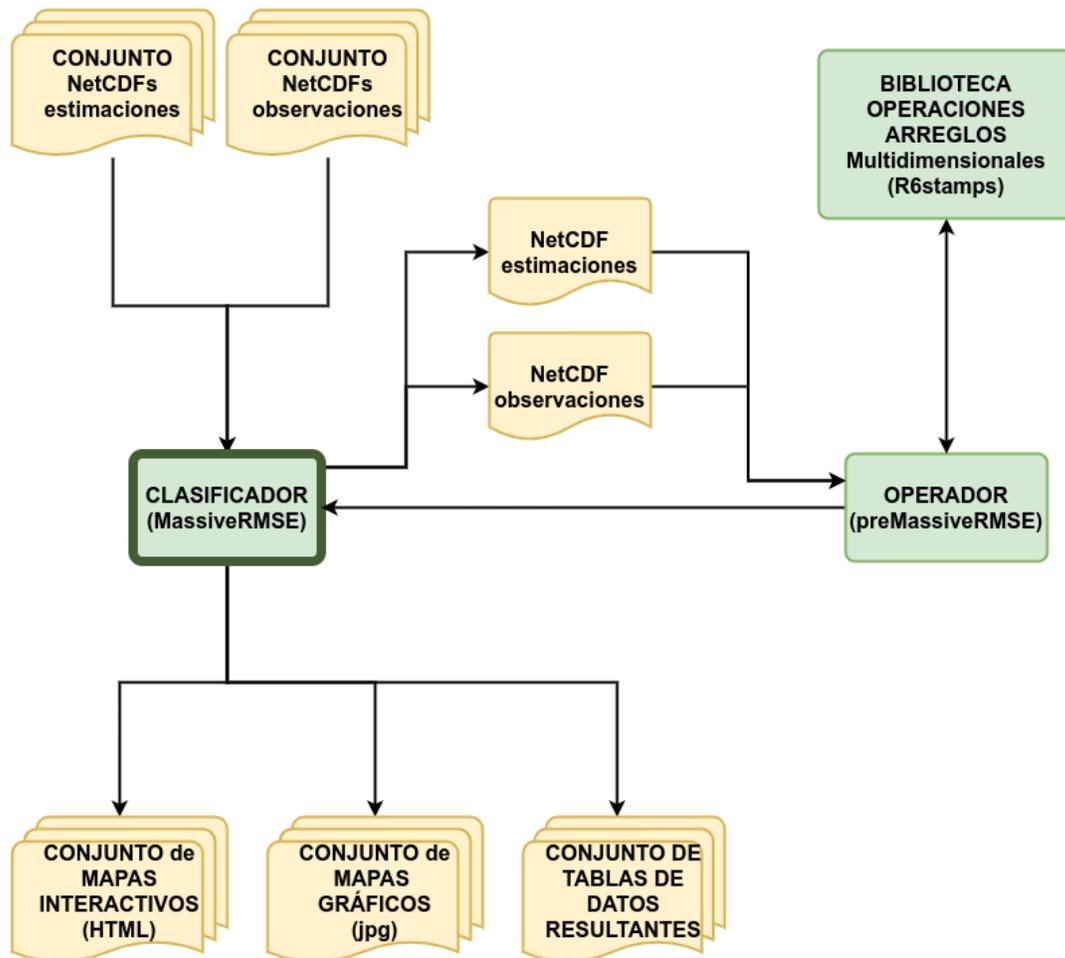


Figura 4.4. Arquitectura del sistema para la generación de mapas de RMSE: **MassiveRMSE**.

- Clasificador.** Este es el módulo rector del sistema que recibe como entrada los distintos conjuntos de datos tanto de los valores estimados como de observaciones para todas las variables consideradas. Los datos vienen en un formato NetCDF y los nombres de los archivos se deben construir de acuerdo con una sintaxis que permite identificar, la variable de la que se trata, el período mensual con año, si los datos provienen de estimaciones u observaciones, y en el caso de estimaciones, si el pronóstico es a 24, 48 o 72 horas. La primera tarea del Clasificador consiste en invocar al módulo Operador, mediante la selección las parejas adecuadas de archivos NetCDF, uno de estimaciones y otro de observaciones y su envío al módulo. Esta tarea, que se hace para cada pareja seleccionada, se muestra como la salida *superior* del módulo en la Figura 4.4. La respuesta del módulo Operador es un arreglo de valores RMSE para cada pareja de

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 13 de 24	México, 2020	Clave: F.C0.2.04.01

archivos. Estos arreglos se usan para la construcción de los mapas en formatos html y jpg y las tablas de datos resultantes, lo que constituye la salida final del sistema. En R, la función que se usa para invocar este módulo se llama **MassiveRMSE**.

- **Operador.** Este módulo toma un par de archivos NetCDF, uno de estimaciones y el otro de observaciones y calcula el RMSE para cada una de las series de tiempo asociadas a cada punto registrado. Para ejecutar sus cálculos, invoca a las operaciones contenidas en la Biblioteca de Operaciones de Arreglos Multidimensionales. El resultado se entrega como un arreglo bidimensional o malla, correspondiente a las coordenadas geográficas longitud y latitud, que contiene los valores del RMSE calculados en el proceso. En R, la función que se usa para invocar este módulo se llama **preMassiveRMSE**.
- **Biblioteca de operaciones de arreglos multidimensionales.** Esta es una biblioteca de operaciones entre arreglos multidimensionales, desarrollada para facilitar las operaciones matemáticas entre los elementos de los arreglos, tales como sumas, multiplicaciones, exponenciaciones, etc. Esta biblioteca, también permite hacer operaciones de agregación en un solo arreglo, tales como sumatorias, promedios, medianas, etc. Por ejemplo, para el cálculo del RMSE se utilizaron operaciones de resta, exponenciación, raíz cuadrada y la operación de agregación cálculo del promedio. En R, esta biblioteca, desarrollada bajo los principios de la programación orientada a objetos se llama **R6stamps**.

En la Figura 4.5 se muestran las salidas del sistema; en la izquierda se tiene el mapa en formato html, mientras que a la derecha se presenta el mapa en formato jpg. La diferencia entre uno y otro es que el segundo es estático, mientras que el primero es interactivo y permite operaciones de acercamiento (*zoom*) y deslizamiento (*pan*).

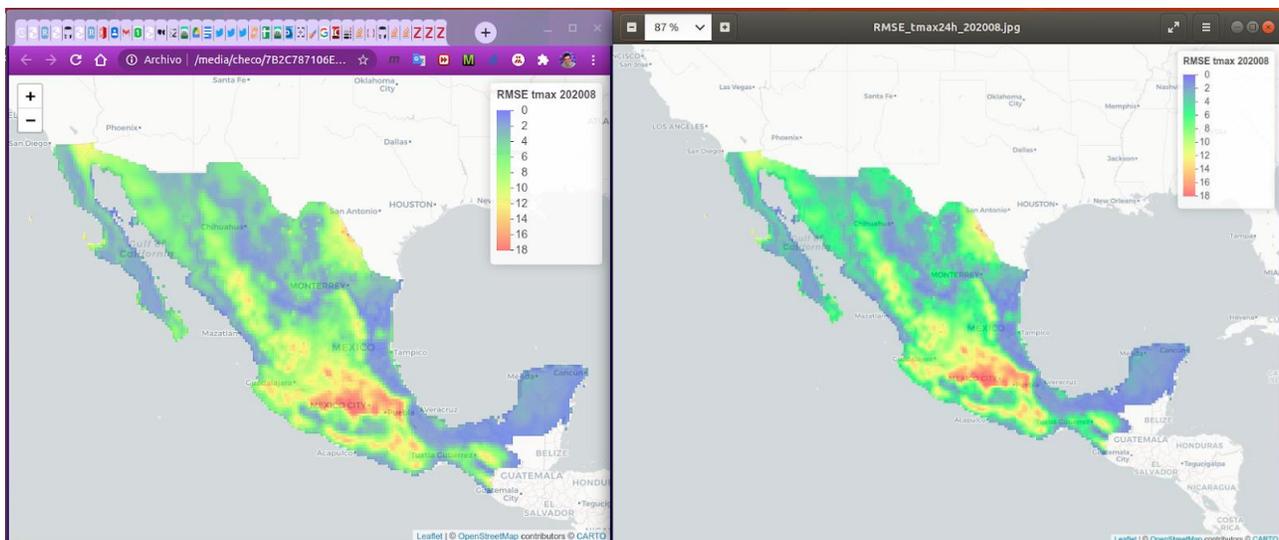


Figura 4.5. Mapas HTML y jpg del RMSE para la variable Tmax correspondiente a un pronóstico a 24 horas para el mes de agosto 2020.

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 14 de 24	México, 2020	Clave: F.C0.2.04.01

5 EVALUACIÓN DEL PRONÓSTICO

5.1 Análisis estadístico por municipios

Análisis de sesgo medio del conjunto global de datos pronosticados contra observaciones.

- **Temperatura máxima.** En el periodo de 24 horas se presentó una subestimación del pronóstico con respecto a las observaciones en todos los meses analizados, con un rango que oscilaba entre -7 y -3°C aproximadamente. En lo correspondiente a 48 horas, se observan valores del pronóstico muy cercano a lo registrado en las observaciones durante los meses de marzo-mayo, mientras que de junio a agosto se presenta una subestimación del pronóstico en el rango de -1 y -2°C aproximadamente. Finalmente para el periodo de 72 horas se identifica una sobrestimación del pronóstico con valores de entre 1 y 8°C aproximadamente, para el total de meses analizados.
- **Temperatura mínima.** Para el periodo de 24 horas se presenta una subestimación del pronóstico en todos los meses del análisis, en un rango de -4 a -7°C aproximadamente. Para 48 horas continua de forma general una subestimación del pronóstico respecto a las observaciones, con valores de entre 2 y -3°C, a excepción del mes de abril donde el error fue cercano a cero, al igual que los primeros meses del periodo de 72 horas (marzo-mayo), mientras que para el resto del periodo analizado se observa una subestimación del pronóstico de entre -1 y -2°C.
- **Precipitación acumulada diaria.** El análisis para los intervalos de 24 y 72 horas muestra en general un sesgo negativo en el pronóstico para el total del periodo (marzo-agosto), con valores de entre -4 y -7 mm, y de 0 a -3 mm respectivamente, mientras que para el intervalo de 48 horas, de marzo a agosto, se observa una sobrestimación del pronóstico con valores de entre 1 y 9 mm aproximadamente.
- **Rapidez máxima del viento.** Esta variable en su análisis para el intervalo de 24 horas muestra valores cercanos a 0 km/h durante los meses de marzo a mayo, mientras que de junio a agosto se aprecia un sesgo negativo con valores de entre -1 y -2 km/h. Por su parte el análisis a 48 horas señala una sobrestimación del modelo para el total de periodo de evaluación, con valores de 2 a 8 km/h aproximadamente. Finalmente, el análisis del intervalo a 72 horas arroja un sesgo negativo con magnitud de 0 a -3 km/h, lo que refleja una subestimación de los valores pronosticados.

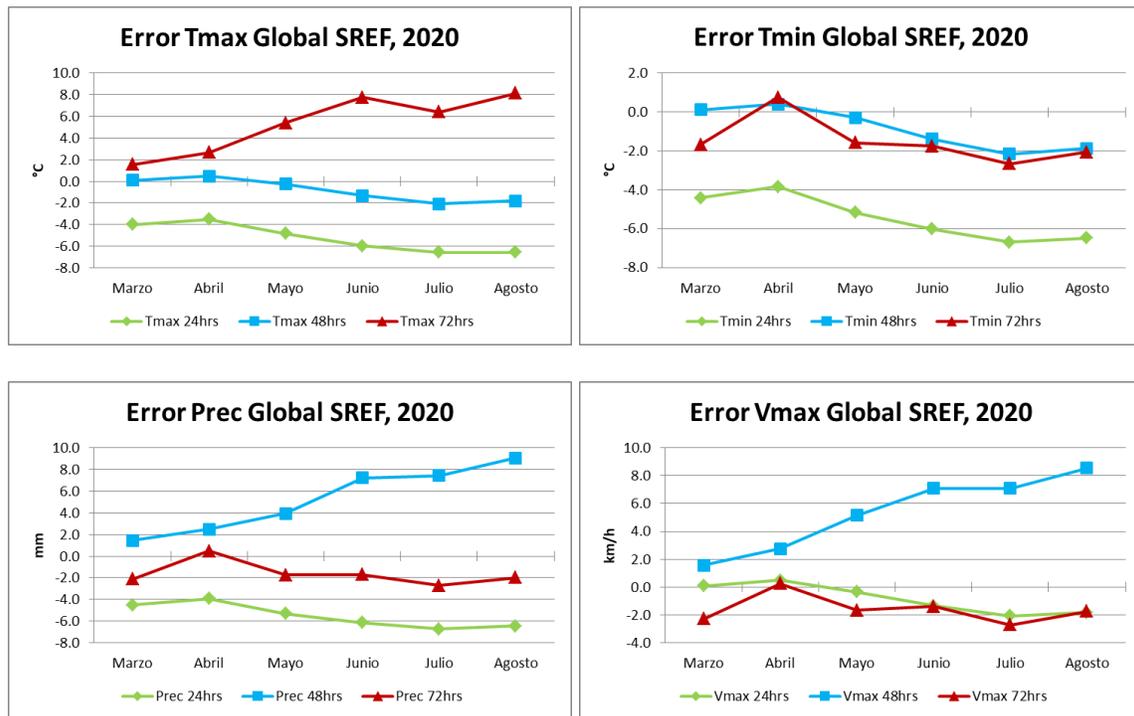


Figura 5.1. Error promedio para el total de municipios analizados, por variable.

En resumen, para el pronóstico a 24 horas se observó una subestimación en general para el total de variables, mientras que para el pronóstico a 48 horas, los valores del sesgo se encontraron cercanos al cero para el caso de las temperaturas y de forma positiva para la precipitación y viento máximo; finalmente para el pronóstico a 72 horas, para las variables temperatura mínima, precipitación y viento máximo, se observan valores de error cercanos al cero así como una ligera subestimación del pronóstico, mientras que para el caso de la temperatura máxima se presenta una sobrestimación con respecto a las observaciones registradas.

Análisis del RMSE medio del conjunto global de datos pronosticados contra observaciones

- Temperatura máxima. El análisis de esta variable muestra en lo general valores de error menores para los intervalos de 24 y 48 horas, siendo el último el de menor magnitud, dentro del rango de +/- 4 y 5°C aproximadamente. Por su parte, el intervalo de 72 horas exhibe en lo general (a excepción de los meses de marzo y abril) los valores más altos de error en el periodo de estudio, de entre +/- 8 y 12°C.
- Temperatura mínima. El error asociado al pronóstico de temperatura mínima indica los valores más bajo para el intervalo de 48 horas, de entre +/- 4 y 5°C, mientras que para los intervalos de 24 y 72 horas se presentaron valores de error de entre +/- 5 y 8°C.
- Precipitación acumulada diaria. Para el caso de la precipitación, el pronóstico para los intervalos de 24 y 72 horas presentan una señal similar, a lo largo del periodo analizado,

con valores de entre +/- 6 y 8 mm, mientras que para el periodo de 48 horas se presentan los valores de error más bajos durante los meses de marzo a mayo (+/- 2 a 6 mm), caso contrario para el resto de los meses donde las magnitudes de error son las mayores de todo el periodo, de entre +/- 10 y 13 mm.

- Rapidez máxima del viento. Finalmente para esta variable se observa de forma similar una señal consistente entre los intervalos a 24 y 72 horas que oscilan entre valores de error de +/- 4 a 7 km/h, mientras que para el intervalo de 48 horas, a excepción de marzo y abril (error de +/- 2 a 4 km/h), se observan los valores más elevados de error de todo el periodo analizado, con magnitudes de entre +/- 8 y 12 km/h.

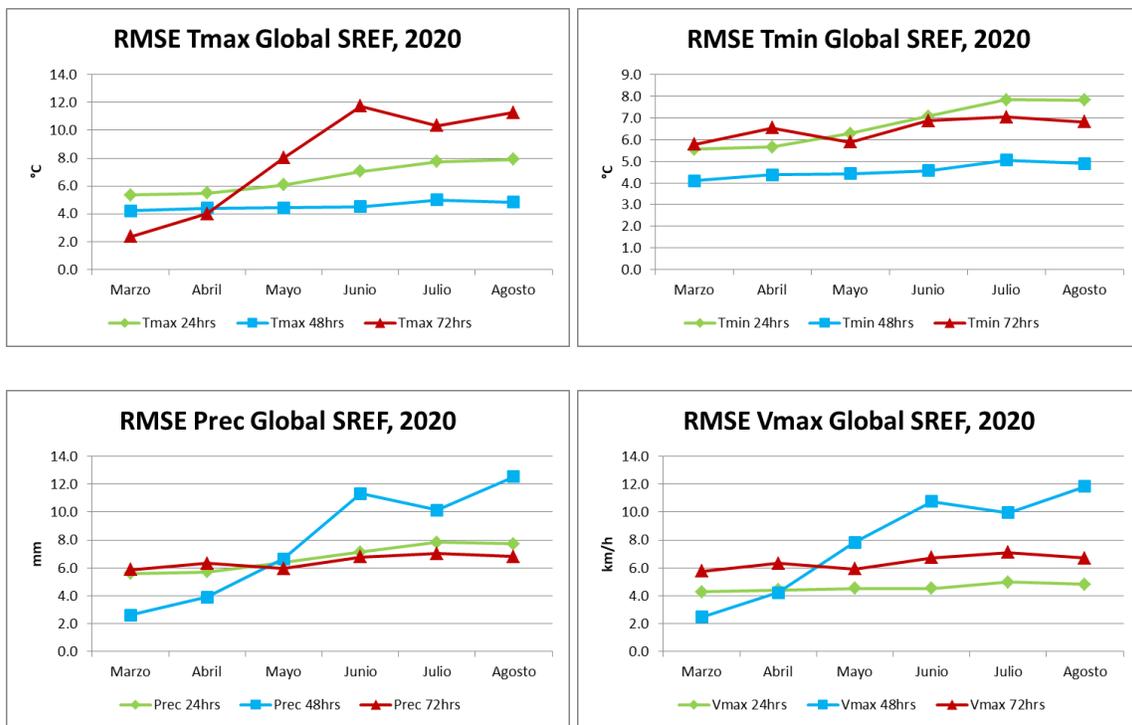


Figura 5.2. RMSE promedio para el total de municipios analizados, por variable.

Como síntesis, se observó que para el caso de las temperaturas máximas y mínimas, en general, los intervalos de 48 horas presentaron un menor error del pronóstico con respecto a las observaciones, con valores de entre +/- 4 y 5°C, mientras que los intervalos de 24 y 72 horas se encontraron en el rango de +/- 5 y 12°C. Por su parte para el caso de precipitación y viento máximo, se observaron de forma general y consistente los valores de error más bajos para los intervalos de 24 y 72 horas, con magnitudes de entre +/- 6 y 8 mm para la lluvia acumulada, y +/- 4 a 7 km/h para el caso del viento máximo, mientras que para el intervalo de 48 horas, a excepción de marzo, abril y mayo (error de +/- 2 a 6 mm y +/- 2 a 8 km/h respectivamente), se observaron los valores más elevados de error, con magnitudes de +/- 10 a 12 mm y km/h respectivamente.

5.2 Análisis a nivel nacional

En este apartado se muestran los mapas a nivel nacional que muestran la evaluación estadística de los ensambles con respecto a observaciones.

- **Temperatura máxima.** El análisis de RMSE para los intervalos de 24, 48 y 72 horas, de forma general muestra valores de entre 2 y 20°C para el territorio mexicano, con máximos predominantes sobre el norte y zonas altas de la mesa central y sureste de México, lo cual puede deberse a errores asociados a la construcción de la malla regular en regiones con una menor densidad de observaciones. Mientras que en resto del país se observan valores generales de RMSE de entre 2 y 8°C, principalmente en zonas costeras, donde se observan las magnitudes más bajas.

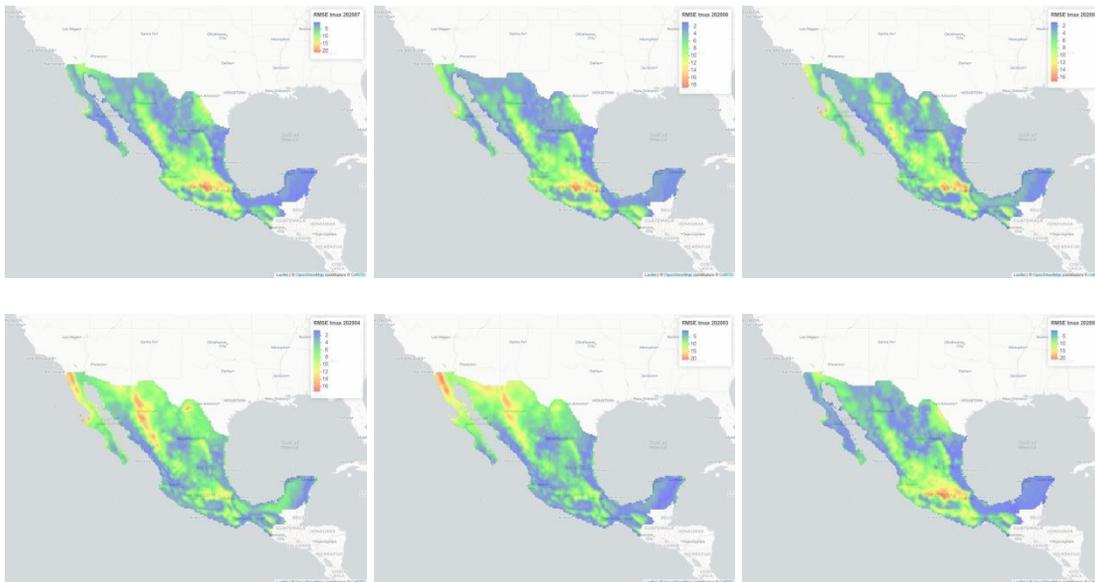


Figura 5.3. RMSE promedio por mes (marzo – agosto 2020) para temperatura máxima.

- **Temperatura mínima.** El análisis de los intervalos a 24, 48 y 72 horas, de forma general presentan valores de RMSE de entre 2 y 14°C para el total de periodo analizado, mientras que la distribución espacial de dichas magnitudes se presentan de forma similar al caso de la temperatura máxima, con máximos localizados sobre el norte, noroeste, zonas altas de la mesa central y sureste mexicano, donde este último se extiende hasta la porción de la Península de Yucatán, mientras que los valores más bajos se observan hacia regiones costeras, zonas de valle de la mesa norte y central mexicana. La distribución de las magnitudes de RMSE puede verse influenciada por errores asociados a la construcción de la malla regular en regiones con una menor densidad de observaciones.

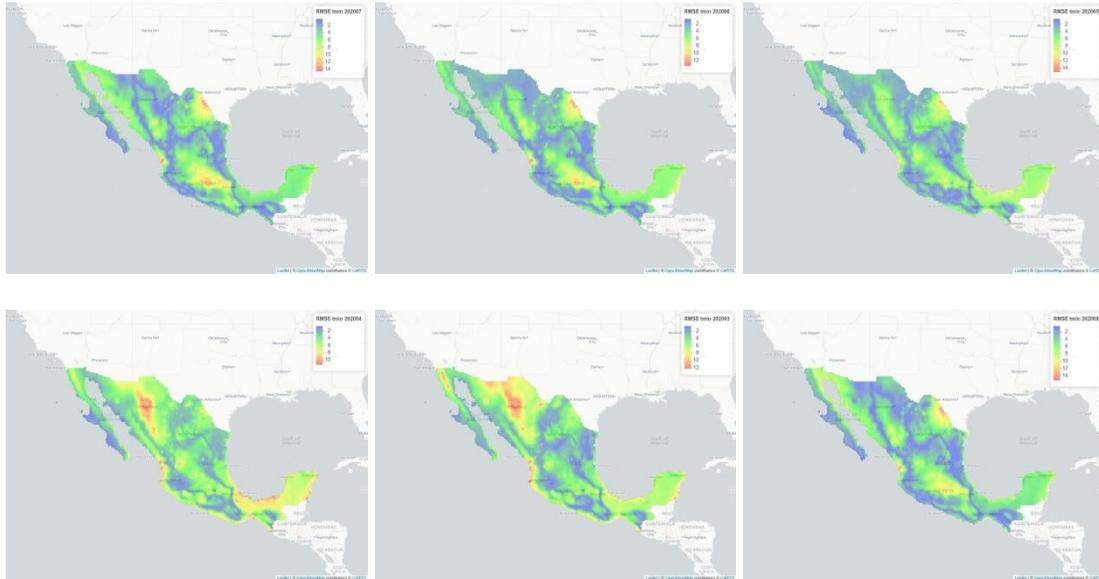


Figura 5.4. RMSE promedio por mes (marzo – agosto 2020) para temperatura mínima.

- Precipitación acumulada diaria. Se observa, de manera general, la existencia de valores de entre 10 y 80 mm, con una distribución similar para el mismo mes entre los distintos intervalos, pero con variación en las magnitudes, por lo que se presentan, de forma localizada, los valores más altos sobre la Sierra Madre Occidental del país, así como en zonas altas de la mesa central y sur-sureste mexicano principalmente, mientras que en el resto del territorio nacional se observan magnitudes menores de RMSE. Esta configuración es atribuible a que en dichas regiones es donde se presenta la mayor actividad de lluvias en México durante los meses analizados.

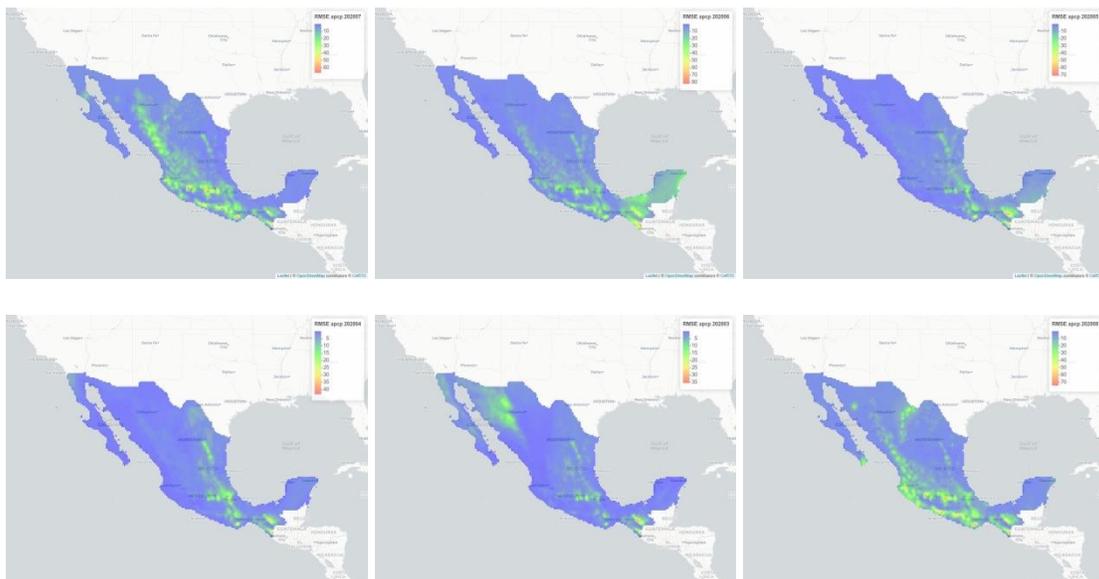


Figura 5.5. RMSE promedio por mes (marzo – agosto 2020) para precipitación diaria.

- Rapidez máxima del viento. Finalmente, en lo que respecta al análisis de RMSE del viento máximo, presentado durante el periodo de análisis, se puede observar de forma general que, a excepción de una particular región en la mesa norte del país, el resto del territorio nacional presenta valores que oscilan de entre 10 y 30 km/h, con las magnitudes más altas desplegadas sobre las porciones norte, noreste y sureste del territorio mexicano, lo que puede estar asociado a los sistemas de circulación local y de mesoescala que se presentan durante primavera y verano en México. Es importante también mencionar que la distribución de las magnitudes del error pueden verse influenciadas a una menor densidad de observaciones empleadas para el análisis de esta variable (por cuestiones de disponibilidad de los organismos operadores), que resulta en errores asociados a la construcción de las mallas regulares para la evaluación.

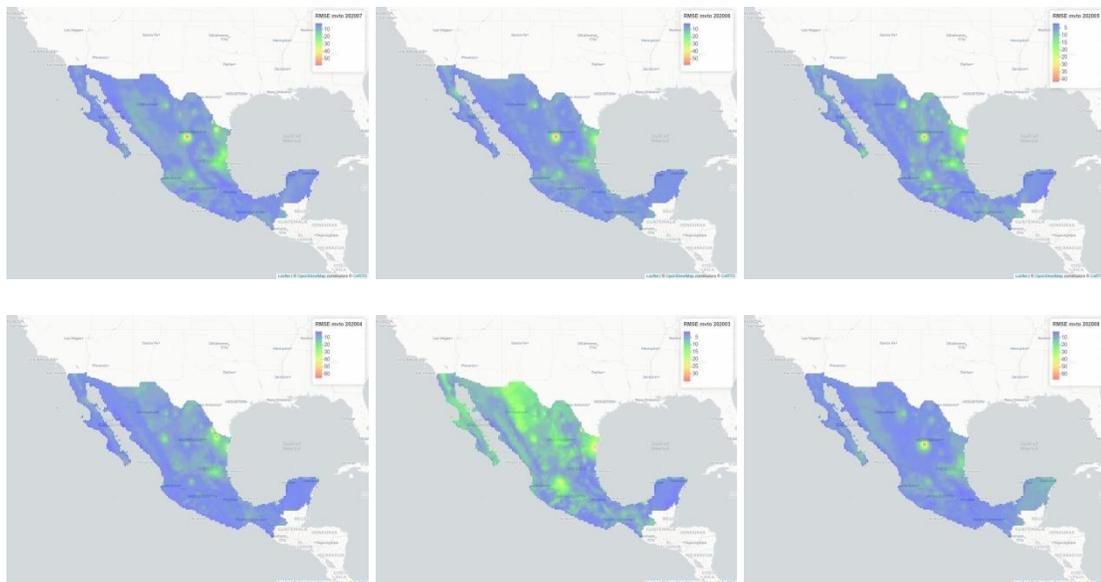


Figura 5.6. RMSE promedio por mes (marzo – agosto 2020) para viento máximo.

6 PUBLICACIÓN DE PRONÓSTICOS Y EVALUACIÓN

Para la publicación de los resultados de este proyecto se llevó a cabo el desarrollo de esquemas dinámicos visuales para la consulta del pronóstico por municipio, así como para el análisis y despliegue de información relacionada con la evaluación estadística espacial y temporal de los pronósticos por ensambles. Los elementos desarrollados están contenidos en dos módulos principales “Pronóstico” y “Evaluación”, los cuales fueron contenidos en una plantilla de página web con formato institucional (<http://galileo.imta.mx/BOLENS/>) como se muestra en la Figura 6.1.

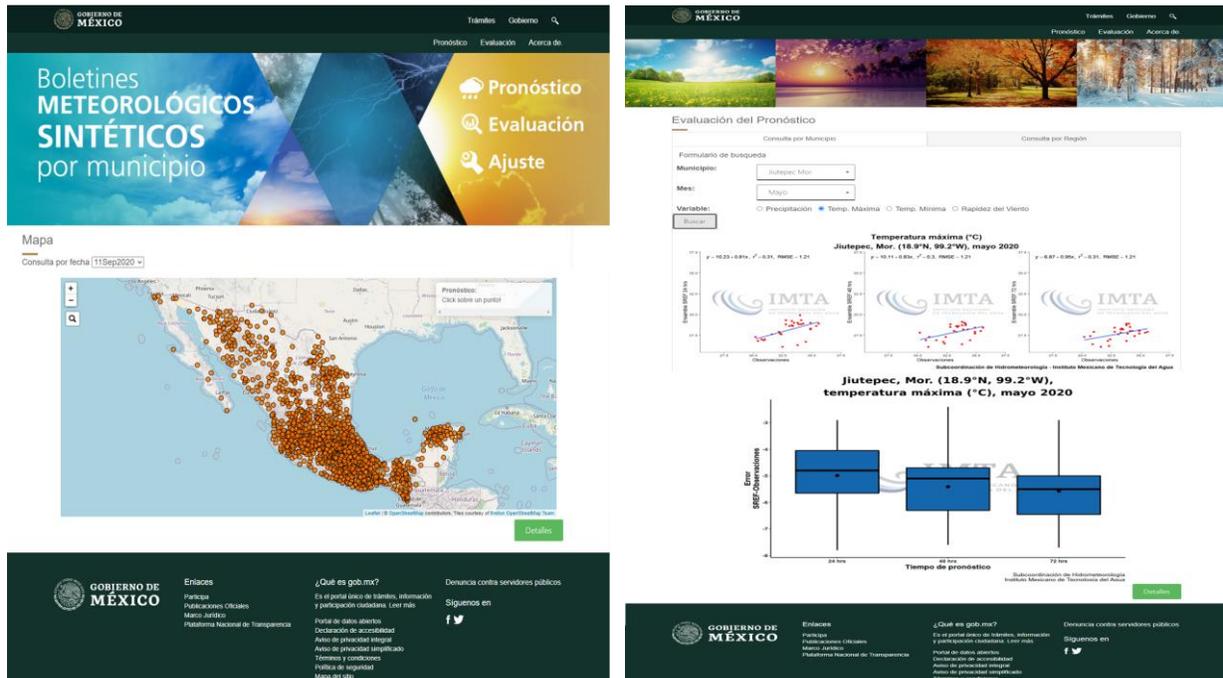


Figura 6.1. Sistema para la visualización de resultados, para los módulos pronóstico y evaluación respectivamente.

6.1 Módulos de visualización

En el módulo de pronóstico se muestran resultados operacionales del sistema de generación de boletines meteorológicos sintéticos en modo de consulta rápida de variables meteorológicas pronosticadas, o a través de descarga de los reportes escritos. Los elementos de consulta desarrollados permiten al usuario elegir el municipio mediante un campo de llenado, así como interactuar con el mapa donde se localizan todos los municipios señalizados.

Módulo de pronóstico. Se requirió utilizar un gestor de mapas con la República Mexicana contenida, para visualizar marcadores que representan a cada uno de los municipios que cuentan con reportes meteorológicos sintéticos diarios, de acuerdo a la información obtenida del sistema.

Criterios:

- Establecimiento de encabezado para el módulo de pronóstico.
- Configuración de un motor de búsqueda por fecha a partir del día actual y 6 días posteriores (alcance del pronóstico).
- Muestra de cada municipio y globos emergentes de información relativa al pronóstico meteorológico, al momento de dar click sobre un marcador del mapa.

- Implementación del motor de búsqueda por municipio dentro del gestor de mapas.
- Descripción de los detalles del producto, mediante elementos desplegables.

Por su parte, en el módulo de evaluación se presentan los resultados de la verificación estadística del desempeño del pronóstico meteorológico sintético por ensambles, para cada uno de los municipios y para la región total de México.

Módulo de evaluación. Se solicitó contener dos tipos de búsqueda.

Criterios:

- Opción de búsqueda por municipio, con el fin de mostrar graficas concernientes a la evaluación del desempeño de los pronósticos meteorológicos, utilizando parámetros de evaluación como diagramas de dispersión, Correlación y RMSE medio, de acuerdo a filtros por variables meteorológicas: viento, precipitación, temperatura máxima y mínima.
- Opción de búsqueda por región, con el fin de mostrar graficas concernientes a la evaluación del desempeño de los pronósticos meteorológicos para el territorio mexicano, utilizando el RMSE medio para el conjunto total de datos, y de acuerdo a filtros por variables meteorológicas: viento, precipitación, temperatura máxima y mínima.
- Descripción de los detalles de cada producto, mediante elementos desplegable.

A continuación se describen las tecnologías empleadas para la implementación de los elementos visuales dinámicos y configuración del proyecto en el servidor:

Tecnologías	Descripción
PHP 5.6	Es un lenguaje especialmente adecuado para realizar conversiones de ficheros y búsquedas en el servidor, compatible con sitios web.
JavaScript	Lenguaje de programación de lado del navegador, para manejar eventos de búsqueda dentro del servidor.
Apache 2	Servidor para alojar paginas html, php, de todo el sitio y poder ser vistas en la nube.
Bootstrap 3	Framework para acondicionar la interfaz del proyecto, que muestra la vista al contenido temático.

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 22 de 24	México, 2020	Clave: F.C0.2.04.01

Leaflet	La biblioteca JavaScript de código abierto utilizada para crear mapas interactivos en el área de trabajo
Servidor Linux	Servidor físico llamado galileo. Donde se encuentra todo el proyecto

6.2 Desarrollo de los módulos de visualización

Conversión de pronósticos meteorológicos por día de ficheros “txt” a tipo “GeoJson”, que emplea como insumo datos de longitud, latitud e información del pronóstico. Para esto se necesita acceder diariamente al contenido del fichero pronostico.txt, generado por el sistema de pronósticos, mediante el lenguaje PHP estructurando cada línea de acuerdo a los requerimientos, y así obtener como salida un nuevo fichero tipo GeoJson.

Acondicionamiento de la interfaz del proyecto de acuerdo a la guía de estilo para gov.mx <https://www.gob.mx/guias/grafica>; además de los módulos que lo componen pronóstico, evaluación, y acerca de.

Utilización de un gestor de mapas para despliegue de información por municipios. En esta sección se ocupó la Liberia Leaflet, a la cual se le integra mediante programación JavaScript con el fichero GeoJson y que ambas tecnología van incluidas dentro del módulo de pronóstico, perteneciente al proyecto. Adicionando lo requerido como consulta por fecha, búsqueda por municipio y la información del pronóstico.

Utilización de un gestor de resultados en formatos gráficos a partir de la base de datos de la evaluación estadística del desempeño de los pronósticos meteorológicos. Este sistema de control cuenta con elementos de búsqueda por municipio, mes y variables, así como despliegue de gestor de mapas para el análisis de resultados regionales.

Finalmente se llevó a cabo la configuración de servidor Linux para en actividades de descarga, análisis y generación de información relacionada con el proyecto en el sitio, empleando los elementos mencionados, imágenes, estilos, documentos pdf, textos y gestor de mapas. El servidor Galileo corresponde a un Dell Power Edge T610 con doble procesador Xeon a 2.8 Ghz, 32 GB RAM y 400 GB de disco duro, el cual cuenta con RedHat Enterprise Linux como sistema Operativo y se encuentran habilitados los servicios de Apache como servidor web con soporte para PHP y MySQL como manejador de bases de datos, así como el conjunto de compiladores PGI y GrADS como herramienta para el análisis y despliegue de datos meteorológicos y climáticos, entre otras utilerías.

 SEMARNAT <small>SECRETARÍA DEL MEDIO AMBIENTE Y RECURSOS NATURALES</small>	Boletines meteorológicos sintéticos para municipios de México	 IMTA <small>INSTITUTO MEXICANO DE TECNOLOGÍA DEL AGUA</small>
Página 23 de 24	México, 2020	Clave: F.C0.2.04.01

7 CONCLUSIONES

El sistema de generación de boletines meteorológicos sintéticos ofrece la capacidad de contribuir a la disponibilidad de información especializada a escala local de forma operacional, lo que resulta útil para la evaluación de situaciones de peligro, atención de emergencias o para la programación de actividades, mientras que la integración de un modelo de ensambles a este producto constituye un elemento que contribuye a la identificación de la incertidumbre asociada al pronóstico. El esquema propuesto y los elementos que lo integran presentan una simplicidad y síntesis en su diseño, de tal forma que su revisión y análisis puede ser realizado de forma rápida y sencilla, por usuarios que no necesariamente requieren ser expertos en el tema.

Por su parte, la evaluación objetiva del pronóstico mediante la obtención de los errores asociados, brinda la oportunidad de identificar y atender posibles errores sistemáticos asociados al pronóstico, que una fase posterior de este trabajo, puedan ser empleados en esquemas de ajustes de sesgo con el propósito de mejorar la calidad de la información generada para la toma de decisiones, planeación de actividades, prevención de desastres y atención a emergencias.

8 BIBLIOGRAFÍA

- García N, Méndez K. y Vázquez F. (2018). Impacto socioeconómico de los principales desastres ocurridos en la república mexicana en 2016. Segob-Cenapred, Versión electrónica 2018, México. Tomado de:
<http://www.cenapred.gob.mx/es/Publicaciones/archivos/384-IMPACTO2016OEFINAL12FEBRERO2018.PDF>
- Cui B., Toth Z., Zhu Y. y Hou D. (2012). Bias Correction for Global Ensemble. Forecast. Weather and Forecasting. American Meteorological Society, Volume 27, Pg. 396-410
- Du, J., G. DiMego, M. S. Tracton, and B. Zhou 2003: NCEP short-range ensemble forecasting (SREF) system: multi-IC, multi-model and multi-physics approach. Research Activities in Atmospheric and Oceanic Modelling (edited by J. Cote), Report 33, CAS/JSC Working Group Numerical Experimentation (WGNE), WMO/TD-No. 1161, 5.09-5.10
- Guijarro J. A. (2018). Homogeneización de series climáticas con Climatol. Agencia Estatal de Meteorología, Guía versión 1.3.1, Islas Baleares, España.
- Jolliffe, I. T., y Stephenson D. B. (2012). Forecast Verification. A Practitioner's Guide in Atmospheric Science. 2nd Edition, Wiley and Sons Ltd, 292 pp.
- Stensrud, D.J., Bao J-W y Warner T. T. (2000). Using Initial Condition and Model Physics Perturbations in Short-Range Ensemble Simulations of Mesoscale Convective Systems. Monthly Weather Review, No. 128, pp. 2077-2107.
- Slingo, J y Palmer D. (2011). Uncertainty in weather and climate prediction. Phil. Trans. R. Soc. A. No. 369, Pp. 4751-4767.
- Wilks, D. S. (2006). Statistical Methods in the Atmospheric Sciences. 2nd Ed. International Geophysics Series, Academic Press, Elsevier, 627 pp.