



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

---

---

PROGRAMA DE POSGRADO EN CIENCIAS DE LA TIERRA

PRONÓSTICO DE LLUVIA  
SOBRE LA CUENCA DEL VALLE DE MEXICO

TESIS

QUE COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS DE LA TIERRA

PRESENTA

JUAN RUBÉN GONZÁLEZ CÁRDENAS

JURADO EXAMINADOR:  
ERNESTO CAETANO  
ÓSCAR FUENTES MARILES  
GUILLERMO MONTERO MARTÍNEZ  
BALDEMAR MÉNDEZ ANTONIO  
DAVID M. ZERMEÑO DÍAZ



México D.F

Agosto, 2011.



Universidad Nacional  
Autónoma de México



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## Índice general

1..	<i>Introduccion</i> . . . . .	1
2..	<i>El riesgo</i> . . . . .	4
	2.0.1. El enfoque ante los desastres en México . . . . .	7
	2.1. El Riesgo en la Ciudad de México por lluvias extremas . . . . .	9
	2.1.1. La Amenaza . . . . .	9
	2.1.2. La Vulnerabilidad . . . . .	14
	2.1.3. Un enfoque científico para la prevención de desastres por eventos hidrológicos . . . . .	15
3..	<i>Datos y Metodología</i> . . . . .	17
	3.1. Datos Meteorológicos . . . . .	17
	3.1.1. Datos de precipitación . . . . .	17
	3.1.2. Control de calidad de los datos del SACM . . . . .	19
	3.1.3. Datos de precipitación estimados por satélite . . . . .	25
	3.1.4. Metodología de interpolación de datos en una malla regular . . . . .	26
	3.2. El Modelo MM5 (Mesoscale Model Version 5) . . . . .	37
	3.2.1. Breve Historia del Pronóstico Numérico . . . . .	37
	3.2.2. Los datos para el MM5, condiciones para el experimento	41
	3.2.3. Desempeño del MM5 en pronósticos de lluvia para el Valle de México . . . . .	44
	3.3. MOS y Modelación Estadística. . . . .	45
	3.3.1. El Climate Predictability Tool (CPT) . . . . .	47
	3.3.2. Adaptación del CPT para el pronóstico de eventos de lluvia horarios . . . . .	49
	3.3.3. Funcionamiento de CPT . . . . .	51
	3.3.4. Calibración del Modelo Estadístico y Pronóstico Cuan- titativo. . . . .	52

---

4..	<i>Resultados</i> . . . . .	55
4.0.5.	Verificación de pronóstico cuantitativo . . . . .	55
4.1.	Análisis de la distribución del área de lluvia . . . . .	57
4.2.	Análisis de la distribución del patrón espacial de lluvia acumulada horaria . . . . .	64
4.2.1.	El problema de la Intensidad. . . . .	64
4.2.2.	El Patrón de lluvia ajustado por el modelo estadístico en la region del Distrito Federal . . . . .	65
4.2.3.	Precipitación local y criterio de comparación . . . . .	73
4.2.4.	Análisis de eventos extremos . . . . .	82
5..	<i>Conclusiones</i> . . . . .	89
5.0.5.	Desempeño del Pronóstico Cuantitativo de Precipitación . . . . .	89
5.0.6.	Area de lluvia, desempeño del QPF . . . . .	90
5.0.7.	Patrón de lluvia, similitud entre patrones . . . . .	90
5.0.8.	Trabajo a Futuro . . . . .	91
6..	<i>Bibliografía</i> . . . . .	93
7..	<i>Indice de Figuras</i> . . . . .	101
8..	<i>Abreviaturas</i> . . . . .	102

*Resumen*

La Ciudad de México está expuesta a diversas amenazas de orden hidrometeorológico, entre las que destacan las precipitaciones intensas. Cada año estas tormentas severas causan severos daños en la infraestructura y en la vida de la capital del país, y aunque la organización de las autoridades ante un posible desastre ocasionado por estos fenómenos ha cambiado en años recientes, todavía falta mucho por hacer en términos de lograr un manejo del riesgo por precipitaciones importantes en el Valle de México. La utilización de modelos numéricos para realizar pronóstico a corto plazo sobre áreas específicas permitiría identificar el patrón de lluvia sobre áreas concretas de la ciudad con el mayor índice de precipitación, lo que nos daría una idea de su vulnerabilidad a eventos de lluvia intensos. Dentro de los modelos utilizados en México para realizar pronóstico de precipitación diario en la ciudad se encuentra el modelo de área limitada MM5. Aunque el desempeño del modelo es aceptable para predecir ocurrencia o no de lluvia en un área como la Ciudad de México, los pronósticos presentan errores que se vuelven sistemáticos al momento de inferir la intensidad y la distribución de la precipitación en la ciudad. Estos errores se ven aumentados en la temporada de verano, donde las lluvias convectivas y los eventos de precipitación extrema son frecuentes. La presente investigación pretende mostrar que si se utiliza un modelo estadístico de corrección, sobre las salidas del MM5, es posible ajustar la distribución espacial de la precipitación horaria pronosticada por MM5 con respecto de campos de lluvia horarios observados en estación. Esta corrección ajusta el desempeño del pronóstico del MM5, mejorando la calidad del pronóstico hasta en 50 % para pronósticos de lluvia acumulada en temporada de verano. Para conocer la eficiencia de estos pronósticos ajustados, se realizó un Pronóstico Cuantitativo de Precipitación, a partir del cual se obtuvo un intervalo de confianza de entre el 30 % y el 60 % para los pronósticos ajustados de lluvia horaria sobre un área relativamente pequeña dentro de la ciudad. Aunque esta técnica en verdad mejora los campos de distribución espacial de lluvia pronosticada, la intensidad de lluvia obtenida no es la adecuada, al presentarse valores que están muy por debajo de la media de precipitación observada. Aunque una buena estimación de la intensidad sigue siendo una pregunta abierta en el campo de la modelación del clima y del tiempo, en nuestro caso particular, la subestimación de la intensidad puede deberse al poco tiempo usado para la calibración del modelo, o a cuestiones inherentes del algoritmo de cómputo utilizado, entre otras hipótesis.

*Abstract*

Mexico City is exposed to various threats of hydrometeorological order, among which are: heavy precipitations. Each year, severe storms cause severe damage to the infrastructure and life of the country's capital, and although the authorities' answer when facing possible disasters caused by these phenomena, has changed for good in recent years, there is still much to do in terms of achieving good risk management by significant rainfall in Mexico's Valley. In particular: the use of numerical models to produce short-term forecast over specific areas, allowing the possible identification of the rain pattern over certain areas into the city with the highest precipitation rates and, therefore giving the chance to assign their vulnerability to intense rain events. Among all the models used in Mexico for make daily precipitation forecasting over the city, one of the most used is the limited area model MM5. Although the model's performance is acceptable to predict the occurrence or not of rain in an area like Mexico City, the forecasts presents systematic errors when inferring the intensity and distribution of rainfall over the city; these errors are increased in the summer season, where convective rainfall and extreme precipitation events are frequent. This research aims to show that when using a statistical correction model, the outputs of MM5 can be adjusted respect to observed hourly rainfall fields even in summer season. This correction adjusts the MM5 forecast's performance, improving the forecast quality up to 50 % for some cases of accumulated rain in summer. To study the efficiency of these adjusted forecasts, we performed a Quantitative Precipitation Forecast, from which we obtained a confidence interval between 30 % and 60 % for hourly rainfall on a relatively small area within the city. Even this technique really improve the forecasted spatial distributions fields, the rain intensity is not the adequate. This because in the corrected fields, the intensity values are always below of the observed ones. Although a good intensity precipitation forecast still is an open question to the modeling field, in our special case, this intensity underestimation could be caused by the small amount of used data used for makes the model's calibration, or could be caused by a problem in the source code of the program used. Among other theories.

## 1. INTRODUCCION

La cuenca del Valle de México ha sido una región en donde por largo tiempo ocurren eventos extremos de precipitación. De acuerdo a datos de la Comisión Nacional del agua precipitaciones mayores a los 60mm en 24 horas no son extraños en la parte suroeste de la ciudad. El registro histórico máximo para la parte sur es de 164.9mm en 24 horas en septiembre 22 de 1986. Sin embargo, según diversos estudios citados en la literatura, los eventos de precipitación intensa son más frecuentes a partir de unas décadas para acá. (Jáuregui, 2000).

El crecimiento explosivo de la ciudad de México ha cambiado radicalmente el paisaje de la cuenca. Lo que era zona de bosques ahora son zonas urbanas. De acuerdo diversas investigaciones, el efecto de la isla de calor ha incrementado la precipitación convectiva (Jauregui y Romales, 1996). En áreas urbanas grandes como en la ciudad de México, los cambios en el albedo de la superficie pueden resultar en una superficie más cálida y por ende afectar el estado dinámico de la atmósfera en términos de su estabilidad, lo que da lugar a condiciones para que se ocurran tormentas intensas. En términos de emisión de contaminantes, más partículas en la atmósfera actuando como núcleos de condensación pueden también resultar en actividad convectiva mayor.

A lo largo de los años, en la ciudad de México numerosas familias se han establecido en áreas de riesgo, como riberas o acantilados, debido a diversos factores sociales de exclusión y pobreza, y a la omisión o negligencia de las autoridades encargadas de evitarlo. De esta manera aumenta la vulnerabilidad de ciertos sectores de la población ante eventos hidrometeorológicos extremos, en particular, los hidrológicos. Los desastres asociados con la precipitación intensa se ha convertido en un serio problema para las autoridades encargadas de la protección civil y del manejo urbano del agua. El proceso de planeación ante eventos de esta naturaleza, surge como una respuesta a los cambios de la sociedad y del medio ambiente, ya que el riesgo se construye socialmente, en el espacio y en el tiempo.

---

Hoy en día los esquemas de prevencion incluyen el análisis diagnóstico y la predicción con al menos 12 horas de anticipación. La mayoría de la predicción del clima y el tiempo en la ciudad de México está basado en métodos empíricos y rara vez son utilizados para tomar medidas preventivas. En el mejor de los casos, el monitoreo de la precipitación es usado para informar a la población y para explicar la naturaleza de un desastre, pero no proponer diagnósticos del riesgo inherente a estos eventos, en zonas con alta vulnerabilidad.

En las actividades de preparación ante un desastre es necesario conocer las características temporales y espaciales para determinar cuando una condición particular constituye un evento extremo y una amenaza para un sector de la población. Sin embargo, éste es sólo uno de los elementos que un esquema completo de mitigación y respuesta temprana debe incluir. El otro elemento importante es un modelo numérico de predicción que sea capaz de determinar con anticipación que regiones pueden ser afectadas por clima y tiempo severo. El estado del arte de los modelos de predicción de mesoescala es tal que la física utilizada en ellos permite resolver las ecuaciones que gobiernan a la atmósfera de manera aceptable, Por otro lado tienen siempre errores sistematicos que se presentan con mayor intensidad conforme se aumenta la resolución de sus pronósticos en el espacio y en el tiempo.

### *Objetivo*

En esta tesis se busca proponer una metodología que permita reducir los errores sistemáticos en las salidas de pronóstico de lluvia acumulada del modelo de mesoescala MM5 (Dudhia, *et al* 1999) utilizando un esquema de reducción de escala y corrección de errores sistemáticos conocida como Modelo de Salida Estadístico (MOS por sus siglas en inglés) y se busca averiguar el desempeño de sus salidas corregidas para el pronóstico de precipitaciones intensas a corto plazo (horas) y a la máxima resolución espacial posible permitida en el MM5, a partir de la asimilación de datos de lluvia horaria observados a una malla de 8 km de resolución. Con las salidas corregidas, se generará un Pronóstico Cuantitativo de Precipitación (QPF, por sus siglas en inglés) que consiste en pronosticar la cantidad de precipitación acumulada en un periodo de tiempo determinado para un área específica (Bushong,



2005). La calidad del QPF será evaluada al comparar la lluvia acumulada pronosticada con la lluvia observada en las mismas zonas (Rossa *et al*, 2007) utilizando dos dominios distintos: una malla de 10 x 8 puntos y una malla de 2 x 2 puntos, que equivalen a un área total de 160 x 128km y 16 x 16km respectivamente. Con esta comparación, es posible asignar un puntaje de efectividad que nos dará una idea clara del desempeño del QPF realizado.

### *Estructura de la tesis*

En el capítulo 2 se discute la definición de riesgo y se analiza la necesidad de adaptar esta definición a un contexto social, se hace un recuento de los desastres en el mundo y se hace énfasis de los desastres asociados a eventos hidrometeorológicos en México. Se analizan a los componentes que definen al riesgo en la Ciudad de México ante eventos de lluvia acumulada y el enfoque actual ante estos desastres de parte de las autoridades . Además de que se da cuenta de las posibilidades de utilizar modelos numéricos de predicción atmosféricos para elaborar un esquema de gestión del riesgo. En el capítulo 3, se realiza una breve introducción del estado de los datos climáticos en México y se eligen las dos bases de datos de lluvia acumulada horaria a utilizar, y se hace un análisis detallado para su interpolación a una malla regular. También se da una breve reseña del pronóstico numérico del tiempo y del clima en México, así como una introducción al modelo de mesoescala PSU/NCAR, mejor conocido como MM5 (Bushong, 2005). Se hacen notar los errores sistemáticos del MM5 en simulaciones de lluvia acumulada en 24 horas sobre el Valle de México. En el capítulo 4 se justifica la elección de la herramienta utilizada para realizar la corrección de errores sistemáticos en las salidas del modelo numérico (CPT, *Climate Prediction Tool*), siguiendo un esquema tipo MOS (Model Output Statistics), y se realiza la adaptación de esta a datos horarios para la calibración del modelo estadístico. En el capítulo 5, se realiza el análisis de las salidas corregidas por el modelo estadístico utilizado realizando un QPF sobre dos dominios distintos; para verificarlo se definen dos categorías: a) análisis del pronóstico por áreas y b) por patrones de distribución de precipitación. En el capítulo 6, se presentan conclusiones sobre los resultados del trabajo considerando la ganancia de utilizar un esquema tipo MOS para corregir las salidas del MM5, también se presentan posibilidades de desarrollar trabajo futuro así como las implicaciones del estudio en el ámbito de la Protección Civil.

## 2. EL RIESGO

**E**l término desastre se refiere a sucesos que causan considerables pérdidas humanas, materiales, económicas o ambientales, alterando el funcionamiento de la comunidad afectada y excediendo su capacidad para hacer frente a la situación utilizando sus propios recursos (capacidad de *resiliencia*) (EIRD, 2007). John Whittow (1980) hace una distinción entre la amenaza o peligro (*hazard*) y un evento catastrófico (*disaster*); a la amenaza la considera como un *evento natural ya advertido que amenaza la vida y la propiedad* y al evento catastrófico lo considera como la materialización de esa amenaza. Cuando algún fenómeno natural (como un huracán o inundación) afecta significativamente la infraestructura, la economía y las condiciones de vida de la gente, hablamos de un desastre causado por un fenómeno natural (y no de un *desastre natural*).

*El riesgo se construye en el tiempo y en el espacio, por lo que la presencia de una amenaza (en este caso una perturbación natural) no implica la causa de un desastre si no está ligada a un contexto netamente social. Por lo que el término "desastre natural" es erróneo.*

Dichos eventos pueden provocar pérdidas de vidas humanas y costos financieros de consideración, con consecuencias a largo plazo que se manifiestan en la disminución de la calidad de vida de las poblaciones afectadas (problemas de salud pública, escasez de alimentos, pérdidas de fuentes de empleo o alimentos). Sin embargo, la definición de desastre anterior es incompleta e inadecuada, ya que por una parte genera la percepción de que todo fenómeno natural extremo es desastroso en sus efectos (que no ocurre en todos los casos) y por otra parte parece restar importancia al hecho de que la vulnerabilidad de las comunidades humanas a los desastres en gran parte tiene que ver con factores sociales. Así, el riesgo debido a la ocurrencia de fenómenos naturales

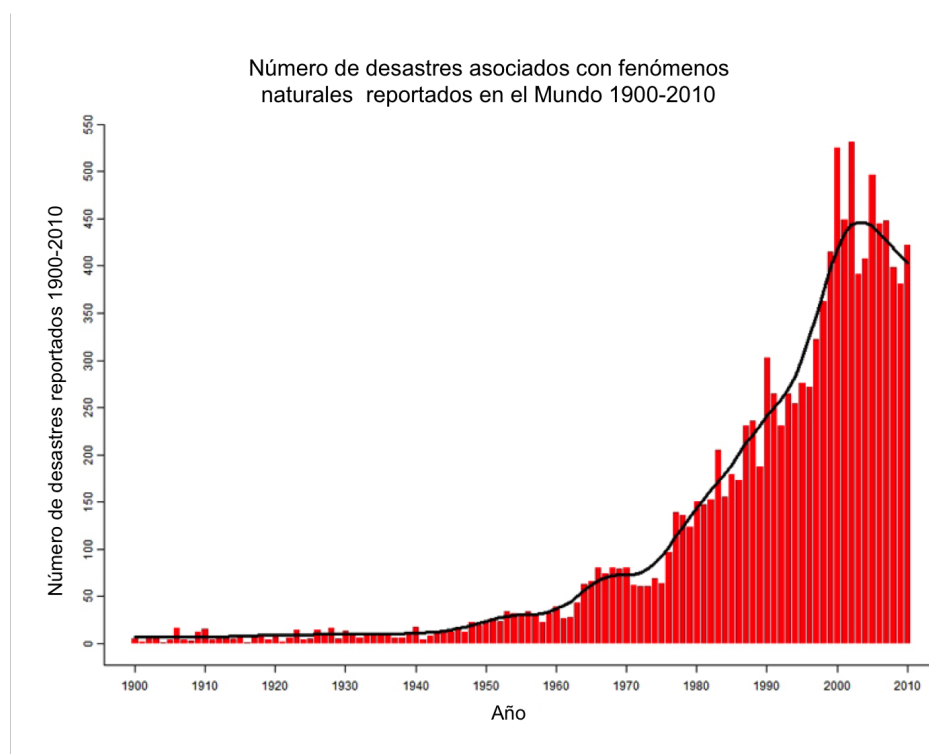
---

se puede determina como el producto de la probabilidad de que se presente el fenómeno (huracán, inundación, sequía, incendio, etc.) y del grado de exposición o vulnerabilidad de la comunidad propensa a ser afectado por el fenómeno en cuestión. (Giroto 2002, CENAPRED, 2001)

Durante el siglo XX en distintas partes del mundo se registraron aumentos, a veces dramáticos, del número de desastres provocados por perturbaciones naturales (PNUD 2004). Independientemente de los cambios climáticos que posiblemente están incrementando la frecuencia y magnitud de muchas de estas perturbaciones, el aumento de los desastres provocados por fenómenos naturales también se debe a una mayor vulnerabilidad de la población humana que se establece en áreas de riesgo (laderas inestables, llanuras de inundación, zonas de matorrales o bosques susceptibles a incendios) y a la degradación de ecosistemas (los manglares, arrecifes de coral, que podrían mitigar el efecto de ciertos fenómenos naturales causantes de desastres (Ries *et al* 2004). Dado que en los países industrializados existe una mayor infraestructura, sus pérdidas económicas por desastres son más elevadas en términos absolutos (Miletti, 1999). Sin embargo, debido a que estos países cuentan con más recursos institucionales, humanos y financieros, normalmente resisten mejor y se recuperan en menos tiempo después de un evento catastrófico. En cambio, en los países en desarrollo, o incluso en las regiones o barrios pobres de países industrializados, una de las causas principales de que los efectos de las perturbaciones de los fenómenos naturales sean mayores, es la vulnerabilidad extrema de dichas comunidades a dichos eventos. Un ejemplo muy ilustrativo y reciente es la inundación de la ciudad de Nueva Orleans, en los Estados Unidos, durante septiembre de 2005 debido al paso del huracán Katrina. El aumento de la vulnerabilidad de los centros de población humanos a los desastres provocados por fenómenos naturales se debe a un conjunto complejo de factores, todos ellos de índole social y económico, entre los cuales podemos mencionar: la desigualdad de la distribución de la propiedad de la tierra y la especulación inmobiliaria que lleva a los más pobres a ocupar zonas marginales de alto riesgo (márgenes de ríos y zonas costeras, laderas de montañas, etc.) el crecimiento poblacional acelerado y la migración rural a las ciudades, la expansión urbana y desordenada, la falta de planeación y regulación del uso de suelo, la ocupación habitacional en sitios silvestres, y a la deficiente información sobre medidas preventivas y el desconocimiento (a veces total) de como enfrentar los fenómenos naturales más comunes. Debido a este alto grado de vulnerabilidad, en promedio hubo

13 veces más víctimas y 20 veces más daños económicos (en relación con el PIB) en países en desarrollo en comparación con las naciones más ricas durante el periodo 200-2004 (CENAPRED 2001, PNDU 2004).

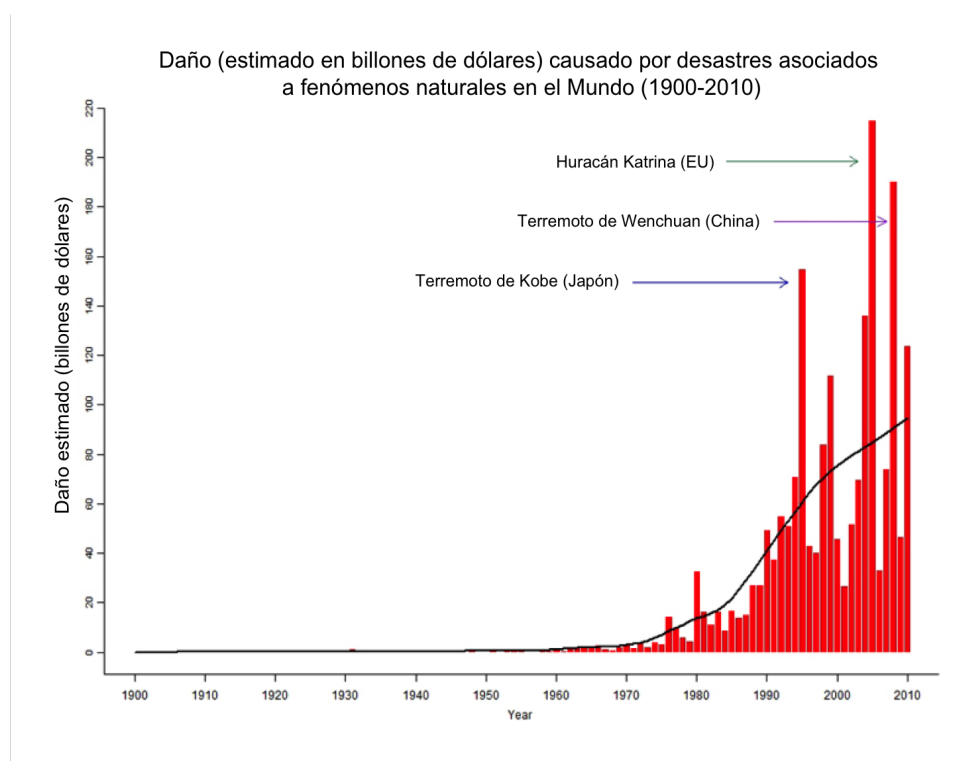
En el mundo, el número de desastres, principalmente de tipo hidrometeorológico, ha crecido en los últimos años (Fig. 2.1). La ocurrencia en años recientes de deslaves, inundaciones, lluvias intensas y granizadas, es más frecuente y costosa que las que se tienen de otros fenómenos geofísicos. De los más de 16,000 desastres que ocurrieron por algún fenómeno natural en países como México, El Salvador, Guatemala, Colombia, Ecuador, Costa Rica, Argentina y Panamá entre 1988 y 1998, un alto porcentaje fueron de origen hidrometeorológico.



**Fig. 2.1:** Número de desastres asociados con fenómenos naturales en el mundo durante el periodo 1900-2010. (*The International Disasters Data Base 2010*)

En el mundo, 296 mil 800 personas murieron en 373 desastres ocasionados

por fenómenos naturales en 2010, lo que ocasionó daños estimados en 109 mil millones de dólares (80 mil 400 millones de euros), según el Centro de Investigación sobre Epidemiología de Desastres de Naciones Unidas. Estos daños son tres veces superiores que en 2009. China, Chile y los Estados Unidos soportaron la mayor parte de los gastos, según la Organización de las Naciones Unidas (ONU).



**Fig. 2.2:** Daño estimado por desastres asociados con fenómenos naturales en el mundo durante el periodo 1900-2010. (*The International Disasters Data Base 2010*)

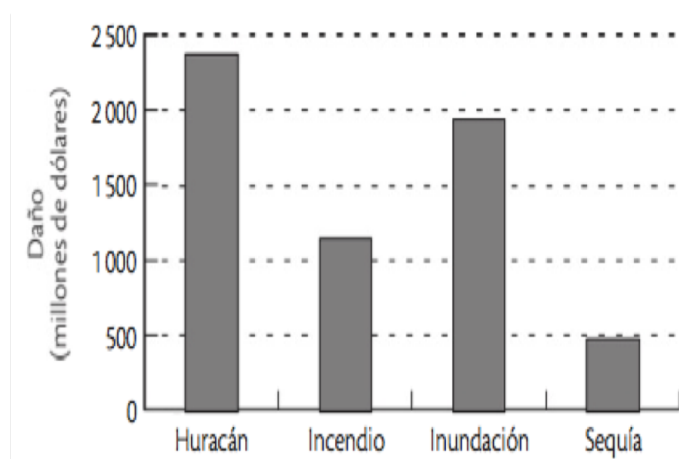
### 2.0.1. El enfoque ante los desastres en México

En México y otros países de América Latina el enfoque principal de respuesta a los desastres causados por fenómenos naturales afortunadamente ha cambiado de un esquema de reacción fundamentalmente, a uno de carácter preventivo, donde lo que se busca es entender mejor las amenazas y los facto-

res de riesgo y el desarrollo de métodos eficientes para detectarlas, predecirlas y mitigarlas. (CENAPRED 2001, Girot 2002). Aunque esto ha significado un gran avance para la prevención ante los desastres, el método aún considera al desastre como un evento puntual y de corta duración, mientras que se continúa ignorando el deterioro progresivo de las condiciones ambientales y sociales, lo que resulta en una percepción de los desastres como algo que es consecuencia solamente de las perturbaciones naturales al tiempo que minimiza el papel de los seres humanos como factor de incidencia en la frecuencia, severidad e impacto socioeconómico de estos eventos (Abramovitz 1999, Girot 2002).

*Aunque cada vez se hace más énfasis en prevenir los desastres causados por fenómenos naturales en México, el patrón actual de inversión es parecido al de América Latina en general, donde las estrategias de desarrollo y crecimiento económico resultan en el deterioro de los sistemas ambientales, lo que se relaciona directamente con el aumento de la vulnerabilidad de la población frente a estos eventos extremos*

Los costos ocasionados por perturbaciones naturales desastrosas en México son siempre muy elevados, de tal modo que la prevención (y por ende la reducción de sus efectos) debe ser una tarea prioritaria para los tomadores de decisiones. Por ejemplo: durante el periodo 1980-1999, los desastres por fenómenos naturales causaron un promedio de 500 muertes y cerca de 700 millones de dólares en daños (CENAPRED 2001). Si consideramos sus efectos indirectos, las cifras actualizadas hasta el 2003 sugieren que estos desastres causaron un total acumulado de pérdidas por 5,950 mil millones de dólares. Así mismo, estos desastres consumieron el 70 % de los recursos del Fondo Nacional de Desastres (Fonden) desde su creación en 1996 (26 562 millones de pesos gastados hasta el 2004, Inegi 2005).



**Fig. 2.3:** Costos económicos causados por los daños de desastres provocados por cuatro tipos de fenómenos naturales en el periodo 1980-2003. CENAPRED 2004

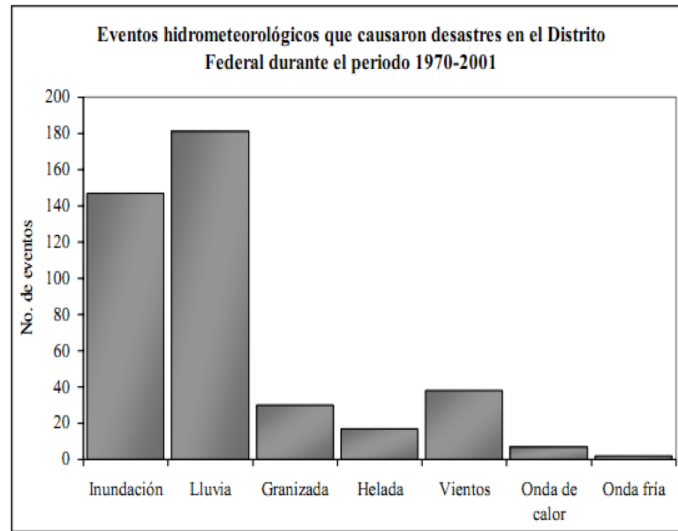
Como un ejemplo, podemos mencionar que aunque en 2008 no se registraron desastres de grandes dimensiones como las inundaciones en Hidalgo, Puebla, Veracruz y Tabasco en 1999, el ciclón tropical Isidore en 2002, los huracanes Emily, Stan y Wilma en 2005 o las inundaciones de Tabasco en 2007; si se registraron alrededor de 30 catástrofes con impactos locales que superaron, cada una, los cien millones de pesos. Así, el monto de daños osciló en 1,275 millones de dólares, colocando a 2008 como el tercer año que mayores pérdidas económicas ha padecido nuestro país en la última década, a consecuencia solamente de desastres de origen hidrometeorológico.

## 2.1. El Riesgo en la Ciudad de México por lluvias extremas

### 2.1.1. La Amenaza

Dentro de las amenazas de tipo hidrometeorológico más conocidas se encuentran: lluvias torrenciales o trombas, inundaciones, vientos intensos (incluyendo tornados) huracanes, granizadas, nevadas o heladas atípicas, sequías prolongadas, olas de calor, incendios forestales. Las tres primeras amenazas (lluvias torrenciales, inundaciones y vientos intensos) son las que con mayor frecuencia afectan al Valle de México (Fig 1.3 ), siendo los eventos de lluvia intensa los que han tenido mayor presencia y más daños han ocasionado

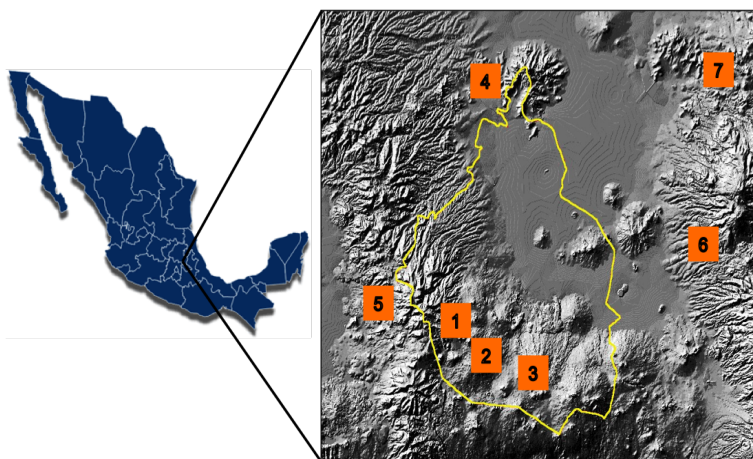
(en un periodo de 30 años), por lo que se puede considerar que los eventos hidrometeorológicos más peligrosos en el Valle de México son las lluvias intensas.



**Fig. 2.4:** Número de reportes periódicos sobre desastres de origen hidrometeorológico en el Distrito Federal entre 1970 y el 2001 (Fuente: *DesInventar 6. La Red*)

Debido a esto, y a riesgo de sonar reiterativo, es necesario llamar la atención acerca de dos elementos que se constituyen como parte importante en la construcción de la amenaza por eventos de lluvia intensa en el Valle de México: la orografía y su régimen de lluvia. El Valle de México está situado en la parte centro-sur del país, entre los 19° N de latitud y 99° W de longitud (aproximadamente a 2200 metros sobre el nivel del mar); la Ciudad de México se encuentra rodeada de montañas importantes entre las que destacan: al sur, la Sierra Chichinautzin y la Sierra del Ajusco; al suroeste, la Sierra de las Cruces, hacia el norte, la Sierra de Guadalupe; al oeste se encuentra el Valle de Toluca y al sureste los Volcanes Popocatepetl e Iztaccíhuatl (Figura 2.5). De este modo, el relieve juega un papel importantísimo como modulador de las circulaciones atmosféricas en el Valle de México. Por otro lado, una amplia variedad de uso de suelo, que va desde el tipo urbano, al bosque de coníferas y al de cubierta de hielo en la montaña, también actúan como forzantes de la circulación atmosférica local (Pérez 2001).





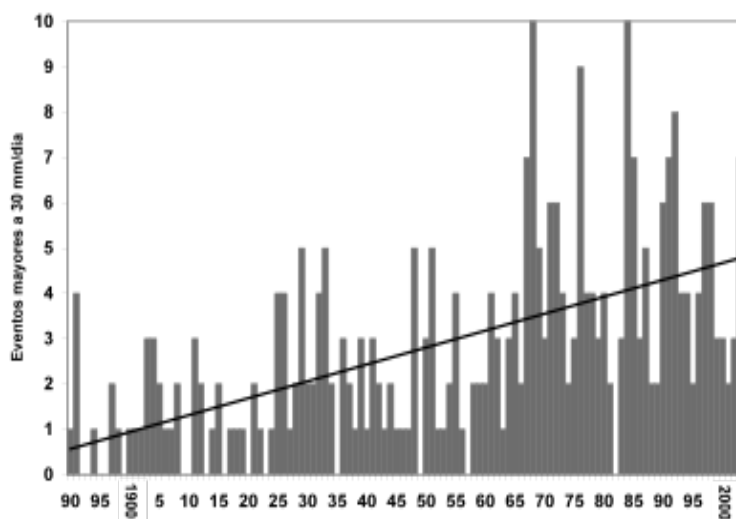
**Fig. 2.5:** Localización geográfica del Valle de México y principales accidentes orográficos. 1. S. de la Cruzes, 2. S. Chichinautzin, 3. S. del Ajusco, 4. S. De Guadalupe, 5. Valle de Toluca, 6. Volcán Popocatepetl y 7. Volcán Iztaccíhuatl

Las lluvias en la Ciudad de México son de tipo monzón (de mayo a septiembre). En comparación con la zona norte y oriente de la Ciudad, las zonas sur y suroeste presentan tanto el mayor número de días con lluvias (alrededor de 110 días) como la mayor cantidad de lluvia acumulada ( 900 mm/año). Las montañas del poniente y sur de la Ciudad, en combinación con el flujo de humedad provocan que, además de tener el mayor número de días con lluvias, también se presenten mas aguaceros (mayores a 20 mm/hr) (Pérez 2004).



**Fig. 2.6:** Precipitación media anual en la Cuenca de México. Periodo: 1941-1995. (Tomada de Jáuregui, 2000)

En los últimos cien años las precipitaciones en la Ciudad de México experimentaron un incremento (Fig. 2.6). De acuerdo con los datos del Observatorio de Tacubaya, la precipitación pasó de alrededor de 600 mm por año a principios del siglo XX, a casi 900 mm por año para finales del mismo siglo. El aumento en la precipitación anual acumulada está asociado al hecho de que el número de eventos extremos (más de 20 mm/hr) también han aumentado, pasando de 1 o 2 por año a principios del siglo veinte, a 6 o 7 por año a finales del mismo.



**Fig. 2.7:** Eventos de precipitación en 24 horas mayores a 30 mm. en la estación Tacubaya. Periodo: 1890-2003. Tomado de: Pérez 2004

Este incremento en la precipitación, se le ha atribuido al aumento de la temperatura en la Ciudad de México debido el efecto llamado "isla de calor" (Jáuregui y Romales 1996). Bajo consideraciones solamente de tipo termodinámico, a mayor temperatura aumenta la capacidad de la atmósfera de contener vapor de agua por lo que la cantidad de agua precipitable es mayor. Aunado a esto, es posible que el cambio de uso de suelo que ha experimentado el Valle de México en los últimos cien años haya llevado a un calentamiento más intenso de la superficie, por lo que las condiciones atmosféricas se vuelven más inestables y se incrementa el potencial de nubes profundas y tormentas más intensas. (Magaña *et al* 2007)

Adicionalmente, la Ciudad de México está asentada, en su porción oriente, sobre un antiguo lecho lacustre constituido por sedimentos de arcillas casi impermeables que limitan la infiltración al subsuelo. Al poniente cruzan la ciudad cursos de arroyos y ríos pequeños cuyo caudal es regulado por presas localizadas en al pie de las montañas y que se encuentran entubados en su cruce posterior por la ciudad.

### 2.1.2. *La Vulnerabilidad*

Las características fisiográficas del Valle de México, en combinación con factores sociales, como la existencia de asentamientos humanos irregulares con vivienda precaria en zonas de riesgo, convierte a los fenómenos de tipo hidrológico en una amenaza constante para la población que habita estas zonas. El Distrito Federal tiene una población de más de 24 mil personas vulnerables a eventos hidrometeorológicos extremos, distribuidas en 168 sitios de riesgo. Debido a esto y sólo muy recientemente, las autoridades realizan esfuerzos por combatir constantemente el crecimiento desmedido de asentamientos humanos en zonas de riesgo, como puede ser la ocupación ilegal de barrancas, no solamente por el valor ambiental que tienen, sino porque al ocurrir una avenida pluvial intensa se producen deslizamientos de materiales como piedras, troncos de madera y basura, poniendo en riesgo a las personas y a sus viviendas.

En una urbe como la Ciudad de México, los episodios de lluvia intensa tipo estratiforme que aunque tienen poca intensidad, si tienen una larga duración en el tiempo, por lo que pueden ser causantes de la saturación del drenaje que se traduce en encharcamientos e inundaciones que tienen graves efectos en la vida cotidiana de sus habitantes por los daños materiales y hasta las pérdidas humanas que pueden ocasionar. Así mismo, las descargas eléctricas nube/suelo dan lugar a cortes de corriente que apagan semáforos con el consecuente caos vial característico de esta época del año en la ciudad.

Los años recientes han puesto de manifiesto la vulnerabilidad de la Ciudad de México ante eventos de precipitación extrema. Las constantes inundaciones al poniente y al sur de la ciudad indican que el riesgo ante aguaceros intensos puede verse incrementado por el aumento de la frecuencia de estos eventos. Por poner un ejemplo; el 2 de agosto de 2006 se registró una precipitación de 50.4 mm en tan sólo 36 minutos, causando inundaciones severas en varias colonias del centro y el poniente del Distrito Federal.

Mientras los sucesos de precipitación intensa aumenten, también lo harán los daños por inundaciones, deslices, aludes y derrumbamientos y se generará también un aumento en la erosión del suelo y un aumento en las escorrentías por inundaciones; que si bien pudieran llevar a un aumento de recarga de algunos acuíferos, también pueden resultar en una mayor presión sobre los sistemas de seguros privados que atienden estos siniestros y sobre

---

los gobiernos locales para auxiliar en caso de desastre

### 2.1.3. *Un enfoque científico para la prevención de desastres por eventos hidrológicos*

En las últimas décadas se ha visto un avance significativo en materia de predicción del tiempo y el clima, y se considera que los pronósticos del tiempo a 24, 48 e incluso 72 horas tienen ya una precisión suficiente para constituir la información más importante al momento de la toma de decisiones, debido a que en gran medida el rápido procesamiento de la información meteorológica alcanzado permite ganar tiempo ante la amenaza de un evento considerado extremo. De hecho, los centros de pronóstico del tiempo de diversos países cuentan ya con modernos sistemas de monitoreo de datos meteorológicos que permiten mantener una vigilancia constante de las condiciones del tiempo. Aunque los pronósticos de precipitación utilizando modelos numéricos a 24 o 48 horas, no tienen aún un alto grado de acierto, principalmente tratándose de precipitaciones intensas, (alrededor del 30 %, según Pérez (2004), y sea también cierto que existen limitaciones en el grado de detalle de un pronóstico numérico, sería una equivocación pensar que se debe esperar a que un pronóstico sea perfecto para comenzar a utilizarlo en la toma de decisiones.

Se puede considerar que se ha llegado a un estado del conocimiento en que los productos de información meteorológica, con sus incertidumbres asociadas, ya pueden ser usados para definir acciones de prevención, principalmente en materia de protección civil. (Magaña *et al* 2007). Falta, sin embargo, establecer los mecanismos mediante los cuales esta información se aplicará y se generará para una región tan específica como lo es el Valle de México. Es decir, mejores predicciones no solucionarán por sí solas el problema de los desastres hidrológicos en la Ciudad de México ya que es necesaria la existencia de una buena *definición* de los objetivos de los productos meteorológicos en términos de las necesidades a satisfacer, aparte del urgente mejoramiento de los incipientes mecanismos de comunicación entre científicos, tomadores de decisiones y la población civil, acerca de la información meteorológica y de pronóstico, lo que resultaría en mejores programas de preparación contra desastres hidrometeorológicos en general.

Por ejemplo: aunque actualmente la Dirección de Protección Civil del Dis-

---

trito Federal realiza distintos esfuerzos para disponer de información meteorológica y realizar así acciones de prevención (evacuaciones o desplazamiento de cuerpos de socorro) ante el anuncio de condiciones de tiempo severo, la información que produce no posee los detalles que permitan definir los lugares puntuales de la ciudad que deban ser alertadas. Si bien se generan pronósticos de lluvia, no se cuenta con un esquema que permita tomar acciones de prevención ante tormentas severas debido a que no se tiene capacidad de determinar cuándo y en dónde pueden ocurrir estos fenómenos, siendo la población civil la más afectada.

De tal modo y dados los avances en el entendimiento de la física detrás de fenómenos hidrológicos, junto con la capacidad de pronóstico de hoy en día, la ocurrencia de un evento extremo no debiera llevar necesariamente a la ocurrencia de un desastre. Sin embargo, la prevención no sólo depende de nuestro entendimiento de la amenaza, sino de la decisión de reducir la vulnerabilidad por parte de las autoridades y la sociedad.

#### *Manejo del Riesgo*

Es importante destacar que la posibilidades de utilizar la técnica MOS para corregir los errores de predicción de un modelo de mesoscala abre implicaciones inmediatas en el ámbito de la Protección Civil en un área como la de la Ciudad de México. Esta se daría en términos de permitir una buena gestión del riesgo, y aún de respuesta ante la inminencia de un posible desastre. Es decir, se podría utilizar un esquema de QPF para realizar escenarios de riesgo a cierto plazo tiempo a partir de los pronósticos de un modelo regional, y aprovechar la física de los mismos para prever las zonas más vulnerables a se afectadas por un evento de lluvia en la ciudad, sea intenso o no. De esta manera se podría trabajar en la reducción del riesgo en zonas muy puntuales (y no de decenas de kilómetros) donde la adaptación y posible mitigación sería más fácil (quizás) de conseguir. Estas zonas, una vez identificadas, tendrían que conseguir la atención de las autoridades de manera inmediata, elaborando planes de contingencia con las comunidades y estableciendo prioridades con los cuerpos de socorro, así como un análisis completo de sus vulnerabilidades a eventos hidrológicos.

## 3. DATOS Y METODOLOGÍA

### 3.1. *Datos Meteorológicos*

La necesidad de contar con buenos datos climáticos se hace fundamental, dado que éstos son el insumo fundamental para la gestión de los riesgos relacionados con las variaciones y cambios en el clima (AGROASEMEX, 2006; Boissonade, 2002). En los últimos años, y particularmente desde que se han identificado los impactos de diversos fenómenos climáticos relacionados directamente sobre la sociedad en su conjunto (Magaña *et al* 1999), se ha observado un interés de las distintas instancias gubernamentales y no gubernamentales para generar instrumentos teóricos y prácticos encaminados hacia lograr un mejor entendimiento de la variabilidad climática. Las bases de datos meteorológicos se caracterizan por su confiabilidad. En el caso de México, aunque son diversas las bases de datos con las que se cuenta, sobre ellas no se ha aplicado un proceso de control de calidad, homogeneidad y análisis espacio-temporal sistemático para generar un conjunto de datos estándar que sirva de base para la generación de productos operativos y de investigación de alta eficiencia. Las bases de datos climáticas de calidad permitirían no sólo el estudio de la distribución de la variabilidad climática regional, sino también avanzar en su modelación y en la valoración de los impactos de distintos fenómenos climáticos en la sociedad.

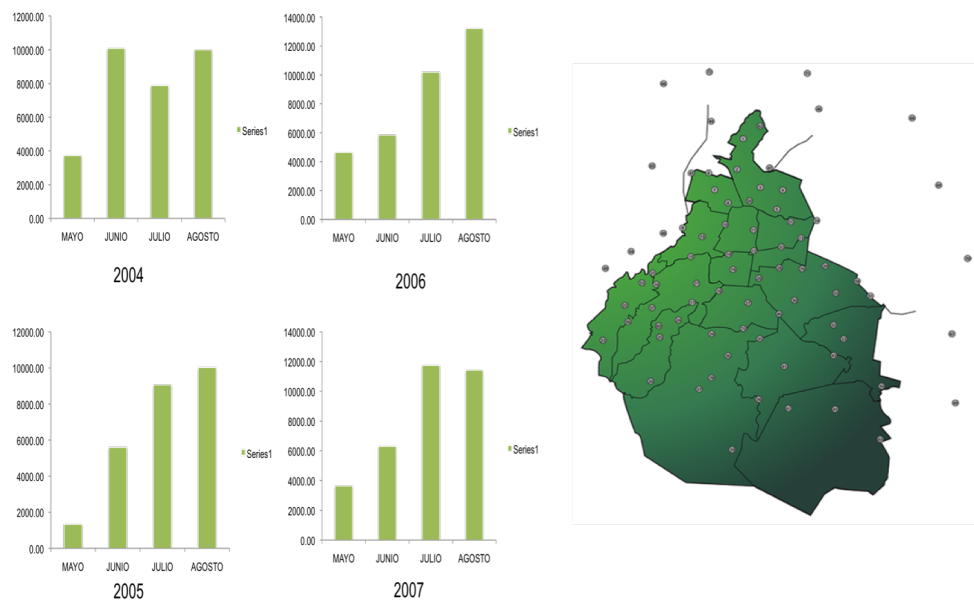
#### 3.1.1. *Datos de precipitación*

Los datos de precipitación observada utilizados para este estudio se obtuvieron de la red pluviométrica del Valle de México, a cargo del Sistema de Aguas de la ciudad de México SACM (antes Dirección General de Construcción y Operación Hidráulica (DGCOH)). Esta red pluviométrica es la más completa en cuanto a cobertura espacial y temporal sobre la Ciudad de México. Actualmente, la red está compuesta por casi 80 estaciones plu-

viométricas (Fig. 3.1) que registran datos al menos cada 3 minutos desde 1993 (SACM 2010). Durante episodios de lluvia, los datos son transmitidos en tiempo real vía radio. Esto permite, entre otras cosas, analizar con gran detalle la evolución y la propagación de las tormentas. La continuidad de los registros en esta base de datos permite caracterizar las lluvias de acuerdo a su distribución espacial, intensidad, duración, etc.

La base de datos del Sistema de Aguas permite diferenciar la variable de interés, pudiéndose obtener ya sea intensidad o altura de lluvia (lluvia acumulada). También es posible elegir la hora de medición inicial y final.

Para este trabajo se obtuvieron datos de precipitación acumulada horaria para el periodo de mayo a agosto del año 2003 al 2008, para representar lluvias de verano, que es, entre otras cosas, cuando el flujo de vientos, predominantemente del este, modula la intensidad y la distribución de la lluvia sobre el Distrito Federal. Los datos de precipitación diaria están medidos a partir de las 8am y hasta las 07:45 del día siguiente.



**Fig. 3.1:** Precipitación acumulada registrada en las 78 estaciones del SACM en los años 2004, 2005, 2006 (izquierda) y distribución espacial de las estaciones de la red de la SACM (derecha)



### 3.1.2. Control de calidad de los datos del SACM

Muchas técnicas estadísticas de análisis de datos requieren de series de tiempo *completas* (Daley, 1991; Nicholls, 1987; Richman y Lamb, 1985). Sin embargo, cuando el objetivo es generar resultados a partir de un análisis de datos obtenidos por métodos instrumentales, existirá siempre un error entre lo que efectivamente sucedió, y lo que se registró con el dispositivo. Aunado a la incertidumbre asociada con la medición del dispositivo, si el método de captura es manual, los errores inherentes de tipo tipográfico y de falta de rigor puede llevar a resultados erróneos. Cuando el método es automatizado, si importan las condiciones de las estaciones climatológicas ya que esto varía de acuerdo al estado de funcionamiento de las estaciones, que va desde las que están en muy buen estado hasta las que requieren mantenimiento urgente o no funcionan. Tristemente, estos datos sí se consideran al momento de digitalizar los datos y comenzar con el análisis. De este modo, el mantenimiento de una red de medición es una tarea prioritaria y elemental que debe ser permanente. Se han desarrollado diversas técnicas para ‘estimar’ los datos faltantes. Sin embargo, no hay una técnica que pueda considerarse la mejor y un dato estimado es, en realidad, algo que nunca sucedió (Schneider, 2001).

Informar sobre los errores y las dudas de los datos no implica que las medidas sean erradas o fallidas, más bien es todo lo contrario. La magnitud del error describe cuán seguros estamos de la exactitud de los datos, así que los mayores errores informados indican menor seguridad. A decir de la Organización Meteorológica Mundial en sus ”Directrices sobre la gestión de datos climáticos” (WMO Programa Mundial de Datos y Vigilancia del Clima 2007), las comprobaciones efectuadas para determinar la calidad de una observación pueden revestir diversos grados de complejidad. Por ejemplo pueden consistir en:

- Comprobaciones de la franja de valores climáticos (por ejemplo, ¿es coherente el dato con la climatología?)
- Coherencia espacial y cronológica (por ejemplo, no deberán excederse las diferencias plausibles entre los valores de precipitación de una estación y los de estaciones vecinas).

La red de estaciones de la SACM presenta deficiencias, debido al no funcionamiento de algunas estaciones, además de datos incongruentes o faltan-

tes. Esto no significa que la base de datos de la red de estaciones de la SACM sea mala; a decir de los encargados del departamento de información del SACM, la base de datos de observaciones en pluviómetro, no es la adecuada para hacer meteorología. El Sistema de Aguas utiliza una esta base como información para una respuesta inmediata ante precipitaciones importantes en orden de desplegar infraestructura de contención. Para evitar utilizar estimaciones en la creación de la base de datos principal que se utilizó para análisis (la red de la SACM), se comparó el registro de lluvia horaria en cada estación para cada día, con los registros de la base de datos del ERIC III (Extractor Rápido de Información Climática), que contiene información del banco de datos histórico nacional del Servicio Meteorológico Nacional (SMN) y de la Comisión Nacional del Agua (CONAGUA). Algunas estaciones tienen información desde 1900, aunque la mayoría de los datos (88.4%) corresponden al periodo de 1960 a 2007, para algunas pocas estaciones se cuenta con información hasta 2008. (Servicio Meteorológico Nacional. 2000). otra fuente de datos útil para comparar fue la base de datos de estaciones del programa PEMBU (Programa de Estaciones Meteorológicas del Bachillerato Universitario: <http://pembu.unam.mx/version/index.html>) que surge en 1992 y que hoy en día cuenta con 20 estaciones dentro de la ciudad de México. En el caso de que una estación de la red de la SACM presentara un valor faltante, se utilizó en su lugar el valor reportado por ERIC III o por PEMBU, siempre y cuando la posición de las estaciones entre las tres bases de datos, no presentara una diferencia en distancia mayor a 8km, siendo esta el tamaño promedio de una nube de lluvia.

Una vez construida la base de datos principal, se validó el registro de valores atípicos (si existió o no) estableciendo un rango estadístico para la precipitación acumulada y fuera del cual ningún valor fue admitido como válido. Cuando se examina una serie temporal de datos, es posible hallar algunos puntos que difieren dramáticamente del resto de los datos. En algunos casos es posible considerar que estos valores no son consistentes con el resto por lo que se procede a eliminarlos. Sin embargo, hay que entender el origen del problema que se desea corregir antes de eliminar estos valores extremos y por lo tanto alterar toda la base de datos. La eliminación de un valor atípico tiene un mayor efecto sobre la desviación estándar que sobre la media de la datos, así que eliminar un valor atípico da lugar a una desviación estándar menor que podría resultar en considerar algunos puntos como valores atípicos, cuando no lo son. En nuestro caso, un valor atípico con respecto

de la media de precipitación acumulada horaria no significa necesariamente que éste dato sea inconsistente respecto a los demás, sino que puede representar un valor extremo de precipitación que, por definición, está por fuera del comportamiento medio (Wilks 1995). Entonces el objetivo hasta aquí no es eliminar estos valores sino solamente identificarlos. El criterio estadístico utilizado es el de considerar un valor como atípico cuando este sea mayor o igual a 4 desviaciones estándar ( $\sigma$ ) (Programa Mundial de Datos y Vigilancia del Clima, 2007). A partir de la media de precipitación acumulada horaria registrada en la base de datos principal, se calculó el valor de la desviación estándar la cual resultó de 2.03 por lo que el umbral definido para caracterizar a un valor atípico es de 8.13mm/hr.

Varianza	Desviación Estándar	Umbral de Valor Atípico
4.1	2.03	8.1

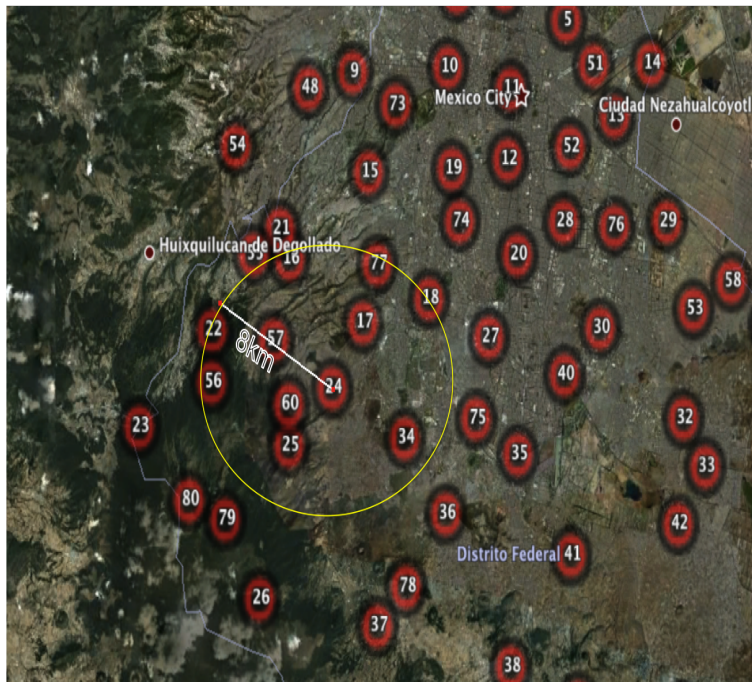
**Fig. 3.2:** Valores obtenidos para categorizar un valor atípico en la base de datos de la SACM

Colocar los datos en un formato visual puede facilitar análisis adicionales. De tal manera que el despliegue de las series de tiempo de la SACM fue diseñado para mostrar valores de precipitación acumulada horaria en el tiempo y en el espacio para cada registro simultáneamente. A través de una función  $f(x, t) = y$  ( $x$  =posición de la estación  $E_n$ ,  $t$  =tiempo y  $y$  =precipitación acumulada horaria en estación), podemos establecer cuando y en donde se presentó precipitación en todo el Valle de México, de acuerdo al reporte en estación.

En la herramienta de cómputo MatLab se generaron 738 series de tiempo correspondientes a la precipitación acumulada horaria en cada estación, para cada día, de cada mes en el intervalo de tiempo considerado para el estudio (2003-2008).

En orden de verificar si estos datos atípicos ocurrieron o no, se consideró un esquema de comparación espacial que consistió en definir a un valor

de precipitación atípico como verdadero (que si ocurrió) en estación cuando al menos 3 estaciones dentro de un radio de 8km ( que equivale a la distancia media entre estaciones) registraran un valor de precipitación acumulada horaria cuya diferencia con respecto al valor a validar no rebasara el 30 por ciento. Utilizando las series de tiempo y el software *Google Earth* se analizaron cada una de las series de tiempo diarias para cada estación al tiempo que se validaba la existencia o no de un valor atípico comparando los valores de estaciones cercanas respetando el criterio anterior. Así mismo, se reutilizó la base de datos del ERIC y PEMBU para verificar la existencia de estos valores atípicos, comparando el valor reportado por la SACM con el reportado por el ERIC y el PEMBU en ese momento.

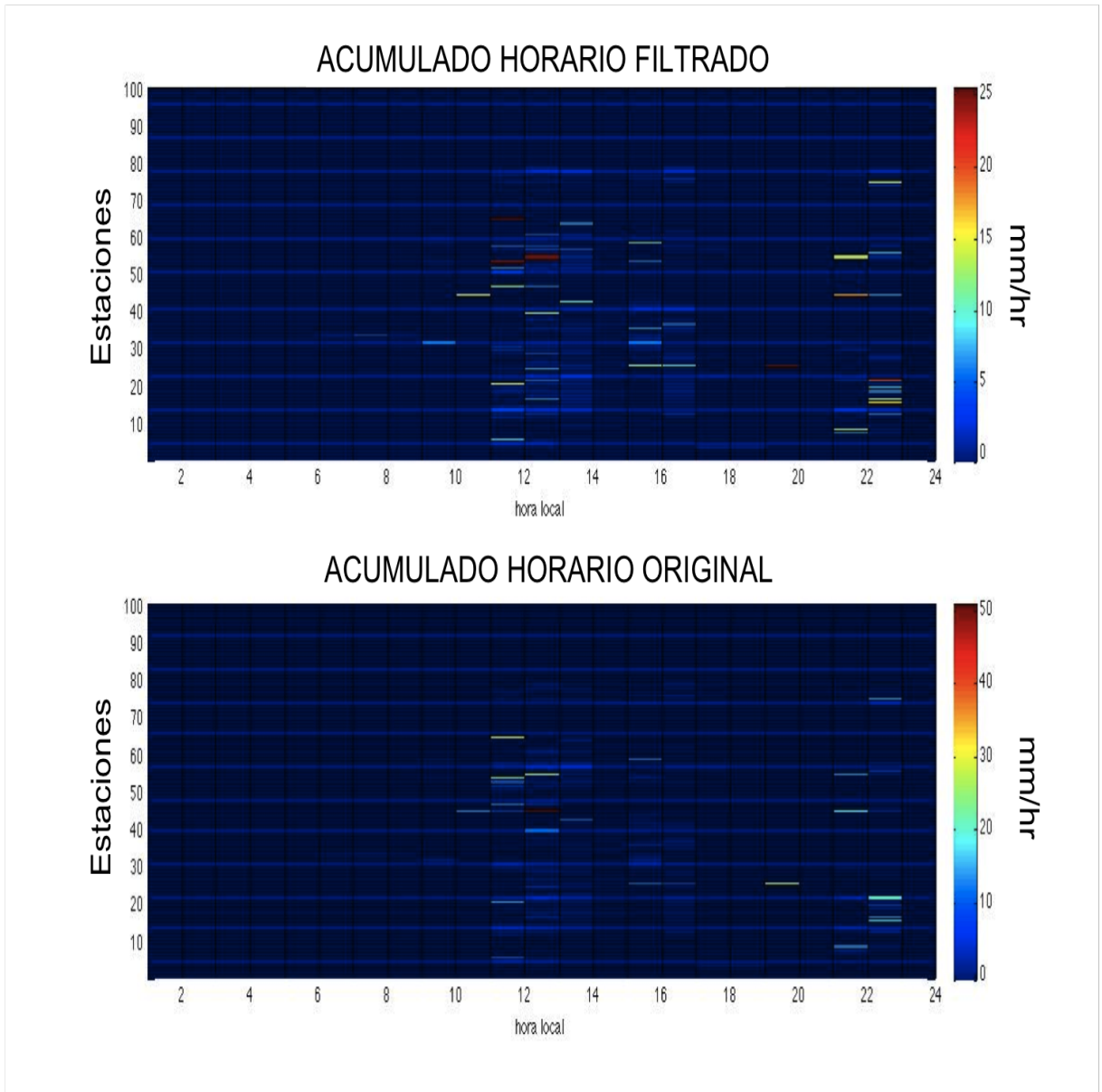


**Fig. 3.3:** Se muestran algunas estaciones de la red de la SACM (círculos en rojo) y el radio de 8km definido para validar eventos de lluvia atípicos

Así se eliminaron valores que bajo el criterio definido no existieron y se verificaron valores que bajo el primer análisis puramente estadístico hubieran

sido despreciados. Este análisis *geo gráfico* permite también filtrar valores erróneos al poder tener idea de la posición exacta de cada estación y por lo tanto, de su correspondiente régimen de precipitación acumulada horaria tomando en cuenta la presencia de irregularidades importantes en el terreno, como montañas o cuerpos de agua, que regulan entre otras cosas, el flujo de humedad y de vientos necesarios para la formación de precipitación. De este modo registros de 20 mm de precipitación acumulada en una hora, son plausibles en la zona suroeste de la Ciudad, pero son improbables en la parte norte o centro. (Magaña *et al* 2003)

El cambio del comportamiento de las series de tiempo para la precipitación acumulada después del control de calidad es claro. Como ejemplo se muestra el caso del 20 de agosto del 2005. (Figura 3.4)



**Fig. 3.4:** Comparación entre SACM filtrado (arriba) y sin filtrar (abajo) para el 20 de agosto del 2005. En el eje ordenado mm/hr, y el eje de las abcisas horas.

### 3.1.3. Datos de precipitación estimados por satélite

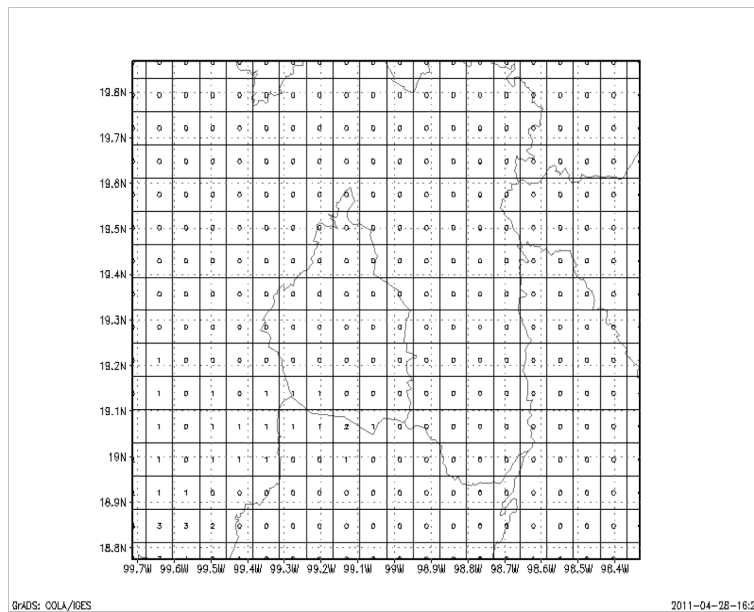
Los productos meteorológicos generados por percepción remota de alta resolución, como la precipitación, están cada vez más disponibles debido al desarrollo de plataformas con mayor resolución espacial. Sin embargo, los ejemplos de los usos del funcionamiento de estos productos son poco frecuentes, siendo la razón principal la falta de confianza en la exactitud cuantitativa de la dichos productos. La técnica de estimación llamada *morphing* generada por el Centro de Predicción Climática de la Administración Nacional Oceánica y Atmosférica de la Administración (NOAA), es la estimación de la precipitación mundial con la mejor resolución: 30 minutos y  $0,088^\circ$  de latitud, y  $0,088^\circ$  de longitud (Joyce et al. 2004).

La técnica "morphing", utiliza las precipitaciones estimadas de observaciones de microondas. Luego propaga estas características en el espacio utilizando vectores de movimiento derivados de observaciones de 30 minutos en el infrarrojo por satélites geoestacionarios (Joyce et al. 2004). Básicamente, los datos de microondas son utilizados para obtener información detallada sobre la estructura microfísica de un conjunto de nubes. Eventualmente las imágenes infrarrojas son utilizadas para un control continuo y de alta resolución espacial de la evolución de las estructura nubosas. De esta forma, CMORPH combina la exactitud de las estimaciones de microondas pasivas y la mayor resolución temporal y espacial de datos en infrarrojo, ( además de que pueden ser útiles cuando no existan datos de mircoondas en alguna localidad.)

CMORPH es, por tanto, una técnica de mezcla y no un procedimiento de algoritmos de estimación de precipitación, por lo que no hereda los problemas inherentes de un algoritmo de estimación. Aunque estas estimaciones están disponibles en una malla con una distancia de separación de 8 km (en el ecuador), la resolución de las estimaciones derivadas primeramente del satélite son más gruesas que eso, del orden de 12 x 15 kilómetros más o menos. La resolución mas "fina" se obtiene mediante métodos de interpolación.

Así pues, CMORPH ofrece datos de estimación de precipitación con resolución temporal de 30 min y resolución espacial de 8km ( $0.08^\circ$ ) en el ecuador a partir de diciembre de 2002. Cada registro del CMORPH se compone por arreglos de malla de 4984 x 1649 con incrementos de  $0.07^\circ$  de latitud y  $0.07^\circ$  de longitud. Los arreglos están orientados desde el Norte hacia el Sur, empezando con una latitud de  $59.96^\circ$  4N y desde el Oeste hacia el Este a partir de una longitud de  $0.03^\circ$ E.

Para esta investigación, se obtuvieron datos horarios de estimados de precipitación acumulada horaria para el periodo de 2003 a 2008 y se generó un dominio de 168km x 128km centrado entre los 98.4° y 99.7° de longitud oeste y entre los 18.8° y 19.8° de latitud norte, que equivalen a 168km x 128km (ó 20 x 16 puntos de malla) y representa la zona que abarca la Cuenca del Valle de México, ver figura 3.5



**Fig. 3.5:** Dominio definido para la base de datos del CMORPH

#### 3.1.4. Metodología de interpolación de datos en una malla regular

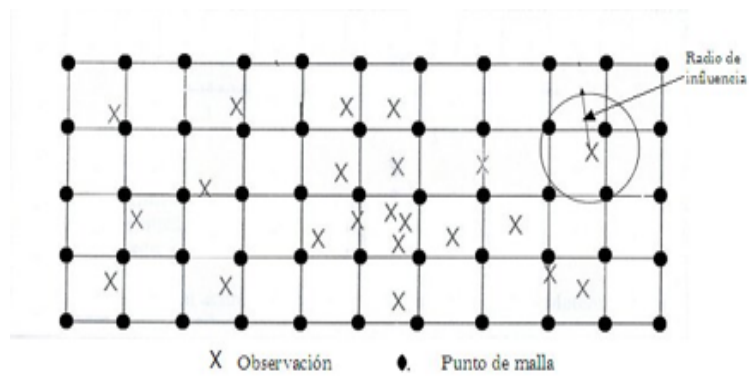
El construir climatologías, calcular anomalías o comparar pronósticos numéricos con observaciones es más sencillo si se realiza un análisis objetivo, el cual consiste en construir campos en mallas regulares a partir de observaciones distribuidas irregularmente en el espacio por medio de un proceso de interpolación.

El correcto desarrollo de bases de datos en mallas regulares es de utilidad para el entendimiento de la variabilidad espacio-temporal del clima. De tal



modo que desarrollamos una base de datos de lluvia acumulada horaria en mallas regulares siguiendo el método propuesto por Cressman (1959), el cual mejoró la técnica para análisis objetivo propuesto por Panofsky (1949) y Gilchrist (1954) al considerar un método de aproximaciones sucesivas en vez de estimaciones basadas en polinomios.

El esquema para el análisis objetivo de Cressman (Fig. 3.7), crea patrones en una malla regular a partir de una condición inicial o campo preliminar (usualmente una predicción, un promedio climatológico o una simple persistencia climática, entre otras posibilidades) la cual es modificada por datos distribuidos aleatoriamente. Este esquema está basado en un radio de influencia anisotrópico para modificar los valores de puntos de malla alrededor de cada observación.



**Fig. 3.6:** Esquema del Análisis de Cressmann

En sentido general, el método de Cressman se basa en la realización de diferentes niveles de interpolaciones (iteraciones), cuya cantidad está determinada por el número de radios de influencia que se especifique. En cada iteración, el valor de cada punto de malla se determina utilizando los valores de las estaciones que se encuentren dentro del radio de influencia. Antes de pasar a la siguiente iteración (siguiente radio de influencia) se realiza una estimación de errores, los cuales son utilizados para producir un factor de corrección, que es aplicado a cada punto de malla mediante una fórmula de distancia ponderada.

De tal modo que el análisis de Cressman consiste en tener una primera aproximación de la variable a analizar y aprovechando los datos reportados como

observados (de estaciones), se corrige la primera aproximación al comparar la diferencia entre el campo preliminar y las observaciones alrededor. El radio de influencia es entonces reducido en iteraciones sucesivas para incrementar la precisión. Así pues, para cada estación, un error se define como la diferencia entre el valor medido en la estación y el valor obtenido al interpolar desde la posición de esa estación hacia la malla. El factor de corrección ponderado ( $W_j$ ), que depende exclusivamente del radio de influencia ( $R$ ) y la distancia entre el punto de malla y la estación ( $r_j$ ), es aplicado entonces a cada uno de los errores presentes dentro del radio de influencia de un punto en la malla, definiendo así un factor de corrección para ese punto.

$$W_j = \frac{R^2 - r_j^2}{R^2 + r_j^2}$$

Si  $r < R$  y  $W_j = 0$  si  $r > R$

El valor corregido en cada punto de malla ( $G_{i+1}$ ) es calculado como el valor obtenido en la iteración anterior ( $G_i$ ) más la suma de los productos de los factores de corrección ponderados ( $W_j$ ) y la diferencia entre el valor real en estación ( $P_j$ ) y el valor interpolado para esa estación ( $P_j^e$ ) dividido entre la suma de los factores de corrección ( $W_j$ ).

$$G_{i+1} = G_i + \frac{\sum_{j=1}^N W_j (P_j - P_j^e)}{\sum_{j=1}^N W_j}$$

donde  $i$  representa la  $i$ -ésima iteración y  $j$  a la estación  $j$ .

Se generó una base de datos de lluvia acumulada horaria en una malla de 21 x16 puntos de malla de resolución espacial (8km) centrada entre los 18.8° y 19.8° de latitud norte y entre los 98.4° y 99.7° de longitud oeste. Como campo preliminar se utilizaron los datos del CMORPH de estimación de precipitación acumulada horaria y como observaciones los datos de precipitación acumulada, también horaria, de la red de estaciones pluviométricas de la SACM.

Debido a que una de las desventajas del método de Cressman es que no existe un método para definir el radio de influencia óptimo (y tiene que hacerse por ensayo y error) se realizaron comparaciones del comportamiento del campo interpolado con el campo observado. Al final, se consideró que

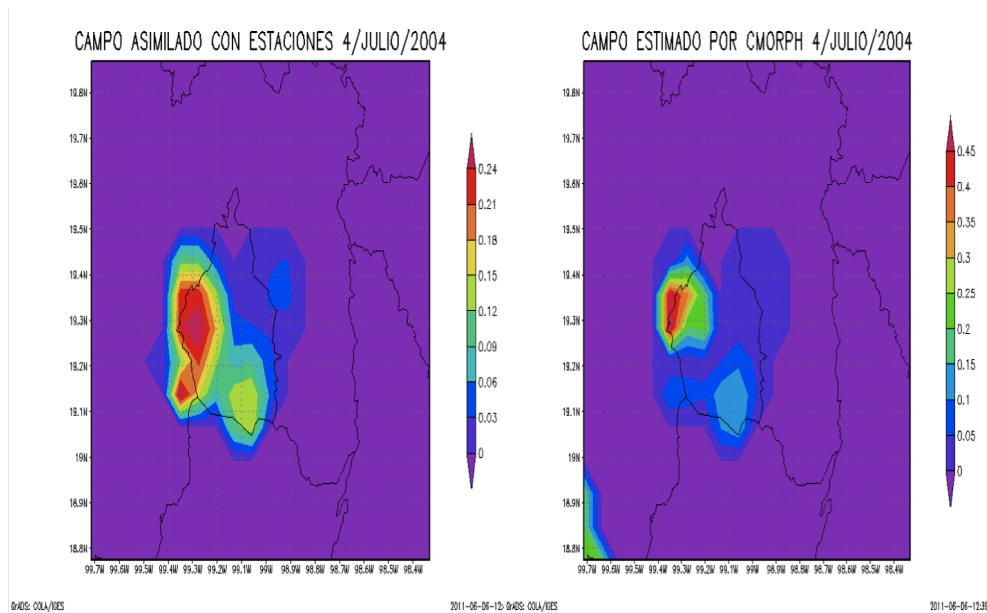
3 iteraciones eran suficientes para representar el campo observado correctamente ya que si el radio de influencia aumentaba, el patrón espacial de los datos interpolados se ampliaba, pero no se modificaba, con respecto al campo observado del SACM. Se consideró un valor máximo para el radio de influencia de 12km (aproximadamente la distancia media entre estaciones) y radios sucesivos de 8 y 4km respectivamente.

El método de Cressman tiene la ventaja de ser más simple y más eficiente en cómputo, además de que ha demostrado ser más preciso que otros métodos de análisis objetivo como la interpolación simple o la estimación por ajustes polinomiales. No obstante, también presenta desventajas, las principales (además de la ya mencionada) se deben a que éste puede volverse inestable si la densidad de la malla a estimar es mucho mayor que la densidad de las observaciones disponibles (Pérez, 2004), además de que el método de Cressman es sensible a los errores presentes en las observaciones; puede producir valores extremos ficticios, y, si el número de estaciones dentro del radio de influencia disminuye, la calidad de las interpolaciones se reduce también, de tal modo que los campos interpolados con menos estaciones para modificar el campo preliminar, no serán capaces de captar todas las características espaciales del fenómeno.

Para la evaluación de la base de datos en mallas así generados, se comparó su distribución espacial de lluvia con respecto a la base de datos del CMORPH (SACM). Al mismo tiempo, se comparó que la cantidad de lluvia estimada en la malla versus los datos de estación fuese la correcta, o en su caso, que reproduzca cada uno de los eventos de lluvia individuales.

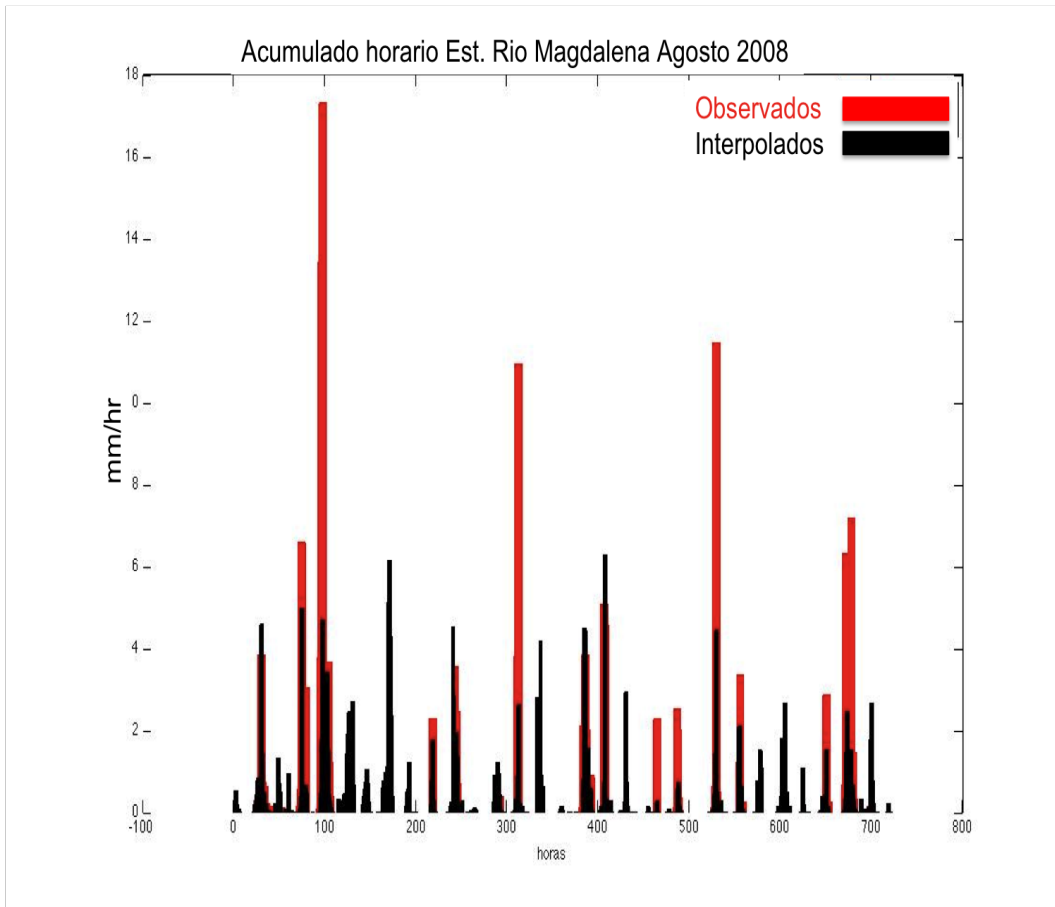
Los campos de lluvia con una resolución como la del CMORPH pueden representar el comportamiento de la lluvia en áreas muy locales, pero tiende a suavizar el patrón y a subestimar la presencia de lluvias en algunas zonas, cuando se le compara con los campos observados (Zeweldi et al, 2008). Sin embargo, cuando estos campos son modificados por los registros de estaciones, la información que se gana en términos de espacialidad es muy buena. Para ejemplificar lo anterior, se muestra la comparación del patrón espacial de lluvia observada horaria en la base de datos de estimados por satélite y los campos de lluvia corregidos por observaciones de estaciones, para el día 4 de julio de 2004 a las 03Z. Existen zonas de lluvia que si existieron, que no estima CMORPH (ver Fig. 3.8), sin embargo, el campo espacial asimiliado

si reconoce estas mismas zonas.

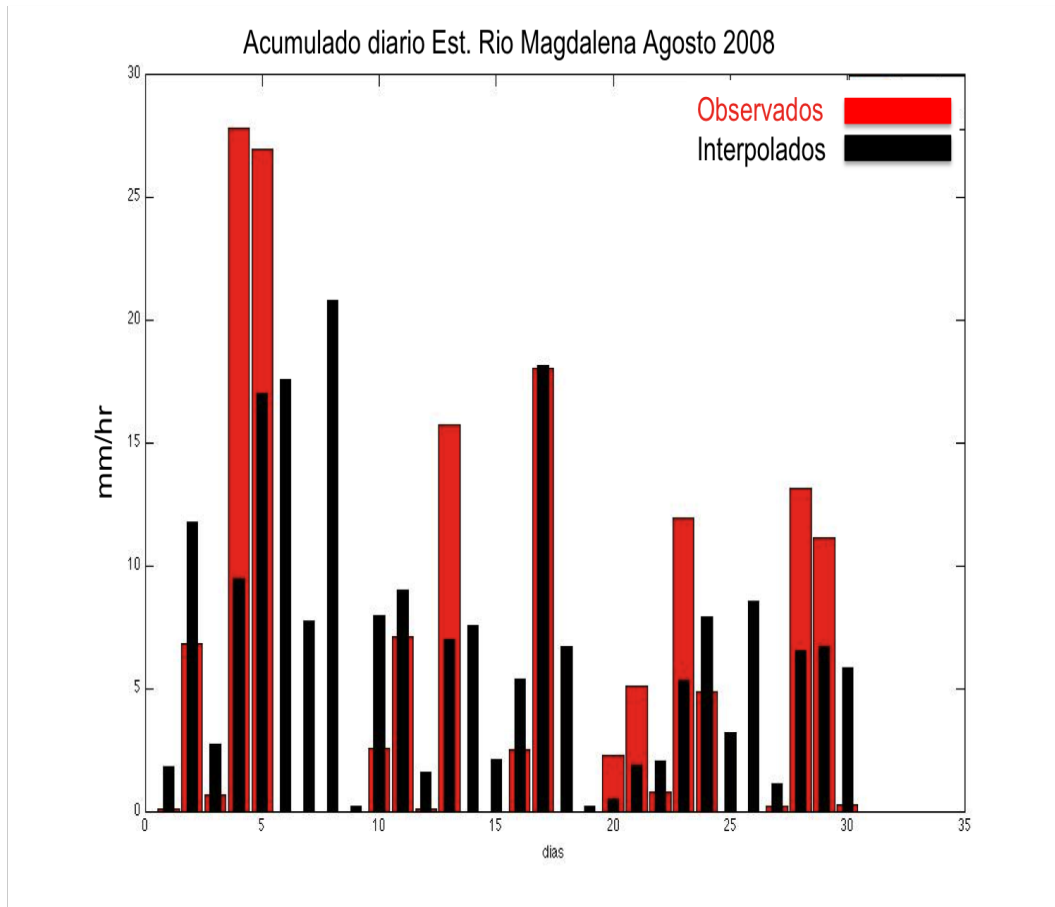


**Fig. 3.7:** Campos de lluvia asimilado con estaciones de la SACM (der) y el símil estimado a través de cmorph (izq)

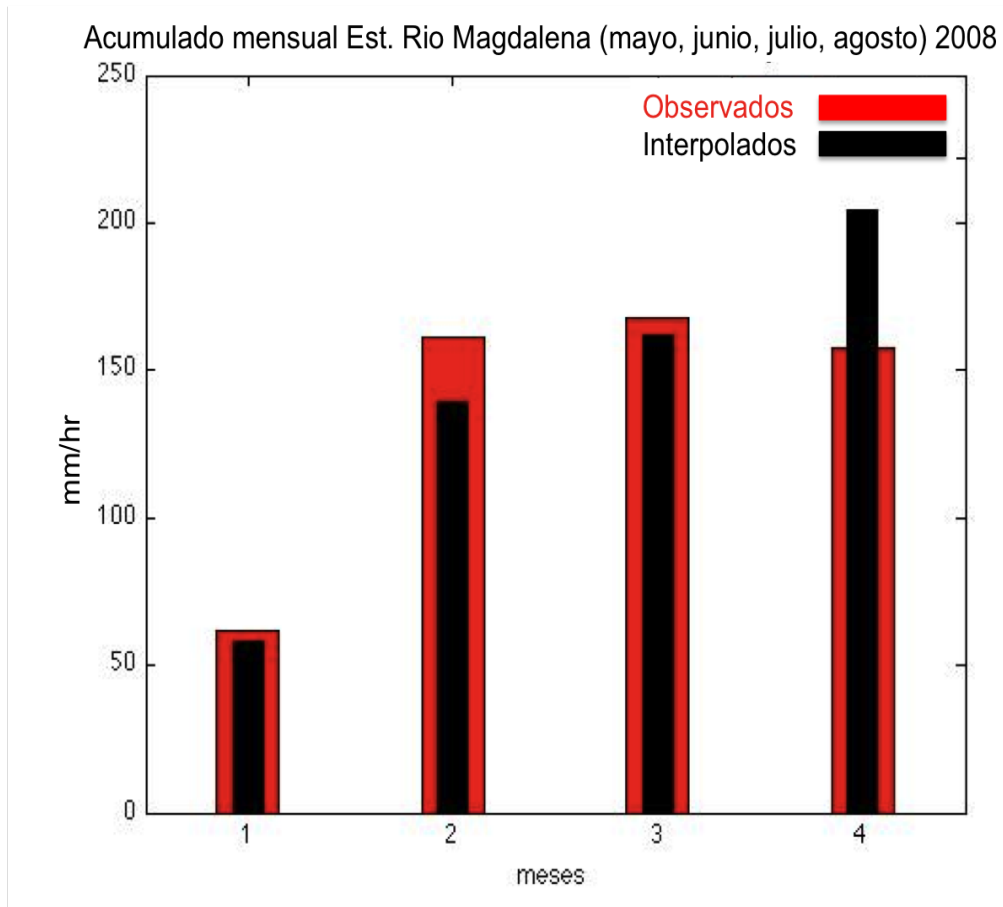
Para verificar que los campos asimilados reprodujeran correctamente la cantidad de lluvia acumulada reportada en estación, se comparó la lluvia acumulada estimada en la base de datos interpolados, con la serie de tiempo de todas las estaciones de la red del SACM en tres escalas temporales (horaria, diaria y mensual). Como ejemplo se muestran las gráficas que comparan las series para las estaciones Río Magdalena, ubicada en la delegación Magdalena Contreras, y El Coyol, en la delegación Gustavo Madero (Figuras: 3.8 y 3.9, 3.10 y 3.11).



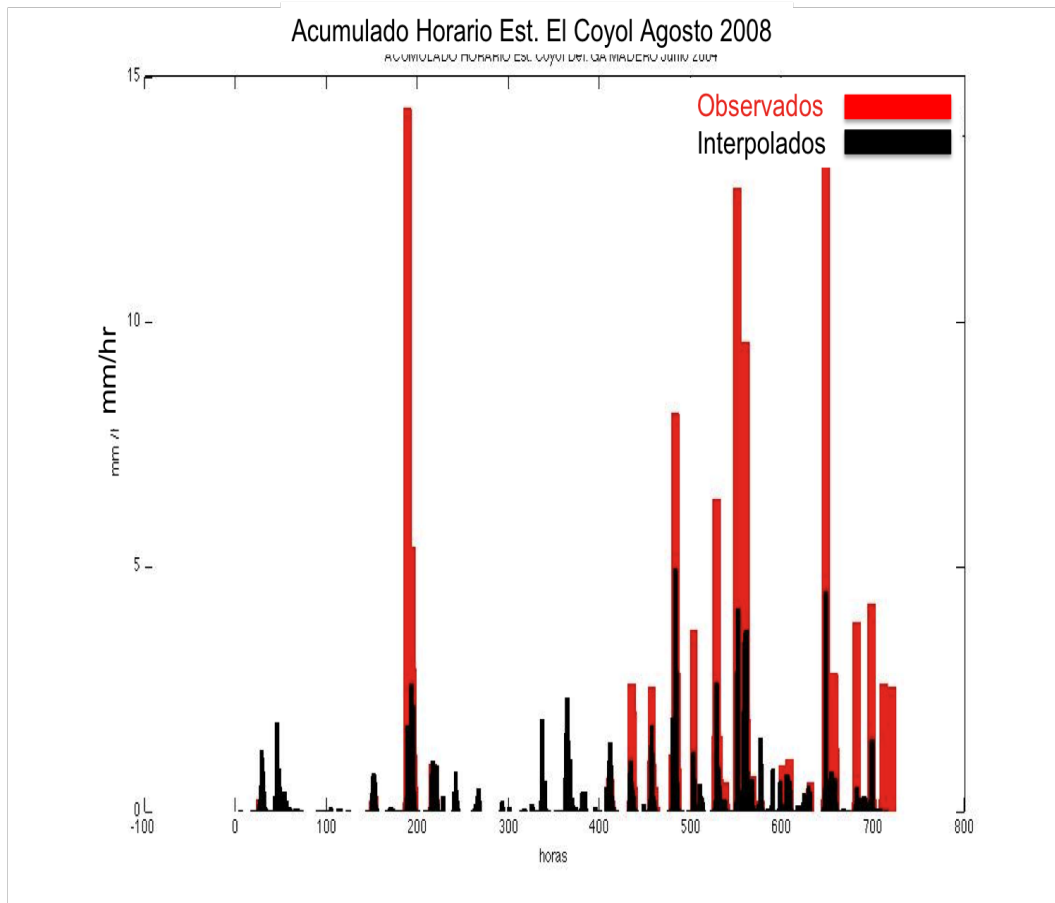
**Fig. 3.8:** Acumulados horarios, diarios y mensuales para la estación Rio Magdalena (M.C). Observados en rojo, interpolados en negro



**Fig. 3.9:** Acumulados horarios, diarios y mensuales para la estación Rio Magdalena (M.C). Observados en rojo, interpolados en negro

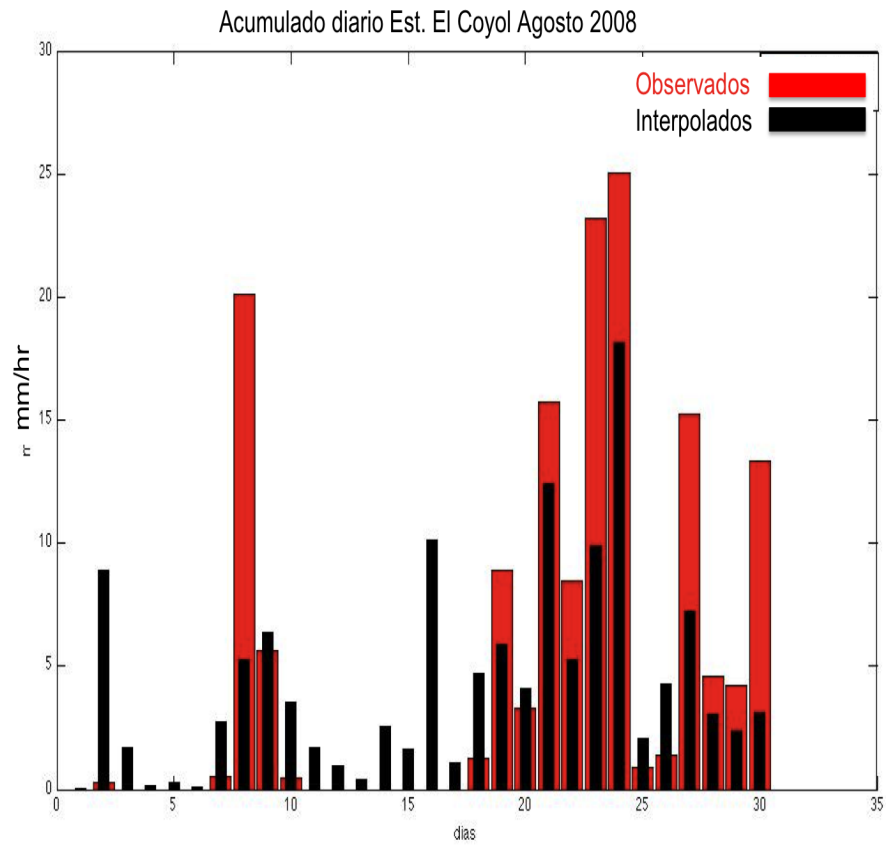


**Fig. 3.10:** Acumulados horarios, diarios y mensuales para la estación Rio Magdalena (M.C). Observados en rojo, interpolados en negro

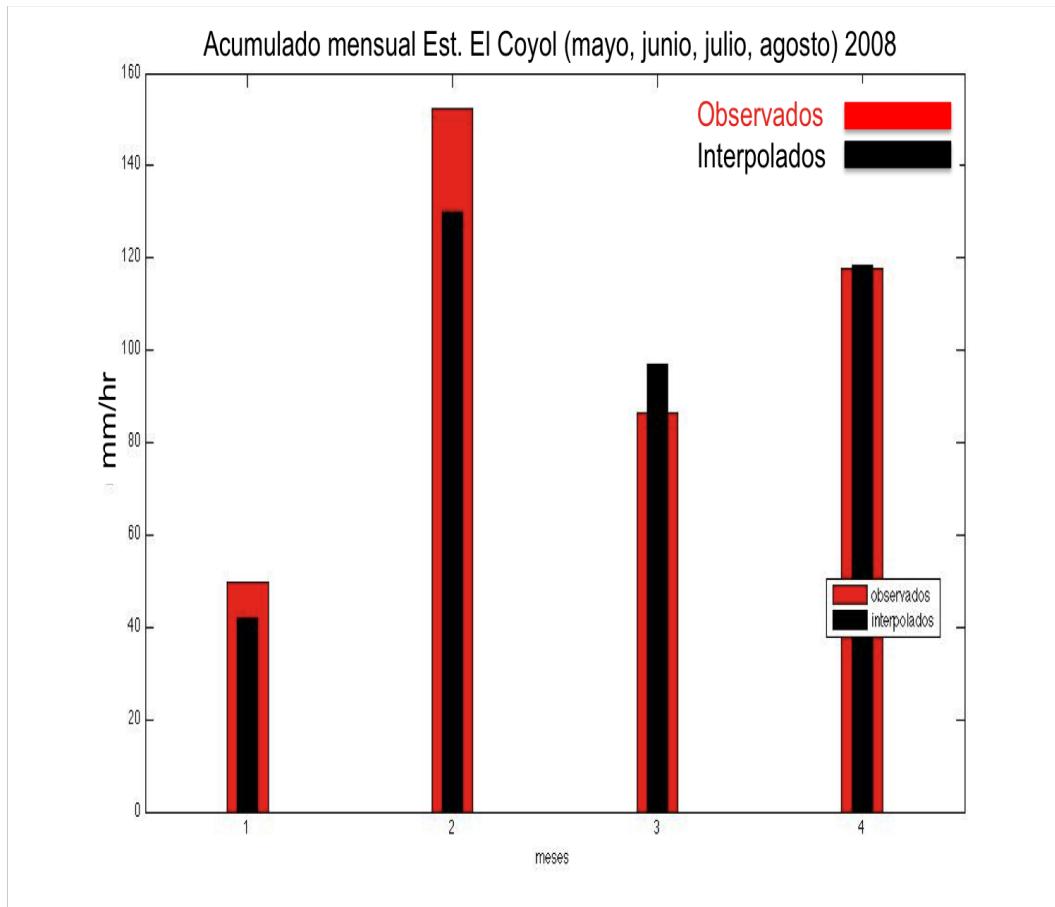


**Fig. 3.11:** Acumulados horarios, diarios y mensuales para la estación Río Magdalena (M.C). Observados en rojo, interpolados en negro





**Fig. 3.12:** Acumulados horarios, diarios y mensuales para la estación Río Magdalena (M.C). Observados en rojo, interpolados en negro



**Fig. 3.13:** Acumulados horarios, diarios y mensuales para la estación Río Magdalena (M.C). Observados en rojo, interpolados en negro

De acuerdo a esta comparación, la base de datos interpolada reproduce fielmente cada uno de los eventos de lluvia horarios, sin embargo, siempre que se interpole a punto de malla (utilizando un método de ponderación de distancias en este caso), se corre el riesgo de reducir la intensidad de lluvia en las series de tiempo y por ende a la subestimación de valores extremos. Eso es muy claro en las series de tiempo mostradas, en donde la intensidad de lluvia horaria se ve reducida casi por un factor de 7. Aún así, en las gráficas de lluvia diaria y mensual, se puede observar que el comportamiento de los valores interpolados siguen muy de cerca al de los reportados en estación. Ciertamente los resultados no son idénticos, pero el patrón de comportamien-

---

to de la lluvia diaria y mensual es capturado con bastante precisión en la estimación a punto de malla.

### 3.2. *El Modelo MM5 (Mesoscale Model Version 5)*

#### 3.2.1. *Breve Historia del Pronóstico Numérico*

Los pronósticos numéricos son cálculos objetivos del cambio de la atmósfera. Son llamados objetivos en el sentido en que dado un conjunto de datos de entrada siempre producen el mismo tipo particular de datos de salida. A diferencia de los pronósticos del pasado hechos con métodos subjetivos, el pronóstico objetivo es consistente y se puede estudiar, reevaluar y mejorar. Modelar numericamente es el proceso mediante el cual se obtiene un pronóstico objetivo del estado futuro de la atmósfera mediante la solución de un grupo de ecuaciones que describen la evolución de un grupo de variables (temperatura, velocidad y dirección del viento, humedad, presión) que a su vez definen el estado de la atmósfera. El proceso se inicia con el análisis del estado actual de la atmósfera utilizando un pronóstico a muy corto plazo y adicionando las observaciones disponibles para ello con el fin de lograr una mejor descripción del verdadero estado actual de la atmósfera. Este método es similar a la aproximación Sinóptica esto es: datos son colectados por un sistema de telecomunicación global en un intervalo de tiempo determinado (digamos 0000 o 1200 GMT) para especificar las condiciones iniciales. Las ecuaciones del modelo son resueltas entonces para varios segmentos de un mapa de tiempo para calcular que tanto se espera que las condiciones esperadas cambien para un tiempo dado, digamos, 10 minutos. Con tales cambios añadidos a las condiciones iniciales, un nuevo mapa es generado válido, en el ejemplo, para las 0010 o 1210 GMT. El nuevo mapa es tratado como un nuevo set de condiciones iniciales, probablemente no tan exacto como las medidas tomadas en las 000 y 1200 GMT pero aún muy aproximadas. El proceso se repite para generar un pronóstico para las 0020 o 1220 GMT y así en adelante. En principio, el proceso podría repetirse indefinidamente. En la práctica, pequeños errores en las mediciones de las condiciones iniciales, entre muchos otros debidos a la naturaleza no lineal del sistema atmosférico, se acumulan con cada iteración realizada en el modelo numérico. Eventualmente estos errores se vuelven lo suficientemente grandes y el pronóstico deja de tener validez.

La era de la predicción numérica realmente empezó en la década de los 50. Al tiempo que la capacidad de cómputo creció, también lo hizo la complejidad, velocidad y capacidad para los detalles en los modelos, cuando nuevas observaciones estuvieron disponibles a partir de fuentes como los satélites y los sistemas de radar, también se crearon nuevas técnicas lo suficientemente sofisticadas para la procesar a los datos en los modelos. Todos los modelos numéricos de la atmósfera se basan en el mismo grupo de ecuaciones que gobiernan la atmósfera pero difieren entre sí en las aproximaciones y suposiciones hechas en la aplicación de las mismas y en como son resueltas para la representación de los diferentes procesos físicos.

En la década de los 60's, los centros encargados de realizar pronóstico empezaron a utilizar versiones simplificadas de las ecuaciones Navier-Stokes pero, debido a limitaciones computacionales y operacionales, sólo eran utilizadas en un hemisferio a la vez. Sin embargo, pronto fueron usadas para realizar pronósticos globales.

Los modelos más sofisticados distan de ser perfectos, principalmente cuando la escala espacial a utilizar se reduce y las parametrizaciones de nubes dejan de ser válidas, llevando a errores en pronósticos de la cantidad de precipitación (Stern et. al, 1986).

Desde las primeras simulaciones y pronósticos numéricos de la circulación de la atmósfera y del tiempo meteorológico se ha trabajado en la evaluación de los resultados usando distintas metodologías. La comparación con campos observados es la única forma de medir la calidad de un modelo. Dicha evaluación puede ser cualitativa o cuantitativa (Wilks 1995). La primera está basada en el análisis de gráficas, mapas de contornos y diagramas. La segunda opción utiliza índices estadísticos para comparar similitud entre campos. Ejemplos de tales índices son el error cuadrático medio o la desviación estándar.

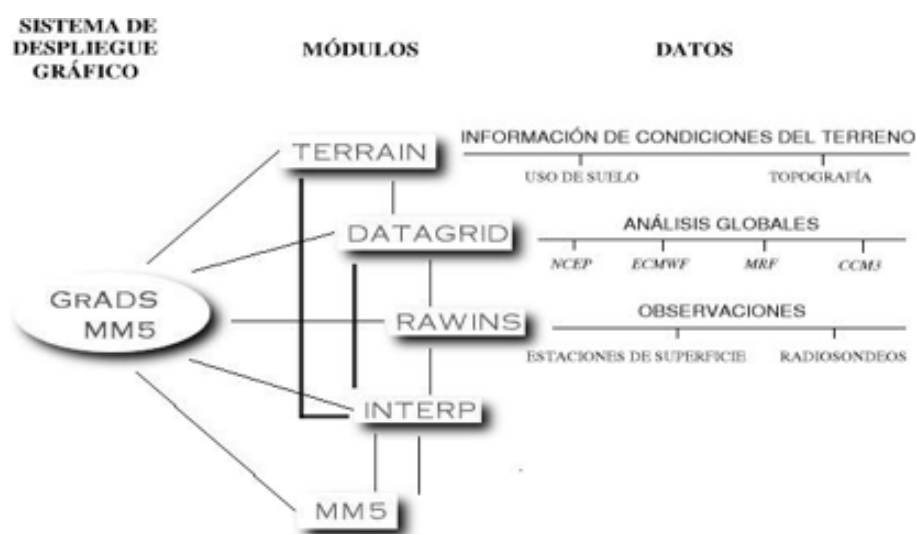
En los primeros pronósticos del tiempo se utilizaron reportes locales y cartas sinópticas para la evaluación puntual de campos de viento, temperatura y altura geopotencial (Charney *et al* 1950, Saucier 1955). En principio, la comparación de patrones era difícil por la falta de observaciones en todo el dominio de interés. Sin embargo, a partir de la Segunda Guerra Mundial, la necesidad de conocer el estado futuro de la atmósfera obligó a realizar más mediciones meteorológicas que permitieran analizar las condiciones sinópticas en una región. El primer intento de realizar un análisis objetivo fue hecho por Panosfky (1949) para el campo de altura geopotencial a 700 mb. Sin embargo, no fue sino hasta el trabajo realizado por Gilcrhist y Cressman (1954),

---

basado en expansiones polinomiales, que se logró generar un campo observado en una malla regular que permitiera comparar observaciones y pronósticos.

Uno de los modelos de área limitada más populares es el Mesoscale Model Version 5 (MM5). La quinta generación del MM5 (MM5-PSU/UCAR, por sus siglas en inglés) es el resultado de diversos desarrollos en la Universidad de Pennsylvania (Warner y Anthes 1978), y en el National Center for Atmospheric Research (NCAR, por sus siglas en inglés). Por ser un modelo de dominio público para la comunidad de las ciencias atmosféricas, sus cambios y adecuaciones son probados por numerosos usuarios. Es por ello que el MM5 se ha convertido en un modelo cada vez más confiable y robusto. En México, el modelo numérico MM5 se ha utilizado para realizar pronósticos a corto plazo y obtener diagnósticos, determinando soluciones accesibles en problemas ambientales y de cambio climático regionalizados (Magaña, *et al*, 1998) y es empleado actualmente por el Servicio Meteorológico Nacional para emitir pronósticos válidos por 72 horas y por el Centro de Ciencias de la Atmósfera con pronósticos con extensión de 48 horas.

El MM5 es un sistema de modelación compuesto por módulos de preprocesamiento y postprocesamiento de datos con una estructura como la mostrada en la Fig. 3.11, y que se describen a continuación:



**Fig. 3.14:** Diagrama de flujo del funcionamiento del MM5

**TERRAIN:** Preprocesamiento que define los dominios madre y anidado, la resolución y el tipo de proyección del mapa. Además, asigna los valores de la topografía y las características del terreno.

**DATAGRID:** Preprocesamiento con el que a partir de los datos de gran escala de viento, humedad, temperatura, etc., (campos iniciales) se construyen las mallas de los dominios a trabajar con la resolución deseada, generando así análisis de datos interpolados con alta resolución espacial, listos para ser procesados por los módulos de RAWINS o INTERP. La información utilizada puede ser histórica, para estudios de diagnóstico o simulación, o en tiempo real para pronóstico operativo del tiempo.

**LITTLE R:** Preprocesamiento donde se realiza un análisis objetivo de la información utilizando datos observados en superficie de estaciones meteorológicas y de altura, como radiosondeos, globo piloto, etc. La asimilación de datos tiene la finalidad de mejorar los campos de las variables utilizados como condición inicial.

**INTERP:** Preprocesamiento donde se interpola de los datos iniciales de coordenadas de presión a coordenadas verticales sigma. Este tipo de conversión de coordenadas permite una mejor representación de la orografía

(Yong-Fu y Zhong 1987).

Mientras que para las condiciones de frontera para áreas oceánicas se puede utilizar, por ejemplo, la temperatura del océano como delimitante; para zonas continentales, la topografía y el uso de suelo son fundamentales para el MM5 como condición de frontera, pues en gran medida son los parámetros responsables de generar circulaciones de mesoescala. La versión 2 MM5 incluye ya coordenadas verticales que se pueden utilizar para representar a la topografía de la superficie.

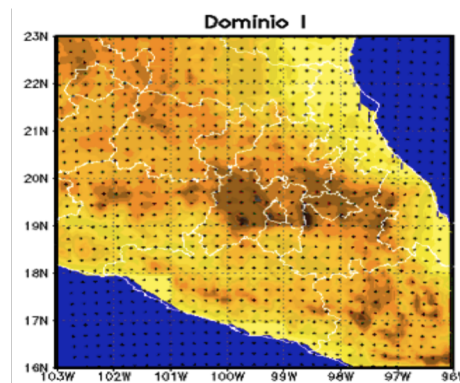
### 3.2.2. Los datos para el MM5, condiciones para el experimento

Como el MM5 es un modelo de área limitada, requiere de condiciones iniciales y de frontera. Las condiciones iniciales se obtienen a partir de análisis de malla de variables meteorológicas interpoladas a una resolución espacial seleccionada en el modelo. Para producir condiciones laterales o de frontera es necesario tener datos en una malla global que cubra el periodo de tiempo completo en el cual el modelo es integrado. Para las modelaciones numéricas de este estudio, las condiciones iniciales y de frontera corresponden a predicciones operativas y campos predichos del NCEP FNL *Operational Model Global Tropospheric Analyses* (Joyce *et al* 2004). Los datos están en mallas globales de  $1^{\circ} \times 1^{\circ}$  continuamente cada 6 horas (00Z, 06Z, 12Z, y 18Z) y los análisis están disponibles en superficie y en 26 niveles de presión desde los 1000 hasta los 10mb. Dentro de los parámetros que encontramos en los datos están: presión en superficie, presión a nivel del mar, altura geopotencial, temperatura de la superficie del mar, humedad relativa, intensidad y dirección del viento, movimiento vertical, vorticidad, entre otros.

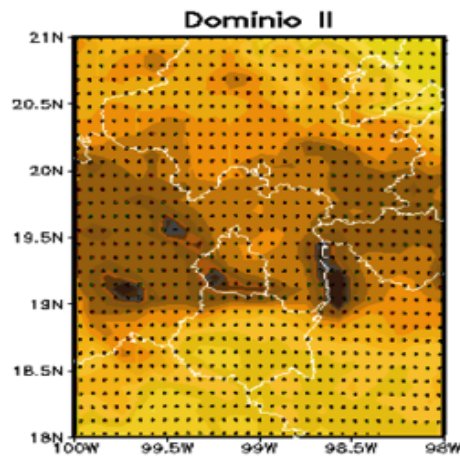
Desde 1999 el MM5 ha sido adaptado para funcionar operativamente por el Grupo de Meteorología Tropical del Centro de Ciencias de la Atmósfera y realizar el pronóstico para la cuenca de México. La configuración del modelo incluye:

- Un dominio madre de  $31 \times 31$  puntos con una resolución espacial de 24km centrada a  $92.2^{\circ} \text{O}$  y  $19.1^{\circ} \text{N}$ . La zona central de la Republica Mexicana
- Un dominio anidado de  $31 \times 31$  puntos con una resolución espacial de 8km.
- Proyección tipo Lambert.

- 23 niveles sigma en la horizontal desde la superficie hasta 100mb.
- Topografía del terreno con una resolución de 9km.
- Uso de suelo con resolución de 17km.
- Un esquema de parametrización *cúmulus* Kain-Fritsch.
- Esquema de radiación cada 30min.



**Fig. 3.15:** Dominio 1 utilizado en el MM5, que abarca el centro de la República Mexicana



**Fig. 3.16:** Dominio 2 utilizado en el MM5 que abarca los estados de México D.F, Morelos, Estado de México, Hidalgo, Tlaxcala y una parte de Guerrero y de Puebla



---

La parametrización para nubes *cumulimimbus* y las condiciones de frontera y de radiación pueden ser escogidas en el MM5 de una variedad de opciones, las que fueron seleccionadas para este estudio han probado reproducir adecuadamente las características sinópticas del clima del valle de México (Magaña *et al*, 2003). Las salidas de la simulación numérica y la predicción numérica operacional son almacenadas para cada hora a fin de tener una descripción temporal relativamente detallada de los sistemas meteorológicos.

Las precipitaciones más intensas se presentan en los meses de verano, debido a que durante la mitad caliente del año, la zona subtropical de alta presión se desplaza hacia el norte y la zona expuesta a los vientos alisios experimenta un aumento en la intensidad de estos, al tiempo que modifica su latitud y altura, por lo que prácticamente todo el país queda bajo la influencia de estos vientos que soplan de Este a Oeste al tiempo que recogen humedad del Golfo de México. La parte centro y sur de la República queda bajo la influencia de la zona inter tropical de convergencia que se desplaza hacia el norte en esta época. (Amparo de Miranda 2003). Así mismo, en el verano, existen mayor cantidad de tormentas convectivas, las cuales tienen su origen en el calentamiento de la superficie terrestre. El aire, al entrar en contacto con esas “zonas cálidas” se calienta más que en los alrededores, lo que da lugar a corrientes verticales con las que asciende aire caliente húmedo. Estas corrientes al llegar a la capa de la troposfera, se enfrían rápidamente, produciéndose la condensación del vapor de agua y formándose nubes muy densas del tipo *cúmulus*. Este tipo de precipitaciones se ven reforzadas por la interacción del flujo de humedad con las montañas que rodean al Valle de México, ya que el viento al encontrarse con un obstáculo se eleva, generando el mecanismo descrito antes. Las lluvias convectivas se presentan solamente en áreas reducidas ya que el ascenso y descenso de las corrientes sólo dan un espacio muy local (Ahrens 2000). Diversos autores (Jáuregui 1996) han estudiado, además, la relación de las llamadas *islas de calor* con la formación de lluvia convectiva sobre zonas urbanas. Así pues, para estudiar precipitaciones abundantes se obtuvieron datos horarios para los meses de verano: mayo, junio, julio y agosto de los años 2006 y 2007 respectivamente.

### 3.2.3. Desempeño del MM5 en pronósticos de lluvia para el Valle de México

El MM5 es capaz de reproducir las condiciones diarias del tiempo meteorológico en forma cualitativa. Sin embargo, en un análisis cuantitativo comienza a exhibir errores tanto sistemáticos como aleatorios. En simulaciones de precipitación realizadas a 24 horas en todo el Valle de México (Pérez, 2004) se observa que modelo reproduce entre un 40 % y un 70 % la ocurrencia de un evento de lluvia. Incluso la estructura de los patrones de lluvia pronosticados captan con buena aproximación la circulación en superficie. Sin embargo, estos pronósticos presentan limitaciones en cuanto a la predicción de la cantidad de lluvia, específicamente cuando se trata de un evento de precipitación extrema, cuando el MM5 muestra una marcada incapacidad para reproducir la cantidad de lluvia (aún con un margen de error) sobre la ciudad, reduciéndose la confianza en las simulaciones a tan sólo un 20 % o 30 % de acierto (Pérez, 2004). Aún así, el MM5 genera una buena aproximación del patrón de lluvia, más aún cuando el evento de lluvia no es forzado por la componente orográfica. El problema anterior puede estar asociado a las limitaciones en las parametrizaciones de nubes y a las condiciones iniciales. y para su corrección se requeriría de un mayor entendimiento en la microfísica de nubes, la cual posteriormente permitirá una mejor representación de los procesos convectivos y su asimilación numérica a escalas menores. En segundo lugar, se necesitaría la incorporación sistemática de datos de estaciones de superficie y altura, de una red relativamente densa. El MM5 llega a alcanzar hasta un 77 % de confianza en el pronóstico de lluvia acumulada diaria que puede reflejar la eficacia del MM5 para simulaciones de lluvias sin presencia de eventos intensos (por ejemplo en el invierno o meses con escasa precipitación). Sin embargo, la situación es distinta conforme avanza la temporada de lluvias. En el mes julio en donde además del aumento en la cantidad de lluvia acumulada, se presentan más casos de lluvias intensas, el MM5 sólo puede alcanzar poco menos del 40 % de confianza en las simulaciones. Así que en las simulaciones del MM5, resulta difícil estimar no sólo la cantidad acumulada de lluvia, sino también la ocurrencia de un evento extremo de lluvia. Pero a pesar de las limitaciones relacionadas con errores inherentes al modelo MM5, los productos del MM5 funcionan como una herramienta de pronóstico del tiempo al poder determinar la ocurrencia de lluvia o no lluvia. Durante el periodo de lluvias, la calidad del pronóstico hecha por MM5 disminuye. Esto significa que en particular en el verano,

existen errores dentro del pronóstico en ubicación e intensidad de precipitación. Esta investigación busca corregir esos errores aplicando una técnica tipo MOS.

Según Pérez (2004) el modelo es capaz de reproducir en forma aproximada los patrones de viento en superficie sobre el Valle de México, llegando a simular los cambios locales, relativamente pequeños, asociados a la compleja topografía del centro de México. Sin embargo en cuanto a magnitud del viento, el MM5 sobreestima su valor en un orden de 2 a 6 m/s. La intensidad y la dirección del viento en MM5 son bastante sensibles a la condición de uso de suelo. En ese sentido, se requeriría un análisis de capa límite para corregir los errores sistemáticos. Otro error presente en las simulaciones del MM5 para el Valle de México es que es necesario definir con mayor detalle la presencia de las montañas que rodean al Valle (en altura) ya que éstas definen el flujo de vientos predominantemente del Este que al final, es uno de los parámetros que modula a la formación de lluvia y su distribución en una zona como el Distrito Federal

### 3.3. *MOS y Modelación Estadística.*

La resolución espacial típica de los GCM (300 km X 300 km) no permite considerar las condiciones del clima regional (Magaña *et al*, 2010). Sin embargo, éstas se pueden obtener a través de la aplicación de técnicas de reducción de escala (downscaling) estadísticas o dinámicas a las salidas de un GCM. Las técnicas estadísticas de reducción de escala se refieren a métodos en los que los cambios regionales o locales, correspondientes a procesos de menor tamaño que la resolución espacial del GCM, son calculados como función del clima de gran escala. En estas técnicas, las variables del clima regional o local se obtienen generando un modelo estadístico que las relaciona con las variables de gran escala del modelo de circulación general.

La reducción estadística del clima provee una aproximación alternativa para usar la dinámica de los modelos numéricos en el pronóstico del clima a baja resolución espacial. Las técnicas de reducción dinámicas, sobre todo utilizan los Modelos de Clima Regional (RCM) que son similares a los modelos de clima global pero poseen mayor resolución espacial de tal modo que, dependiendo de su resolución, se puede también resolver explícitamente

algunos de los procesos que son parametrizados en los GCMs.

La base de la reducción de escala estadística es que existen relaciones entre el clima local y el de gran escala (dinámicas o estadísticas), que permiten determinar condiciones locales a partir de información de baja resolución, y que estas relaciones permanecen válidas aun bajo condiciones climáticas futuras (Caetano *et al*, 2007), de tal modo que los métodos de reducción de escala estadísticos identifican, dentro de datos históricos de dos conjuntos de variables, una relación *significativamente* consistente entre los valores observados de ambos conjuntos. Una relación *significativamente* consistente es aquella que está suficientemente presente en los datos como para sea improbable que haya ocurrido por casualidad y que provee un nivel de confianza razonable a partir de la cual establecer una predicción (Simon *et al*, 2007). Esta relación se puede obtener a través de establecer funciones de transferencia, construidas a partir de la clasificación de patrones de circulación y que constituyen las relaciones estadísticas entre los valores de condición de gran escala en un punto de altura o superficie y las condiciones del clima en un sitio específico.

En los métodos de reducción de escala estadísticos, las variables del clima regional o local (llamadas predictandos) se obtienen generando un modelo estadístico (una función de transferencia) que las relaciona con las variables de gran escala del GCM (llamadas predictores). Como primer paso en la reducción de escala se requiere definir relaciones estadísticas a partir de datos observados y de modelo, por lo que en primer lugar se necesita identificar variables climáticas de gran escala o variables independientes o predictores (Wilby *et al.*, 2000). Son buenos candidatos a predictor aquéllos que:

- Son física y conceptualmente relevantes con respecto a las variables del sitio (la variable dependiente o predictando).

- Están fuerte y consistentemente correlacionadas con el predictando, es decir el parámetro local que se trata de reproducir

- Están disponibles en archivos de datos observados o salidas de GCM.

- Son modelados adecuadamente por el GCM.

Una vez que las funciones de transferencia han sido calculadas, se debe

---

proceder a la verificación del modelo estadístico. Esto se lleva a cabo utilizando un conjunto de datos independiente de aquellos con los que se construyó el modelo. La verificación del modelo estadístico consiste en comparar los valores reportados (observados) en la estación con los predichos por el modelo estadístico para el mismo periodo. Una vez que se determina la calidad del modelo, éste puede ser usado para la generación de predicciones. Determinar qué tan bien o mal lo hace el modelo es en cierta medida subjetivo y dependerá de la zona bajo estudio y de los datos disponibles.

### 3.3.1. El Climate Predictability Tool (CPT)

Existe una gran variedad de herramientas de reducción de escala de las salidas de los GCMs, tanto dinámicas como estadísticas, que en su mayoría son de acceso libre, entre las herramientas dinámicas tenemos al: MM5, PRECIS, RegCM3. Dentro de las herramientas estadísticas tenemos al: SDSM y CPT.

Para la corrección de errores sistemáticos en las predicciones de lluvia acumulada horaria del MM5 se optó por utilizar la herramienta de cómputo llamada Climate Predictability Tool (CPT) la cual permite relacionar patrones de gran escala con condiciones regionales del clima al establecer las funciones de transferencia de un campo de alta resolución a uno de baja resolución para cada registro de una serie de tiempo. CPT fue diseñado específicamente como corrector de errores sistemáticos a salidas de GCM mediante la técnica conocida como Model Output Statistics (MOS) (Glahn, *et al* 1972), que es análoga a la reducción de escala y similar a la llamada *perfect prog* al relacionar patrones espaciales en un campo independiente y otro de respuesta. La técnica MOS puede corregir los errores identificados empíricamente en las salidas de modelos.

El esquema CPT permite aplicar las técnicas estadísticas de reducción de escala de regresión por componentes principales (PCR, por sus siglas en inglés), análisis de correlación canónica (CCA, por sus siglas en inglés) y regresión múltiple a cualquier variable, haciendo accesibles otros resultados de estadísticas básicas como: Funciones Empíricas Ortogonales (EOFs, por sus siglas en inglés) y patrones canónicos a través de los mismos EOFs. De tal modo que CPT es una herramienta de regionalización de pronósticos es-

---

tacionales basada en la corrección de errores sistemáticos en las salidas de los modelos que se consigue identificando patrones espaciales.

El CPT fue desarrollado por el International Research Institute for Climate and Society (IRI) para aprovechar las salidas de los modelos de circulación general y estimar con mayor detalle condiciones regionales del clima al tener la posibilidad de hacer predicciones de alta resolución espacial. El CPT, realiza también la validación del modelo ya que utiliza análisis de CCA y PCR sobre cualquier base de datos.

Con la posibilidad de aplicar diferentes técnicas estadísticas (estandarizar los campos de los GCM, muestrear los datos, normalizar los campos, sustituir valores faltantes o aplicar diferentes técnicas para derivar las EOFs) con el objeto de optimizar las ecuaciones de reducción de escala, CPT es una herramienta muy útil para la predicción del clima de mesoescala. Ha de tomarse en cuenta, sin embargo, que la suposición más importante para utilizar CPT y que es inherente a todas las técnicas estadísticas, es suponer que la relación entre el campo de gran escala de baja resolución (proporcionada por el CGM) y el de alta resolución (a nivel local) permanece constante. Dicha condición obviamente no considera los cambios repentinos.

Entre las ventajas de CPT están:

- 1) Es computacionalmente económica y por ello permite hacer gran número de regionalizaciones
- 2) Permite la estimación de incertidumbre, medida a través de la dispersión entre regionalizaciones
- 3) Dependiendo de la escala espacial del campo observado de referencia (eg.  $0.5^\circ \times 0.5^\circ$ ) se pueden realizar regionalizaciones a la medida de las necesidades del usuario
- 4) Se puede implementar para prácticamente cualquier variable climática.

Entre las desventajas:

- 1) Parte de la hipótesis de que las relaciones de las variables en los campos de baja resolución a alta resolución no cambian
- 2) Se debe considerar los errores de reducción de escala en el esquema de CPT
- 3) No permite análisis de eventos extremos de manera directa porque no

construye campos diarios

Una de las suposiciones fundamentales en el proceso de escalamiento mediante técnicas estadísticas es que la relación de la variable independiente (del GCM) y el campo observado, esté basada en procesos físicos. La relación entre las variables en esta investigación es directa, es decir: precipitación de baja resolución con precipitación de alta resolución, por lo que su relación es meramente de corrección de errores sistemáticos, más que del comportamiento de los campos predictores en función del desarrollo de una variable.

### 3.3.2. Adaptación del CPT para el pronóstico de eventos de lluvia horarios

La interfaz del CPT fue originalmente concebida para realizar predicciones estacionales, esto es: CPT opera con datos de entrada mensuales, para una cantidad de años de datos de las salidas de un modelo numérico de circulación global. Como uno de los objetivos de esta tesis es estudiar el comportamiento de eventos de lluvia extremos, fue necesario reducir la dimensión temporal de las predicciones realizadas en CPT. El procedimiento de escalamiento temporal consistió en generar estimaciones horarias que generaran las condiciones diarias, que ajusten por si mismas, la condición de precipitación media mensual.

Para incluir datos de otra resolución temporal es necesario adoptar una escala anual en CPT, pero horaria en la realidad, esto es: cada archivo de entrada debe contener todos los tiempos (años para CPT) de un solo mes, variando solamente el índice del año. Es decir: colocando la bandera de tiempo (en CPT) en el 1 de Enero de 1901 al campo de lluvia correspondiente a las 00Z del primer año de la base de datos, de este modo, las 06Z del mismo año y del mismo mes, equivaldrían en CPT, al 1 de enero de 1902. Así, podemos colocar en CPT todos los campos horarios necesarios.

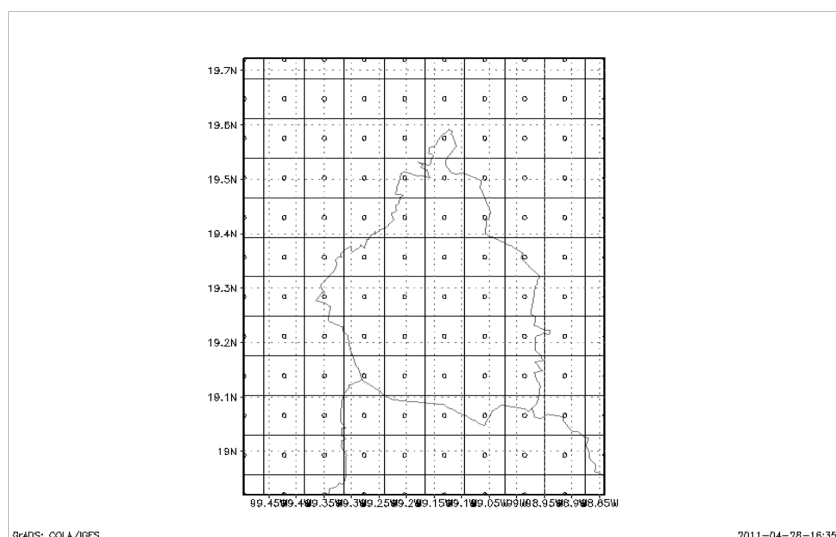
---

TIEMPO EN CPT	TIEMPO REAL
1 Enero 1900	1May200600Z
1 Enero 1901	1May200601Z
1 Enero 1902	1May200602Z
1 Enero 1903	1May200603Z
...	...
1 de Enero 1923	1May200623Z

**Fig. 3.17:** Adecuación de escalas temporales en CPT. Se muestra un día, de las 00Z a las 23Z.

Se adoptó un esquema de procesamiento de datos como el solicitado por el CPT (utilizando dos conjuntos de datos independientes (predicandos y predictores)), pero con una modificación: como predictores, en vez de usar las salidas estacionales de un modelo de circulación global, se utilizó un modelo numérico de mesoescala (MM5) y sus predicciones horarias. Como predictandos, utilizamos el conjunto de datos de observaciones horarias de estación y de satélite (SACM, CMORPH) interpolados a punto de malla. Con estos elementos, se generaron campos de lluvia acumulada horaria en una malla con dominio de 320 puntos de malla 168km x 128km ( 21 x 16 puntos) centrado entre los 98.4° y 99.7° de longitud oeste y entre los 18.8° y 19.8° de latitud norte.





**Fig. 3.18:** Dominio de la malla utilizada (21 x 16 puntos) en el CPT

### 3.3.3. Funcionamiento de CPT

CPT trabaja con campos obtenidos de un periodo histórico de simulación del modelo numérico, y tiene la capacidad de construir una ecuación de transferencia para cada registro (Simon *et al.*, 2007). Las ecuaciones de regresión se construyen con un periodo de calibración y otro de validación cruzada, que en la práctica consiste en probar la función de transferencia original en una muestra de datos independiente. Así, la base de datos del control se divide en subconjuntos. Con el primero se calibra la función de transferencia, mientras que con el otro subconjunto (muestra independiente) se confirma y valida el análisis original. Cuando las ecuaciones de regresión se realizan usando PCR, el esquema construye tantos modelos estadísticos como EOFs puedan ser calculados en la variable independiente (los datos del modelo). En CPT los EOFs pueden ser calculados de varias maneras: utilizando la matriz de covarianza, la de correlación o por mínimos cuadrados. Una vez calculados, CPT realiza regresiones lineales multivariadas para calcular las ecuaciones de regresión utilizando a las componentes principales como predictandos en las ecuaciones de regresión, de tal modo que optimizar el número de EOFs (evaluando el desempeño que tuvo cada combinación de EOFs en explicar las características del campo observado en periodo de validación cruzada.) es importante, pero no determinante en la generación de un buen modelo

estadístico. Para plantear las ecuaciones de regresión, en este trabajo solo se hizo uso del el análisis de Regresión por Componentes Principales (PCR).

El esquema CPT provee una evaluación de significancia estadística y desempeño sobre el periodo de validación mediante índices como: Características Relativamente Operativas (ROC, por sus siglas en inglés), el error cuadrático medio, la correlación lineal etc, (Magaña et al, 2010). También reconstruye con alta resolución el periodo con el que fue calibrado, lo que provee una referencia de sus capacidades y deficiencias en la simulación. *Nuestra hipótesis es que con este modelo de regresión óptimo y el campo de pronóstico del modelo numérico de mesoescala, pueden calcularse pronósticos de alta resolución de lluvia acumulada sobre una región como la del Distrito Federal.*

Ya que dos bases de datos son requeridos por CPT, para establecer funciones de transferencia, utilizamos en la predicción a las dos bases de datos generadas. 1) Como predictores: Campos de lluvia horaria generados por el MM5 del 1 mayo al 31 de agosto de los años 2003, 2004, 2005, 2006, 2007 y 2008, y 2) como predictandos: los campos de lluvia horaria generados a partir de observaciones de estación y de satélite interpoladas a punto de malla utilizando el análisis de Cressman para el mismo periodo de tiempo.

#### 3.3.4. Calibración del Modelo Estadístico y Pronóstico Cuantitativo.

Uno de los primeros criterios que se debe considerar en la generación de un pronóstico en CPT es la definición del período de tiempo a utilizar (debido a la posibilidad de obtener distintos resultados al utilizar distintos intervalos de tiempo). Éste puede definirse de dos maneras: el intervalo de tiempo para el periodo de entrenamiento o calibración, y el primer año del periodo de calibración. El año de inicio para el periodo de calibración que se consideró en la tesis fue el 1 de enero de 1900 (en realidad, el 1 de mayo de 2006, a las 00Z). Se construyeron las funciones de transferencia considerando un periodo de calibración de 2 años de datos horarios, que equivalen a 5760 años en el CPT (1900-5760 en CPT; mayo, junio, julio y agosto del 2001 y los mismo meses para el 2003, en la realidad). Después de realizar diversas pruebas con el análisis de validación cruzado. Los años restantes (2004, 2005, 2006, 2007, 2008) se conservaron para validación, es decir como muestra independiente a partir de la cual se calibrarán las funciones de transferencia.

La elección del número de EOF's utilizados para establecer las ecuaciones de transferencia es de relativa importancia ya que CPT considera en las ecuaciones de regresión, a las componentes principales de los datos originales como predictandos, sin embargo, no hay un método para optimizar el número de EOF's correcto para representar una base de datos y esto tiene que hacerse realizando varias pruebas en las que se mide el desempeño del modelo con respecto al periodo independiente en el análisis cruzado. Realizado esto, se eligieron los primeros 7 modos de variabilidad para la representación del bloque de datos de independiente, debido a que los distintos índices que tiene CPT para la verificación del modelo estadístico en el periodo de calibración, en particular los que miden la habilidad y el error del modelo estadístico, presentarán valores aceptables. Una vez calibrado el modelo, es posible efectuar pronósticos en CPT a partir de considerar una relación lineal con los nuevos valores optimizados de los predictores. Es decir: para observar el desempeño de CPT al corregir errores sistemáticos del MM5 al realizar pronósticos en alta resolución espacial, se realizó un pronóstico cuantitativo de precipitación (QPF, *Quantitative Precipitation Forecast*) el cual consiste en pronosticar la cantidad de precipitación acumulada en un periodo de tiempo determinado sobre una área específica (Bushong, 2005). Para este fin se eligió un periodo de estudio que representara los meses de precipitación mas importante para un verano completo por lo que el QPF se realizó desde el 1 de mayo hasta el 31 de agosto del 2008 (desde el 5070 hasta el 19300 en CPT) sobre la misma malla de 1168km x 128km ( 21 x 16 puntos) utilizada anteriormente.

Un esquema de QPF debe contestar a las siguientes preguntas:

-¿Dónde lloverá?

-¿Cuándo lloverá?

-¿Qué tanto lloverá?

Del mismo modo, el QPF debería poder reconocer cuales patrones espaciales son los capaces de producir lluvia intensa y cuales no. Es necesario mencionar que para trabajar con un QPF, se debe tener cierto conocimiento de la climatología local y de la física que hay detrás de los procesos que producen precipitación y los que no en una escala espacial y temporal alta.

En el área del manejo del riesgo hidrológico, los pronósticos cuantitativos de predicción (QPF) y de estimación de lluvia (QPE) son clave, por ejemplo: para cuantificar potenciales inundaciones, especialmente en escalas cortas de tiempo para cuencas urbanas, o en dar una respuesta en acciones de protección civil ante un evento extremo de precipitación. En este contexto, el pronosticar la lluvia acumulada puede ser entendido como el intento que se hace para reducir la incertidumbre futura del estado del sistema hidrológico y así anticipar acciones de mitigación. Los QPF son una herramienta útil para el usuario final de los productos de pronóstico, siempre y cuando éste conozca los umbrales zonales de precipitación que se tienen que superar para tomar una decisión.

Las autoridades encargadas, sin embargo, a menudo se muestran reacios a crear e invertir en estas acciones en base solamente de las predicciones numéricas cuando su calidad es desconocida. En otras palabras, para que los pronósticos sean útiles y eficaces, la calidad de la incertidumbre del pronóstico debe ser cuantificada.

## 4. RESULTADOS

### 4.0.5. Verificación de pronóstico cuantitativo

EL pronóstico realizado por CPT está representado por campos de lluvia acumulada en una malla reticular de 8km de resolución que se puede comparar directamente con observaciones de lluvia observados para el mismo tiempo e interpolados a una malla de resolución similar. Al comparar el campo pronosticado y el campo observado es posible evaluar cuantitativamente la exactitud de la predicción de la intensidad de la lluvia y de su ubicación (Rossa *et al*, 2007). Para generar una evaluación cuantitativa de un evento de lluvia individual es necesario entonces, asociar a un evento dado, con su contraparte en el análisis, comparando a los patrones de lluvia. En este sentido, se han desarrollado una multitud de medidas de verificación de la exactitud y de habilidad relativa de un pronóstico cuantitativo dependiendo de la naturaleza de los datos. Si los datos son continuos, como la temperatura, el viento, la presión en superficie, es usual que la exactitud de un pronóstico sea verificado usando estimadores estadísticos como: el sesgo, el error cuadrado medio, la raíz cuadrada del RMS, etc. En cuanto a la habilidad relativa, se puede utilizar un método para reducir la varianza. Por otro lado, si los datos son categóricos (que consideren la presencia/no presencia de algún evento, como la lluvia), la exactitud puede ser verificada a través de evaluar una tabla de contingencia para obtener, por ejemplo, un porcentaje de aciertos, el sesgo, la probabilidad de detección de ocurrencia o no de eventos de lluvia, o la razón de falsas alarmas (Rossa *et al*, 2007).

La verificación general de una bases de datos categóricos se logra evaluando una tabla de contingencia donde se asignan categorías (falso, acierto, falsa alarma, etc) a los resultados del pronóstico cuando son comparados con los datos observados para el mismo periodo de tiempo y obtener así un puntaje a partir del cual se calculan diversos parámetros.

De acuerdo a la sección *El problema de la Intensidad*, para el cálculo de por-

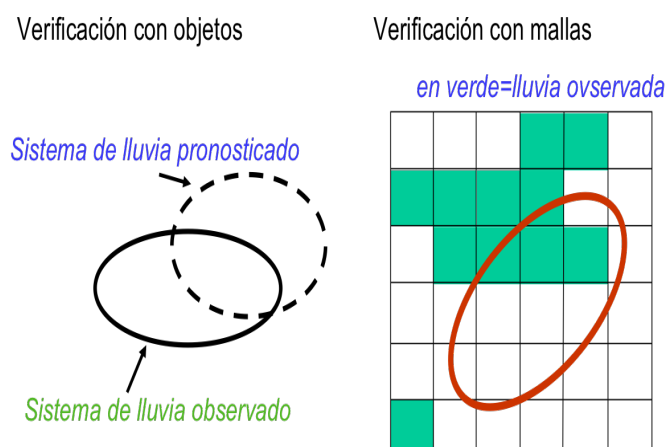
centajes solamente se consideraron los casos cuando hubiera sido observada lluvia, es decir: comparamos el caso de, cuando llueve, cómo pronostica CPT el patrón espacial. Así, no contabilizamos a los registros en donde CPT contabilizara lluvia y esta no hubiera sido observada, ver figura 4.1.

		OBSERVADO	
		si	no
PRONOSTICADO	si	a	b
	no	c	a

a="Acierto", b="Falsa Alarma", c="No detectado"

**Fig. 4.1:** Tabla de contingencia de 2x2

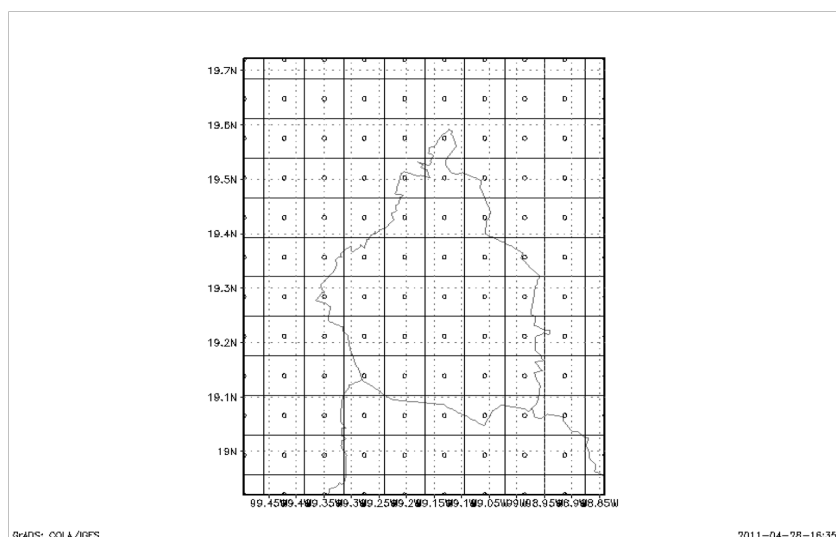
Cuando se está verificando un pronóstico cuantitativo cuya representación se da en una malla con cierta resolución, se utilizan las mismas categorías pero considerando a las áreas de intersección de retículas entre el campo de lluvias pronosticado y el campo observado. Se compara cada retícula pronosticada con su correspondiente observada. En realidad, lo que hacemos es superponer el patrón espacial del campo pronosticado sobre el observado, y contabilizar los casos en los que las áreas coinciden, o no. De esta manera podemos obtener mediciones estadísticas de verificación del QPF. Ver figura 4.2.



**Fig. 4.2:** Representación esquemática de la tabla de contingencia si tenemos datos en malla

#### 4.1. Análisis de la distribución del área de lluvia

Se realizaron medidas de verificación del pronóstico de CPT calculando puntajes de efectividad en una malla de 10 x 8 puntos (160 x 128km) centrada entre los 94.4° y los 93.5° de longitud y los 19.35° y los 19.7° de latitud, zona que abarca el área completa del Distrito Federal.



**Fig. 4.3:** Malla utilizada para el análisis de la distribución de área de lluvia

Se comparó la malla de precipitación horaria SACM ( de las mismas dimensiones) con la malla de precipitación horaria de CPT utilizando un método de comparación de puntos de malla al definir una tabla de contingencia que representara la ocurrencia o no de lluvia en cada punto.

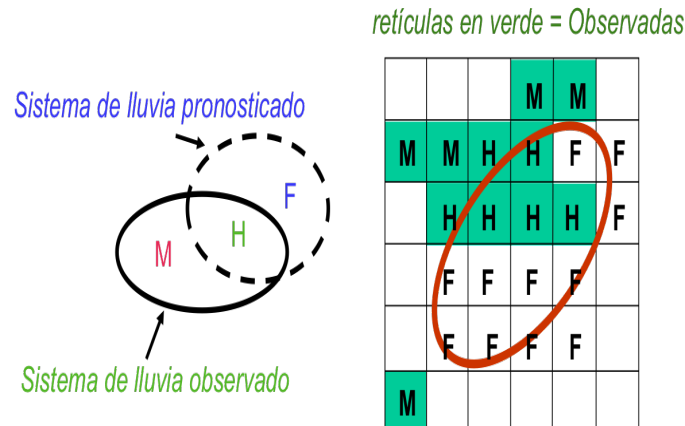
Siguiendo la discusión anterior, para verificar el QPF y obtener medidas de verificación, se utilizaron tres categorías:

*Acierto* o "Hit" (H). Si tanto en el punto de malla observada como en la pronosticada se presentó lluvia.

*Falsa Alarma* o "False Alarm" (F). Si se pronosticó lluvia en un punto de malla cuando no la hubo.

*No detectados* o "Missed" (M). Cuando la lluvia en un punto de malla si ocurrió pero no fue pronosticada en el mismo punto.





**Fig. 4.4:** Representación esquemática de la tabla de contingencia y las categorías utilizadas

Contabilizando cada caso para cada malla en el periodo del pronóstico (2928 registros, equivalentes a mayo, junio, julio y agosto del año 2008) se calcularon tres scores de exactitud :

Proporción de Falsas Alarmas (*FAR*, *False alarm Ratio*): divide el área de la predicción que no se superpone al campo observado por el tamaño de la zona pronosticada. El valor ideal de esta medida es cero.

$$FAR = \frac{F}{H+F}$$

Donde H=Acierto o Hit, M= no detectado o "missed".

Probabilidad de Detección (*POD*, *Probability of Detection o Hit Rate*). Se obtiene dividiendo la superposición entre las predicciones y los campos observados por el tamaño del campo observado: el resultado ideal es de 1

$$POD = \frac{H}{M+H}$$

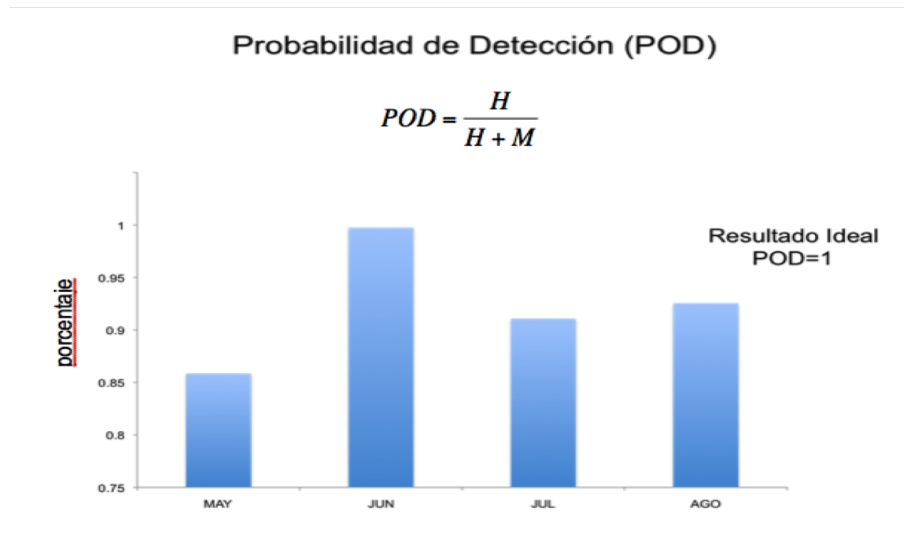
Donde F=Falsa alarma o "False", H=Acierto o "Hit".

Sesgo ó *Bias*. Compara el tamaño del campo de los pronósticos, con el campo observado, con una puntuación ideal de 1 y un rango de 0 a infinito. El sesgo mide la razón de frecuencias de ocurrencia de eventos pronosticados con respecto a la frecuencia de ocurrencia de eventos observados. El rango del sesgo está entre 0 e infinito, con un score ideal igual a 1 (la frecuencia de eventos pronosticados es la misma que la de los eventos observados). De tal modo, el sesgo mide si el sistema de pronóstico tiene una tendencia a subestimar en el pronóstico (Bias menor que 1) o a sobre estimar (Bias mayor que 1)

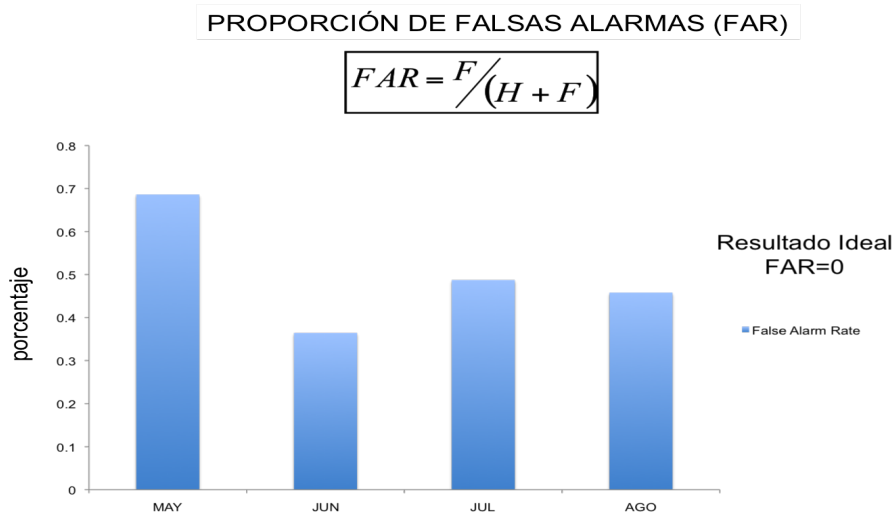
$$Bias = \frac{F+H}{M+H}$$

Donde H=Acierto o "Hit", M= No detectado o "missed", F= Falsa alarma o "False".

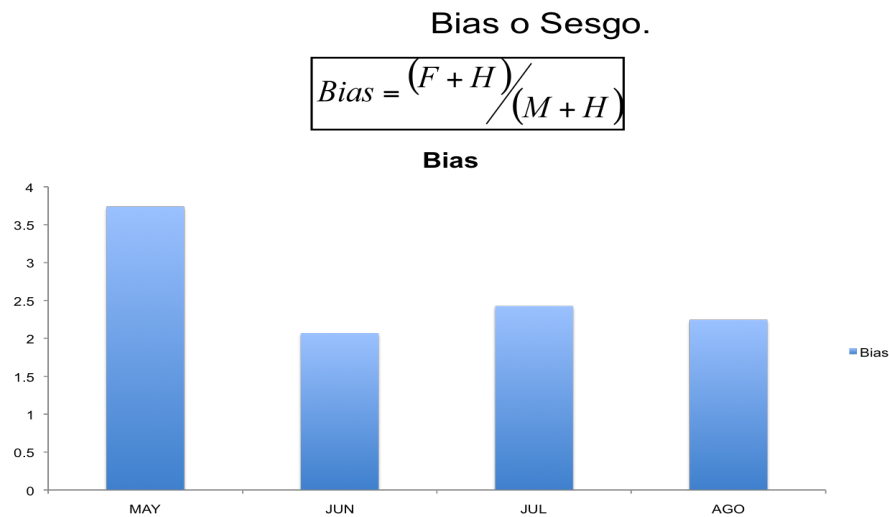
Se graficaron estos índices en función del tiempo, para cada mes del año de pronóstico, y así observar su comportamiento. A continuación, las gráficas expresadas (excepto el sesgo) como porcentajes.



**Fig. 4.5:** Probabilidad de Detección, POD



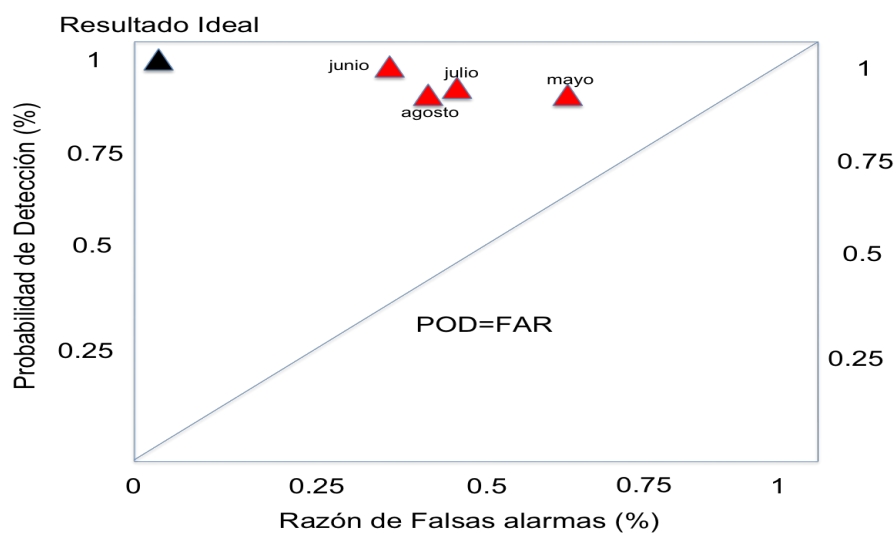
**Fig. 4.6:** Razón de Falsas Alarmas, FAR



**Fig. 4.7:** Sesgo

El primer paso a verificar debe ser si la probabilidad de acierto, para cada vez, es mayor que el FAR, así podemos tener una medida de la habilidad del pronóstico para pronosticar. Esto se consigue graficando a la probabilidad

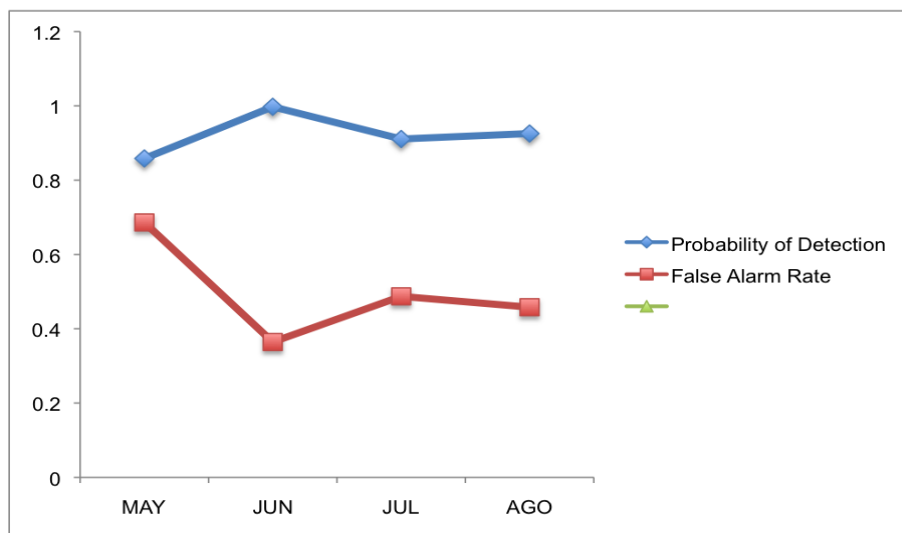
de acierto, en función del índice de falsas alarmas, en curvas conocidas como Curvas de Características Operativas Relativas (ROC, Relative Operating Characteristic). La predicción ideal, estará ubicado en la esquina superior izquierda (0,1) del diagrama ROC, lo que representaría 0% casos de falsas alarmas y 100% de aciertos. Una línea diagonal divide el diagrama de ROC. Puntos por encima de la diagonal representan buenas predicciones mientras que puntos por debajo de la diagonal representan pobres resultados. Siguiendo esta idea, mientras más cercanos estén los puntos a la diagonal, la calidad del pronóstico disminuirá.



**Fig. 4.8:** Diagrama de ROC para POD y FAR

En un primer resultado, el pronóstico que realiza CPT es bueno, ya que según la figura anterior, siempre tenemos casos en donde el número de casos de *aciertos* siempre es mayor al número de casos de *Falsas alarmas* pero de ninguna manera es perfecto, ya que todos los casos se encuentran alejados del punto de predicción ideal. De acuerdo a los resultados, existe *un problema* con la predicción realizada por CPT. Este consiste en la predicción de una lluvia continua y medida en ciertas zonas del DF (ver: *El problema de la intensidad*), lo que se refleja en el número de casos *Falsas alarmas* contabilizados, de este modo, en un mes seco, se presentará un mayor número de casos *Falsas alarmas*, y el índice FAR aumentará, ya que se pronosticará más lluvia de la que en verdad haya existido. Por el contrario, en un mes lluvioso,

el índice FAR disminuirá, debido a que ahora, la cantidad de lluvia pronosticada por CPT, será comparable a la lluvia observada para ese mes. De este modo, FAR y POD se encuentran correlacionados, al graficar estos índices en función del tiempo, esta correlación es apreciable. (ver figura 4.9). Para el mes de mayo, en donde la precipitación horaria es significativamente poca, el número de falsas alarmas es de casi a un 70 % (FAR=0.68618918), mientras que el POD es aproximadamente del 86 % , lo que no implica una alta habilidad para pronosticar, sino por el contrario, demuestra que CPT presenta dificultades para pronosticar lluvia para un mes donde no la hubo, es decir: al tener zonas de lluvia constante, será más fácil que las zonas de lluvia observada se superpongan a las primeras y puedan ser consideradas como *aciertos*. El sesgo es de aproximadamente 3.7, por lo que CPT sobre estima la cantidad de lluvia acumulada considerablemente. (tres veces el resultado ideal de 1).



**Fig. 4.9:** FAR y POD en función del tiempo para los cuatro meses de 2008 considerados del experimento

Los resultados para los meses de junio, julio y agosto del 2008, fueron, a nuestro parecer, muy buenos. Para el mes de junio, aproximadamente el 30 % de los eventos de lluvia no fueron pronosticados correctamente por CPT (FAR=0.364291442) lo que se refleja en el alto índice de probabilidad de acier-

tos (POD=0.997492987), mientras que es sesgo disminuye considerablemente con respecto al mes de mayo. (Sesgo=1.99). Para julio, el caso es similar, pero no tan acentuado, con una probabilidad de acertar en el pronóstico de casi 90 %, y un sesgo de 2.3. Mientras que para agosto la probabilidad de acierto se mantiene en este rango (POD=0.925519493) y un sesgo de 2.38.

## 4.2. *Análisis de la distribución del patrón espacial de lluvia acumulada horaria*

### 4.2.1. *El problema de la Intensidad.*

Para todos los registros, CPT pronosticó una lluvia constante y muy moderada, entre los  $10^{-5}$  (para el caso práctico es cero) y 0.6 milímetros por hora en la mayoría de los casos. Lo anterior, aunque en términos prácticos no tiene gran significado debido a que se consideran a estos casos como de no lluvia, si tiene diversas implicaciones en la eficiencia del modelo, las mas inmediatas son: la falsa existencia de registros de lluvia y la predicción de una intensidad por debajo de la intensidad observada. Por lo que este índice no puede ser considerado para verificación.

La presencia de lluvia continua y persistente se debe en gran medida a la manera en la que CPT realiza la calibración de las funciones de transferencia, ya que idealmente, los datos que conforman a los *predictandos* deberían estar normalmente distribuidos al momento de que CPT realice el cálculo de las PCR (esto porque CPT trabaja con datos de entrada mensuales). Por el contrario, los datos de lluvia a una escala diaria o horaria no siguen una distribución normal debido a la presencia de un gran número de valores de cero y ningún valor menor que el cero; por lo que estos registros siempre tendrán una naturaleza sesgada. CPT es capaz de calcular la intensidad de precipitación media *real*, pero luchará continuamente con el sesgo que presentan las intensidades de los datos.

Aunque no hay una solución simple para resolver este problema, debido a la falta de algoritmos apropiados para calcular intensidades de lluvia en datos no normalizados en el CPT (Simon Mason, en comunicación personal), es utilizar la opción *Transform data Y* en CPT que realiza una transformación de la base de datos de predictandos empírica, hacia una distribución normalmente distribuida antes de realizar la calibración del modelo. Para realizar los pronósticos en CPT vuelve a transformar los datos a su distribu-

ción original y realiza el procesamiento de cómputo posterior de la manera usual. Empero, los coeficientes de regresión en las PCR se calculan utilizando a estos datos transformados.<sup>1</sup> En el caso de datos de lluvia a alta resolución temporal, esta transformación no será tan exitosa debido a que los datos de precipitación acumulada están fuertemente limitados por el cero de precipitación. Sin embargo, un posible método para evitar esta situación sería utilizar la transformación de los *predictandos* hacia una distribución normal y realizar pronósticos definiendo antes un *cero* en los cálculos de CPT, lo que debería disminuir la cantidad de precipitación en los pronósticos, aunque sea un poco. Se podría utilizar después a CPT como herramienta para predecir la probabilidad de lluvia: colocando 0 en los registros de entrada sin lluvia, y 1 en los registros con presencia de lluvia. Al superponer estos resultados con los salidas del pronóstico se podría identificar a las áreas con menor probabilidad de precipitación.

La subestimación de la intensidad de lluvia pronosticada por CPT se debe también, (independientemente de las deficiencias del modelo) al poco tiempo que se utilizó en la generación del pronóstico (6 años de datos horarios) ya que el cálculo de funciones de transferencia están en función de la cantidad de datos de entrada, así como la verificación cruzada, que considera un cierto conjunto de datos independientes, para verificar el pronóstico.

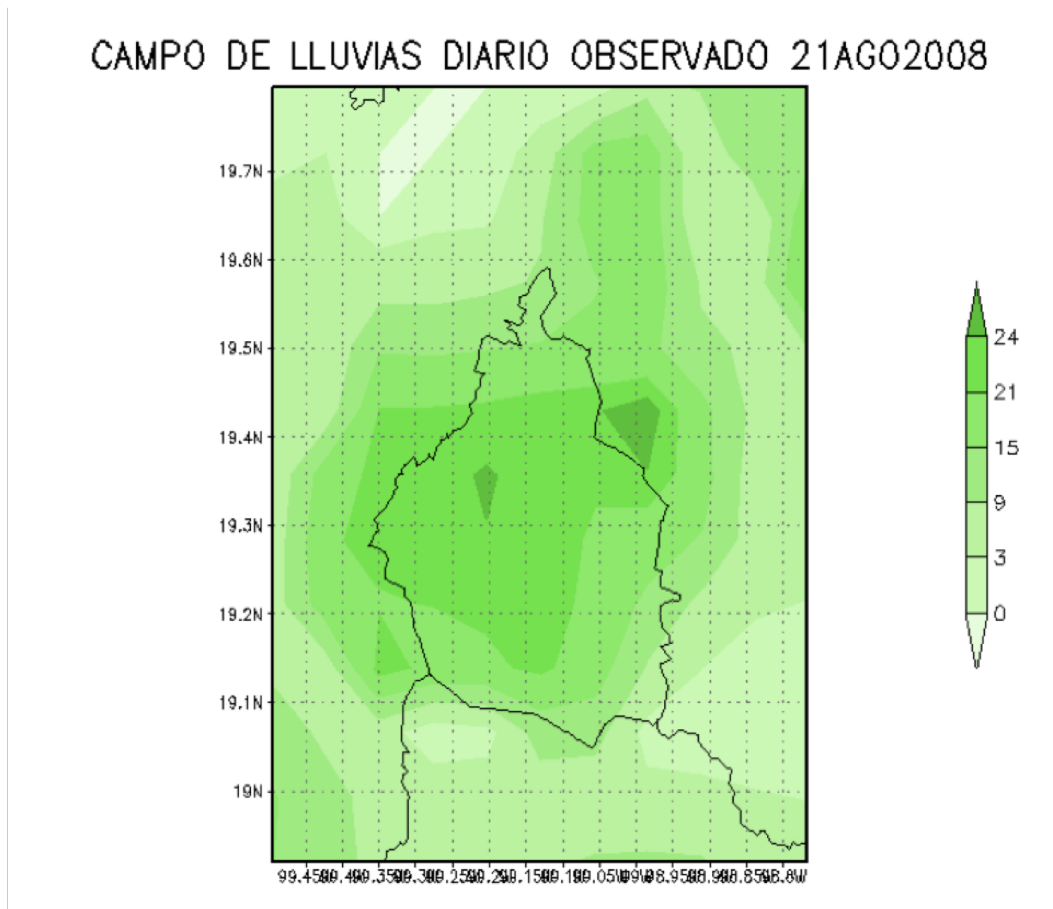
#### 4.2.2. *El Patrón de lluvia ajustado por el modelo estadístico en la region del Distrito Federal*

Para tener una idea mas clara del comportamiento de las salidas ajustadas por el modelo estadístico, una vez verificada la confianza sobre el pronóstico (sección 4.05), se utilizó prácticamente el mismo dominio definido en la

---

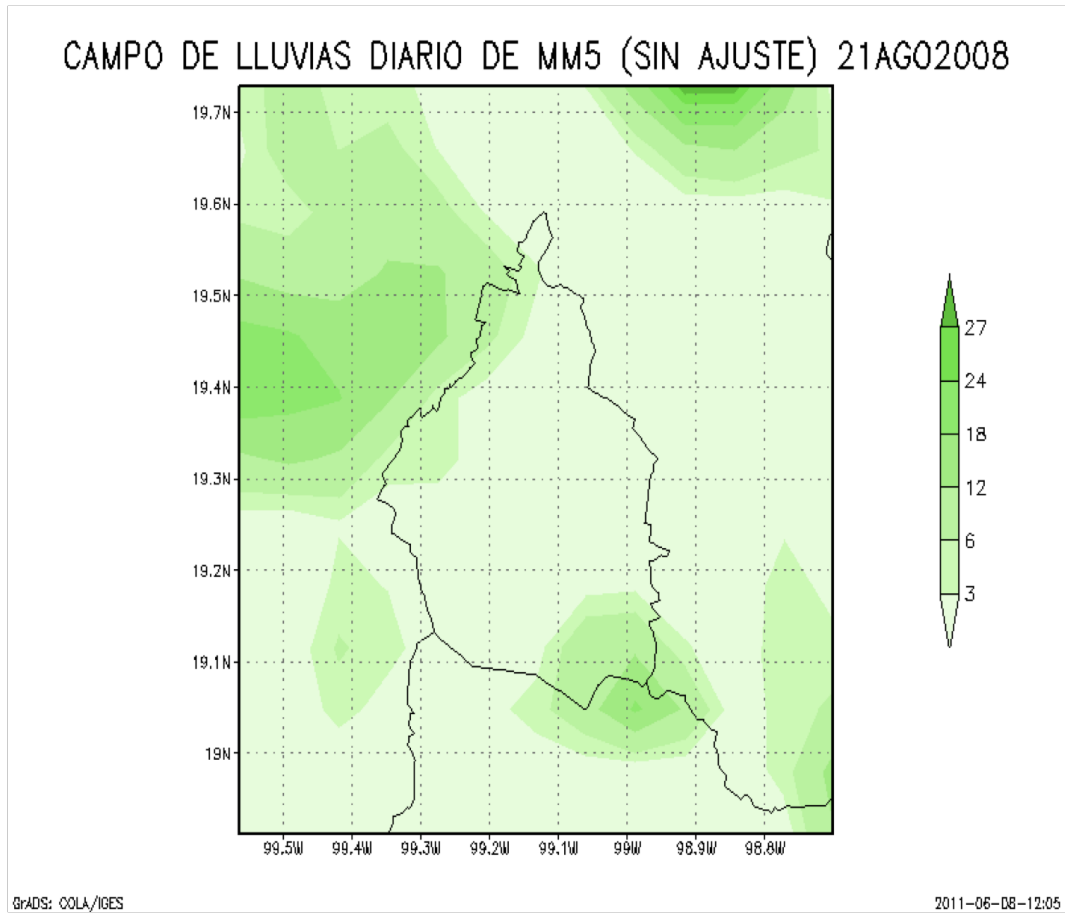
<sup>1</sup> El procedimiento para la transformación de los datos consiste en CPT en convertir a los datos originales en rangos, y expresarlos luego como percentiles de la distribución original, en suma, expresando a la distribución original en una distribución normal de rango unitario. Estos percentiles son usados después para calcular desviaciones estandar de la distribución normal. Los coeficientes de regresión de las PCR son calculados usando estas desviaciones estándar, que son luego transformadas de nuevo a la distribución original interpolando linealmente a los percentiles en los datos originales. Si el percentil se encuentra mas allá del rango de los datos originales, entonces otra interpolación lineal es aplicada basándose en la diferencia entre el valor mas extremo y el segundo mas extremo. (Mason, 2008)

sección 4.05, y sobre este se construyeron campos de lluvia acumulada para varias escalas en el tiempo y se les comparó con campos observados. Se obtuvieron campos de lluvia diarias y algunos eventos de lluvia con corta duración en el tiempo. Como ejemplo se muestran los patrones de lluvia acumulada para el 21 de agosto del 2008 y un evento de lluvia importante sucedido el mismo día (ver Figura 4.10 y Figura 4.11).

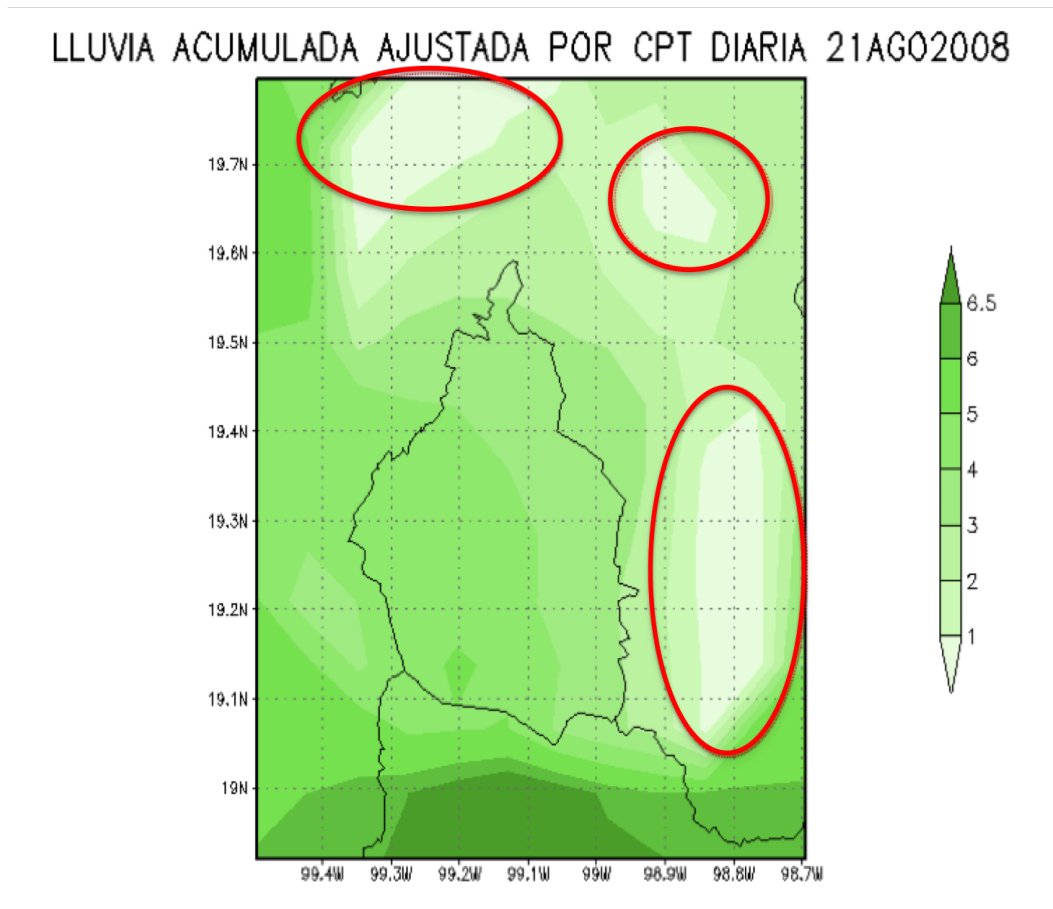


**Fig. 4.10:** Precipitación acumulada en 24hrs observada para el 21 de agosto del 2008



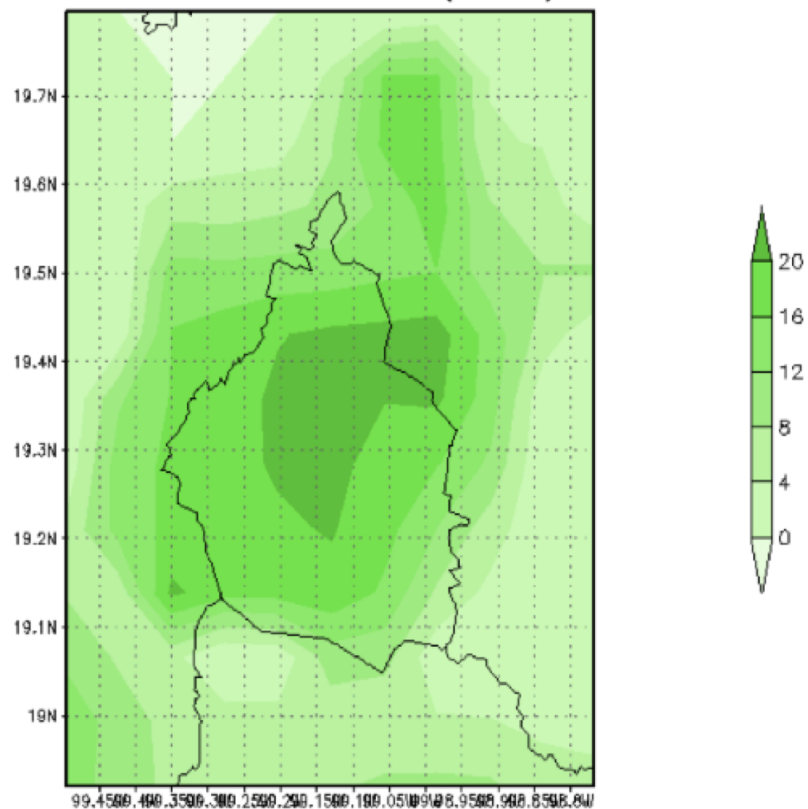


**Fig. 4.11:** Precipitación acumulada en 24hrs simulada por MM5 para el 21 de agosto del 2008

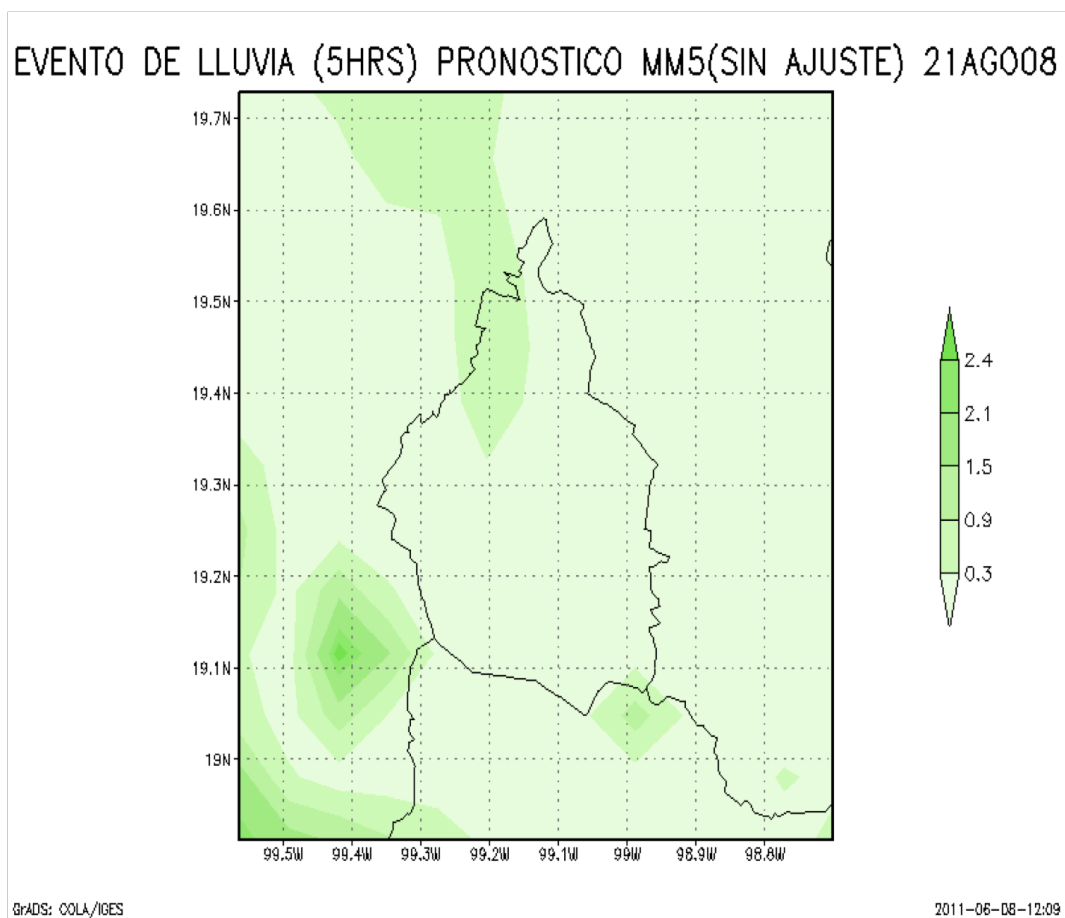


**Fig. 4.12:** Precipitación acumulada en 24hrs ajustada por CPT para el 21 de agosto del 2008 en rojo: zonas con precipitación escasa o nula. Las escalas de las figuras para la intensidad no son equivalentes salvo en los casos del patrón observado y el pronosticado por MM5 (Figuras 4.10 y 4.11). Se utilizó el mismo color, sin embargo, para mostrar la similitud entre los patrones espaciales observados, simulados y corregidos por CPT.

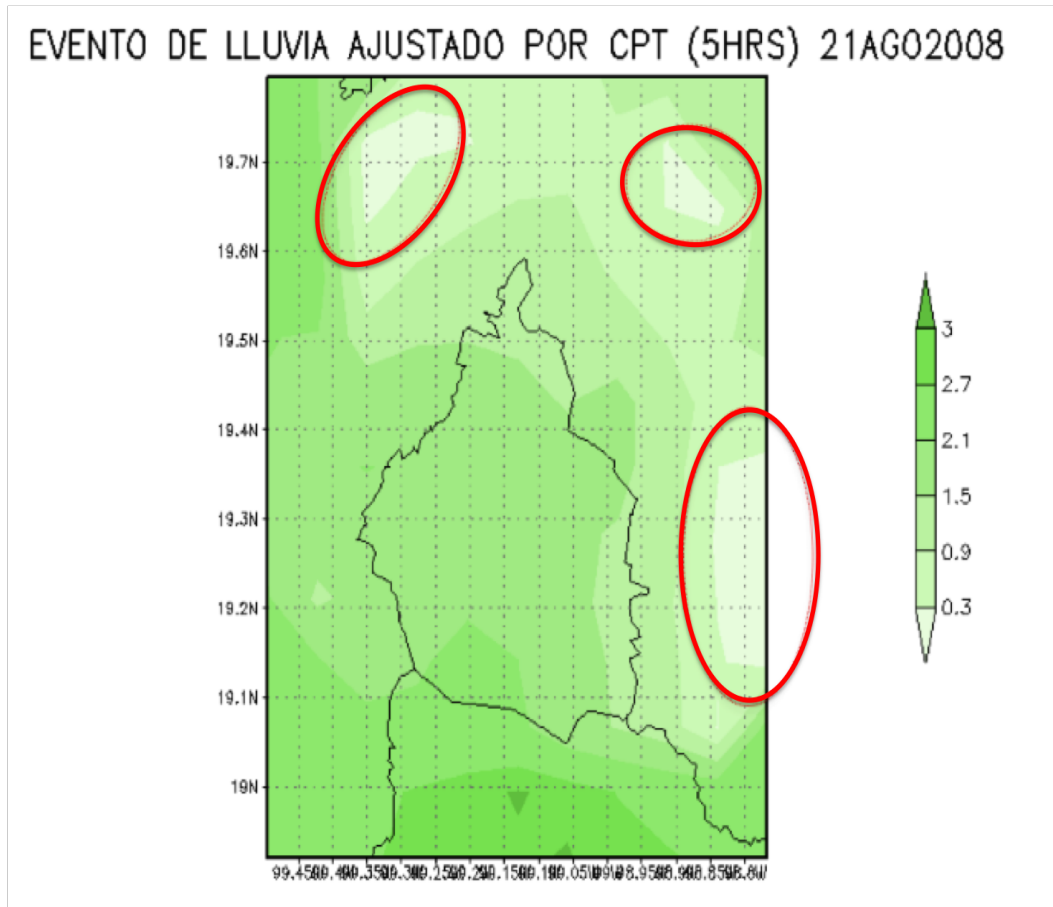
## EVENTO DE LLUVIA OBSERVADO (5HRS) 21AGO2008



**Fig. 4.13:** Evento de lluvia observada (5hrs de duración) del 21 de agosto del 2008



**Fig. 4.14:** Evento de lluvia simulada por MM5 (5hrs de duración) para el 21 de agosto del 2008



**Fig. 4.15:** Evento de lluvia (5hrs de duración) ajustada por CPT para el 21 de agosto del 2008. En rojo: zonas con precipitación escasa o nula. Las escalas de las figuras para la intensidad no son equivalentes salvo en los casos del patrón observado y el pronosticado por MM5 (Figuras 4.13 y 4.14). Se utilizó el mismo color, sin embargo, para mostrar la similitud entre los patrones espaciales observados, simulados y corregidos por CPT.

De acuerdo las figuras 4.10 y 4.11, los pronósticos de lluvia acumulada diaria realizados por el MM5 sin ajustarse por medio de un esquema MOS, si muestran la capacidad para predecir la ocurrencia de lluvia, pero no aciertan del todo en términos de determinar la distribución del patrón de lluvia, cuando se les compara con el campo observado. Sin embargo, cuando los errores

inherentes al MM5 son ajustados por el modelo estadístico de regresión, surge un patrón de lluvias que empieza a ser similar con el patrón observado (Fig. 4.10). En particular, el campo ajustado de lluvia diaria parece representar mejor al patrón observado, sobre todo en la zona que abarca al Distrito Federal, mientras que la calidad del patrón ajustado disminuye en zonas aledañas. De hecho, en zonas hacia el este y al norte, fuera de la ciudad, se observa en todos los registros zonas donde se predice ausencia casi total de lluvia ver (Figura 4.10: Lluvia acumulada diaria ajustada). Si bien existen registros donde si se predice lluvia en estas zonas, el resultado general son zonas siempre con escasa o nula precipitación, que constituyen evidentemente un error en el ajuste, debido en parte a la casi nula existencia de observaciones en esa zona, en donde las estaciones meteorológicas están espaciadas en promedio 25km (Ver Figura 3.1, Distribución de estaciones de la red de la SACM), y a la relativamente poca cantidad de datos históricos utilizados para calibrar el modelo estadístico (6 años de datos horarios), la predictibilidad sobre estas zonas disminuye.

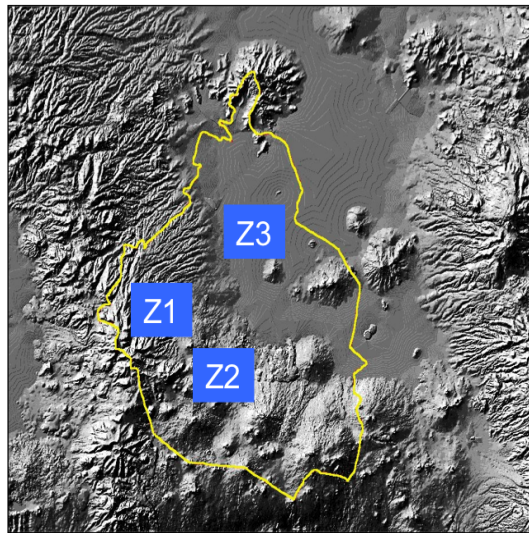
En cuanto a la capacidad de las salidas ajustadas para representar el patrón de lluvia de eventos cortos en el tiempo, se observa que los ajustes realizados a las salidas del MM5, son de gran ayuda para darse una idea de la distribución de la lluvia sobre la Ciudad de México en una escala de horas. Si bien los errores que hemos comentado también están presentes en estos campos, las salidas ajustadas mejoran considerablemente al representar eventos de lluvia generalizados. Sobre todo dentro de los límites de la ciudad. (ver Figura 4.11).

En la sección 4.05 se verificó al pronóstico cuantitativo de precipitación a través de una tabla de contingencia, sin embargo no es posible comparar el patrón espacial de lluvia debido a que en la sección 4.05 sólo contabilizamos retículas de una malla. Así que se realizó un análisis cualitativo para conocer el desempeño del pronóstico ajustado por el modelo estadístico, al pronosticar el patrón de lluvia en las mallas antes definidas, para tres diferentes zonas dentro del Distrito Federal. A través de definir un criterio, se comparó el patrón espacial de lluvia horaria observada contra el patrón espacial pronosticado, y los campos espaciales del MM5 para cada zona. El criterio de comparación debe considerar las semejanzas que tienen los patrones espaciales en las distintas mallas pero, especialmente, debe considerar la posición del máximo de intensidad de lluvia acumulada, debido a que esta podría ser

definitiva al momento de tomar una decisión en el ámbito de la protección civil por una emergencia hidrometeorológica.

#### 4.2.3. Precipitación local y criterio de comparación

En orden de evaluar la eficiencia del esquema MOS para corregir el pronóstico del patrón espacial de lluvia acumulada, se eligieron tres zonas dentro del Distrito Federal que fuesen representativas de la no homogeneidad de la lluvia en la ciudad (debida sobre todo a la orografía del terreno). De tal modo, se eligieron zonas donde hubiera cuerpos montañosos importantes, como el Sur y el Oeste de la ciudad, así como una zona al Norte, que representara solamente el uso de suelo e interacción urbana con los patrones de viento predominantemente desde el Este, sin presencia de montañas importantes. De este modo se eligieron zonas al norte, oeste y sur de la ciudad de México, siendo cada zona equivalente a un área de 16 x16km. La zona 1 se encuentra al oeste de la ciudad y se encuentra entre los 93.5° a 94.5° de longitud, y 19.1° a 19.25° de latitud. La zona 2 se encuentra al sur entre los 99.1 y los 99.25° de longitud y 19.05° a los 19.15° de latitud. La zona 3 al norte entre los 99.1° a los 99.25° en longitud y 19.2° a 19.30° de latitud (ver Figura 4.12).



**Fig. 4.16:** Tres zonas de estudio utilizadas para establecer comparaciones del patrón espacial de lluvia

---

Se dividió cada malla horaria, tanto observada, pronosticada y simulada, en tres subareas de 16km x 16km (equivalentes a un 2 x 2 puntos de malla cada una) y se comparó la distribución de lluvia en cada subárea con la correspondiente en la malla de lluvia observada y en la modelada; otorgando un punto a cada subárea si el patrón espacial fuese *cualitativamente* igual que el de la subárea de la malla observada en cada subárea . (máximo puntaje: 4 puntos).

El criterio de comparación fue el siguiente:

Para considerar que la malla pronosticada o simulada por CPT es un “acierto” se consideró que:

- La posición del máximo de precipitación pronosticada en malla fuese exactamente la misma que el máximo de precipitación en el patrón observado.
- Que el patrón pronosticado tuviera al menos 3 puntos de acuerdo al criterio anterior comparando con el patrón observado para la misma hora.

Para considerarlo como “falso”:

- Que el patrón pronosticado tuviese menos de 3 puntos con el criterio utilizado, independientemente de la posición del máximo de intensidad en la malla observada.

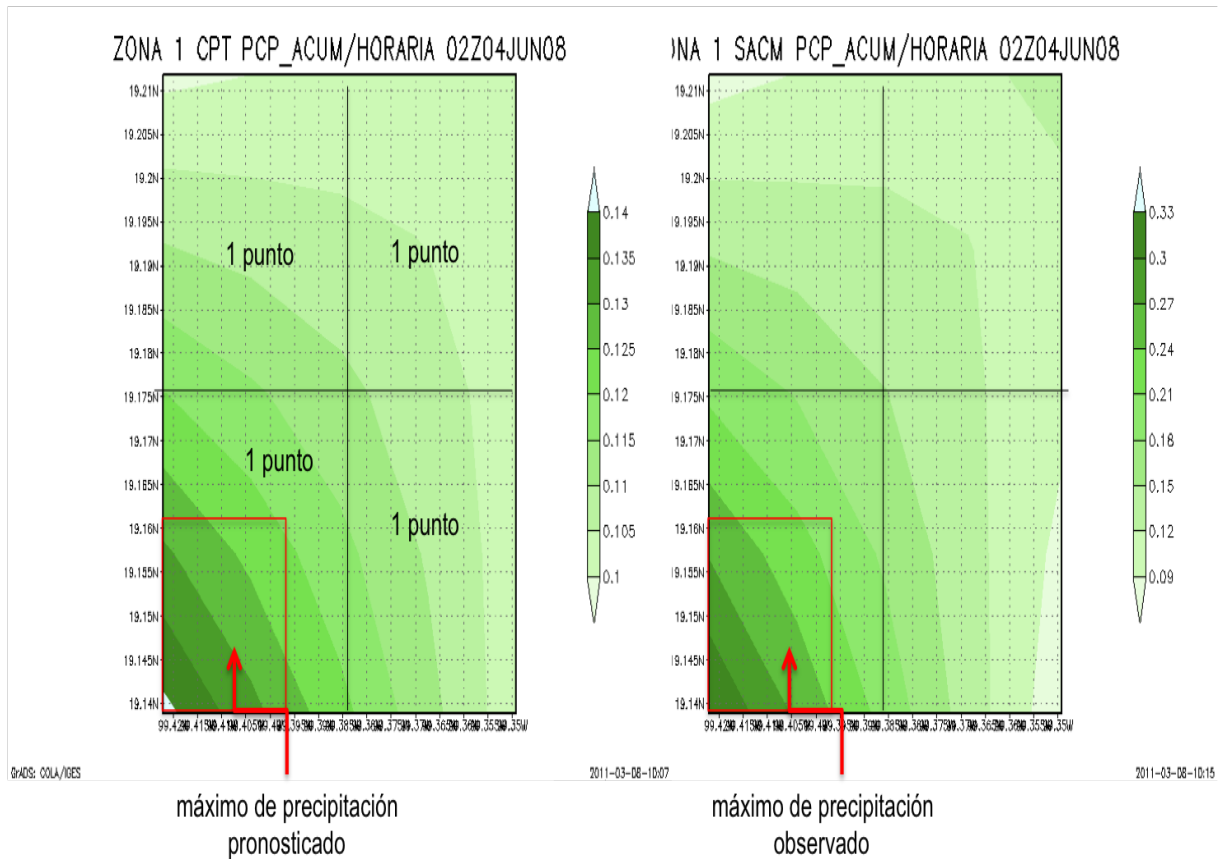
Como ejemplos se muestran los campos de lluvia horarios, así como la ejecución del criterio de comparación, para el: 4 de julio de 2008 (con calificación de: acierto), para el 22 de julio del 2008 (con calificación de: acierto, y para el 29 de agosto del 2008 (con calificación de: no acierto).



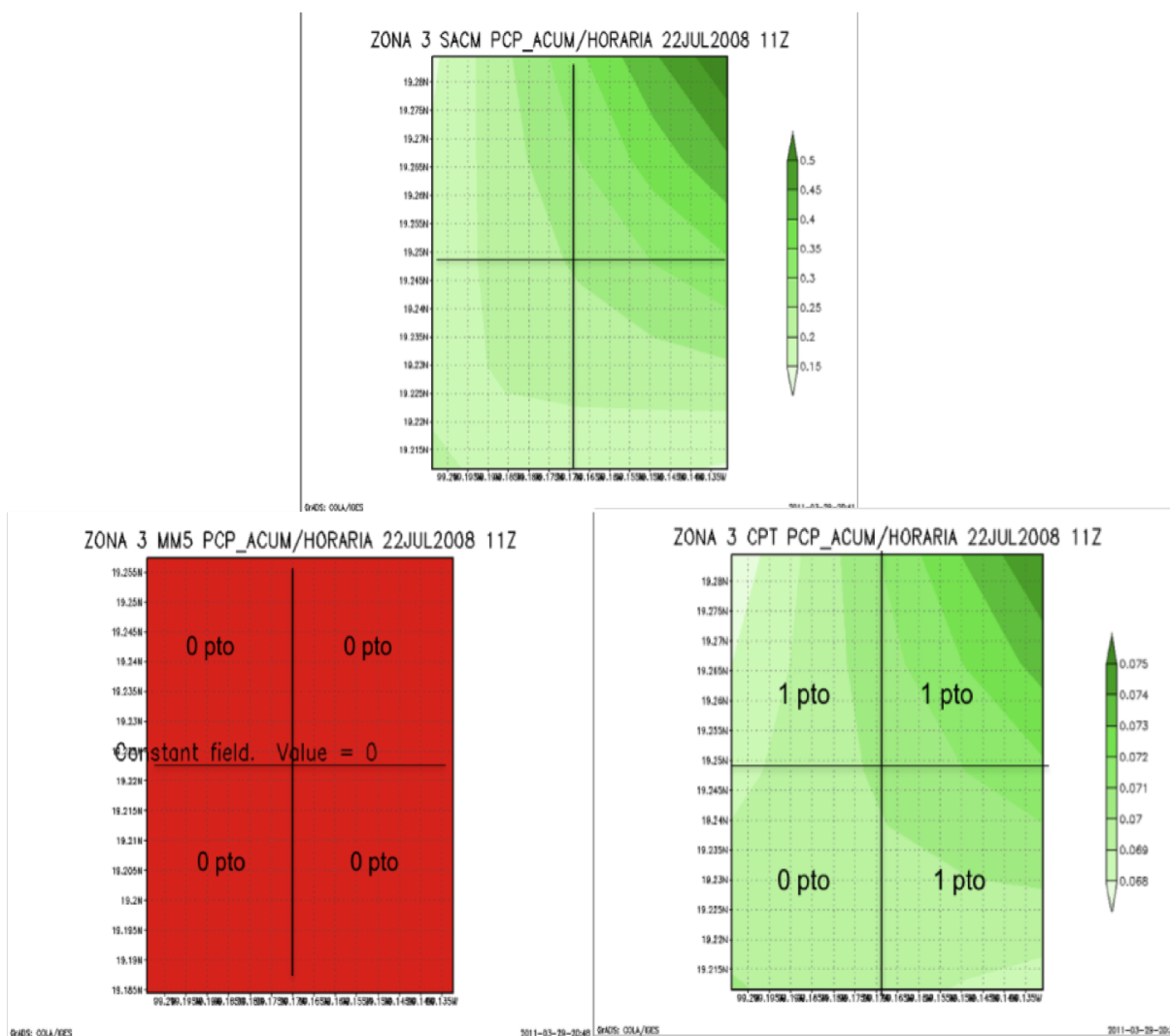
## Ejemplo de pronóstico espacial “acierto”

Pronosticado

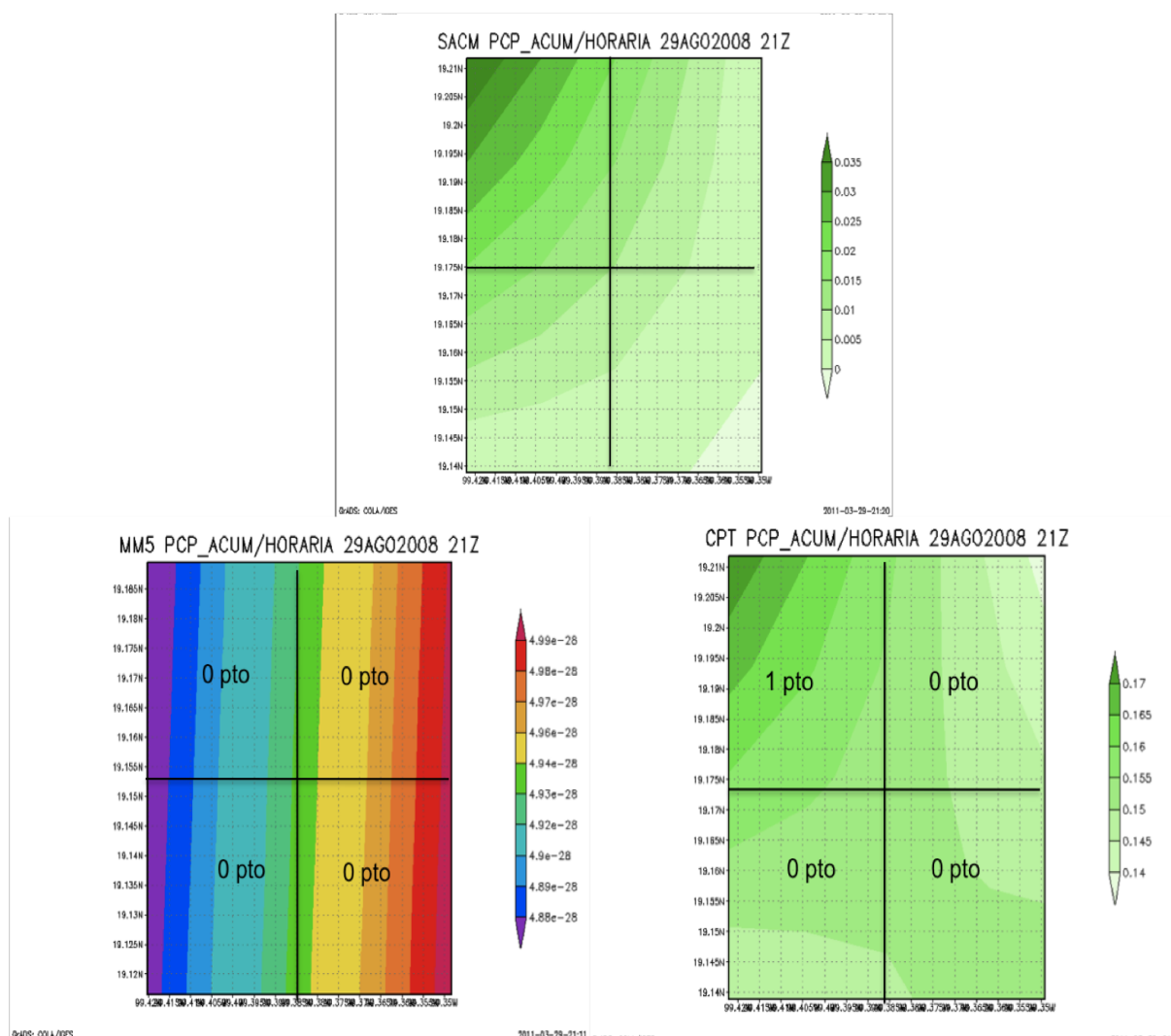
Observado



**Fig. 4.17:** Lluvia acumulada para el 4 de junio del 2008 a las 02Z para la zona 1( Oeste) y el criterio utilizado para darle la calificación de acierto al pronóstico de CPT y falso a la simulación de MM5. Izquierda: campo ajustado por el modelo estadístico, derecha: campo observado



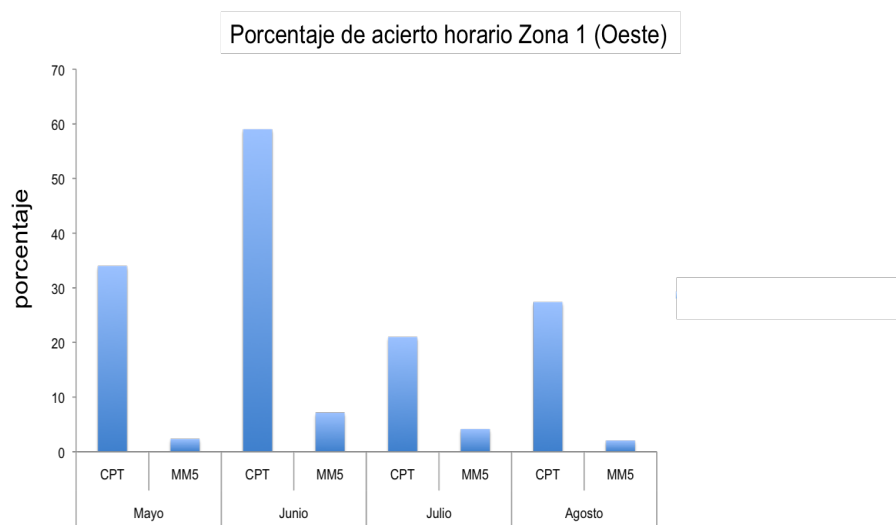
**Fig. 4.18:** LLuvia acumulada para el 22 de julio del 2008 a las 11Z para la zona 3 y el criterio utilizado para darle la calificación de acierto al pronóstico del CPT y falso a la simulación de MM5. Arriba: campo observado, abajo izq: simulación de MM5 sin ajuste, abajo der: campo ajustado por el modelo estadístico



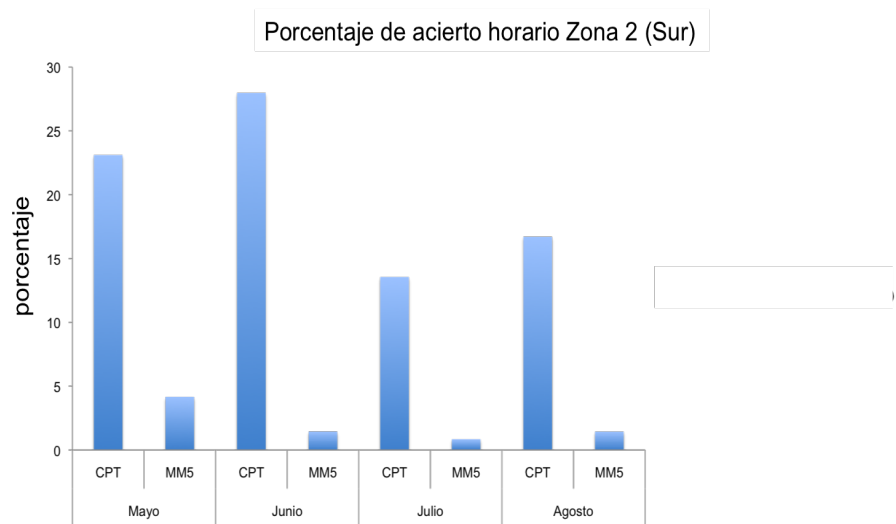
**Fig. 4.19:** Lluvia acumulada para el 29 de agosto del 2008 a las 21Z para la zona 1 y el criterio utilizado para darle la calificación de falso al pronóstico de CPT y falso a la simulación de MM5. Arriba: campo observado, abajo izq: simulación de MM5 sin ajuste, abajo der: campo ajustado por el modelo estadístico

Se generaron series de tiempo donde se muestra el porcentaje de aciertos para cada zona y para cada mes, al comparar los patrones observados, simulados y pronosticados utilizando el criterio anterior (ver figuras: 3.14, 3.15,

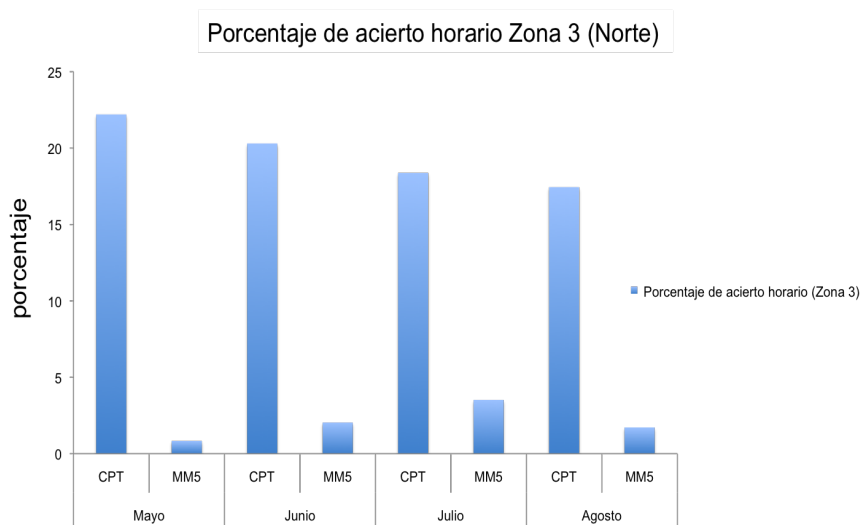
3.16).



**Fig. 4.20:** Porcentaje de aciertos horarios para la zona 1 en el año de pronóstico (2008)



**Fig. 4.21:** Porcentaje de aciertos horarios para la zona 2 en el año de pronóstico (2008)



**Fig. 4.22:** *Porcentaje de aciertos horarios para la zona 3 en el año de pronóstico (2008)*

El ánimo de realizar el análisis del patrón de lluvia es el de estimar la eficiencia de CPT para corregir los errores de las salidas del MM5 y así pronosticar patrones de lluvia acumulada en una malla de alta resolución temporal y horaria, Esto se logra comparando los pronósticos con las salidas del MM5, los pronósticos corregidos y los observados a la misma resolución utilizando el criterio definido con anterioridad. Definir distintas áreas, implica necesariamente observar la respuesta que tiene CPT cuando cambia la orografía al momento de pronosticar patrones de lluvia.

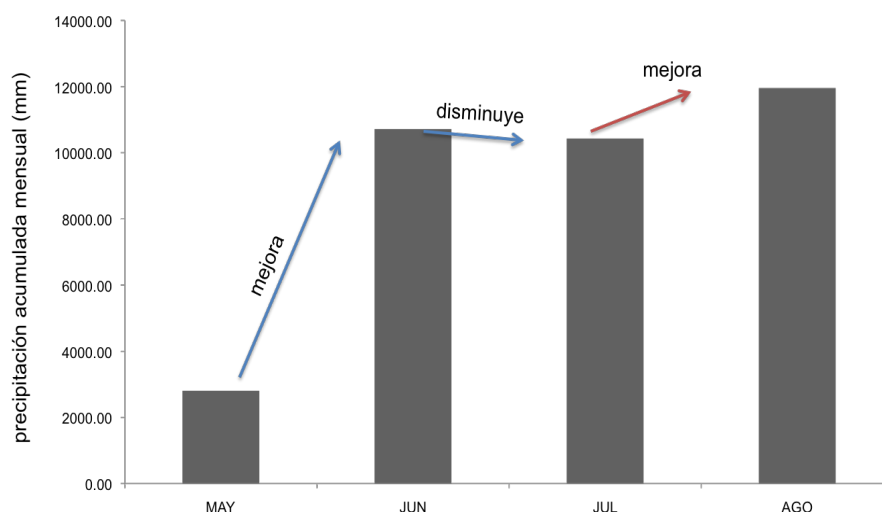
De acuerdo a las figuras anteriores, en la zona 1, al Oeste de la ciudad, el mes que presenta más porcentaje de acierto es junio (aproximadamente de 59%, mientras que el que presenta el menor índice es julio (22%). Para la zona 2, junio es de nueva cuenta, el mes que presenta mayor número de aciertos, y el porcentaje general disminuye considerablemente con respecto de la zona 1 (para junio es de 28%).

En la zona 1 y la zona 2, el comportamiento de los porcentajes es muy similar. Es decir, un porcentaje de aciertos para el mes de junio y un cambio proporcional para los meses de julio y agosto. Mientras que para la zona 3, todos los meses presentan un porcentaje de aciertos parecido, siendo mayo el mes con mayor puntaje (27%).

---

En un primer análisis, parece ser que CPT es sensible a la frecuencia de lluvia observada. Este comportamiento en el porcentaje de aciertos, se ve reforzado precisamente en las zonas montañosas del Distrito Federal en donde se presenta la mayor cantidad de lluvia en el Distrito Federal en el verano (zona 1 y 2). Se observa un particular comportamiento de la lluvia caracterizado por el registro de dos periodos de lluvia máxima en junio y septiembre. En contraste, en julio y agosto disminuye el nivel de lluvias. De acuerdo a esta interpretación, para los periodos de tiempo con mas frecuencia de lluvia, CPT tendría la capacidad de predecir los patrones más eficientemente. Si esta idea es correcta, entonces el comportamiento de los porcentajes de acierto, en función de la frecuencia de la lluvia de verano, debería reflejar un aumento en el periodo de mayo a junio, una disminución de junio a julio, y prácticamente el mismo comportamiento para el intervalo julio-agosto, esperando que entre agosto y septiembre, el indice de aciertos aumente otra vez. De acuerdo a las gráficas anteriores, en el periodo de mayo a junio, el índice de porcentaje aumenta, mientras que de junio a julio disminuye, sin embargo, de julio a agosto el porcentaje de aciertos vuelve a aumentar (aumentando aproximadamente dos puntos porcentuales en la efectividad de aciertos).

Para averiguar porqué, se generó una serie de tiempo de la lluvia acumulada para los 4 meses de estudio, a partir de las mediciones de toda la red de estaciones de la SACM para estudiar el comportamiento de la lluvia acumulada mensual, y se observó si el comportamiento mencionado era consistente con la presencia o no de lluvia, al considerar el aumento o disminución de la lluvia acumulada(ver figura 4.19)



**Fig. 4.23:** Comportamiento del porcentaje de aciertos horarios para la tres zonas en el año de pronóstico (2008) y la lluvia acumulada mensual

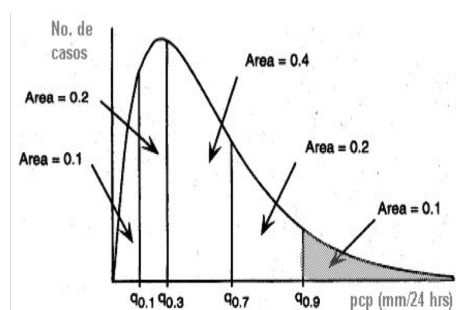
Según la gráfica anterior, la causa de que en el periodo julio-agosto el índice de aciertos aumentase, se debió a que en el año 2008, la lluvia acumulada fue mayor para el mes de agosto.

En el caso del modelo de mesoescala MM5, la efectividad que tiene para predecir los patrones de lluvia es muy baja. Los porcentajes de aciertos, para cada zona y para cada mes, nunca superan los 2%, lo anterior debido a que MM5 funciona con parametrizaciones, que están basadas en las ecuaciones que modelan al comportamiento del clima en una escala mayor, al aumentar la resolución espacial, estas parametrizaciones empiezan a dejar de tener validez. De acuerdo a trabajos anteriores, en el desempeño del MM5 para la simulación de lluvia sobre el Distrito Federal, la confianza en la simulación de lluvia acumulada disminuye conforme la intensidad de las lluvias aumenta. Por el contrario en el CPT, esta situación (en el caso de la predicción de patrones de lluvia) es precisamente la contraria. Aunque CPT presenta, de acuerdo a lo anterior, capacidad de poder pronosticar campos de lluvia a tan alta resolución; solamente un limitado número de casos presentan una predicción casi perfecta, o sea, el patrón observado es casi idéntico al patrón pronosticado. Sin embargo, nos parece importante mencionar que, si en el criterio de comparación, no consideráramos la similitud del patrón de lluvia, y

solamente consideráramos la posición del máximo de intensidad, el porcentaje de aciertos aumentaría, en ciertas zonas, hasta un 70 %.

#### 4.2.4. Análisis de eventos extremos

Según Wilks (1995), un evento extremo puede ser definido como el valor más grande o más pequeño de una variable atmosférica entre un número dado de observaciones. La probabilidad de tener cierta cantidad de precipitación en 24 hrs puede ser obtenida usando una distribución sesgada hacia la izquierda tal como la función de distribución de probabilidad del tipo Gamma. Dividiendo la distribuciones en percentiles, los valores de precipitación por arriba de cierto umbral se les puede asociar una probabilidad de entre el 10 % y el 5 % de ocurrencia (Magaña, Pérez, 2002)

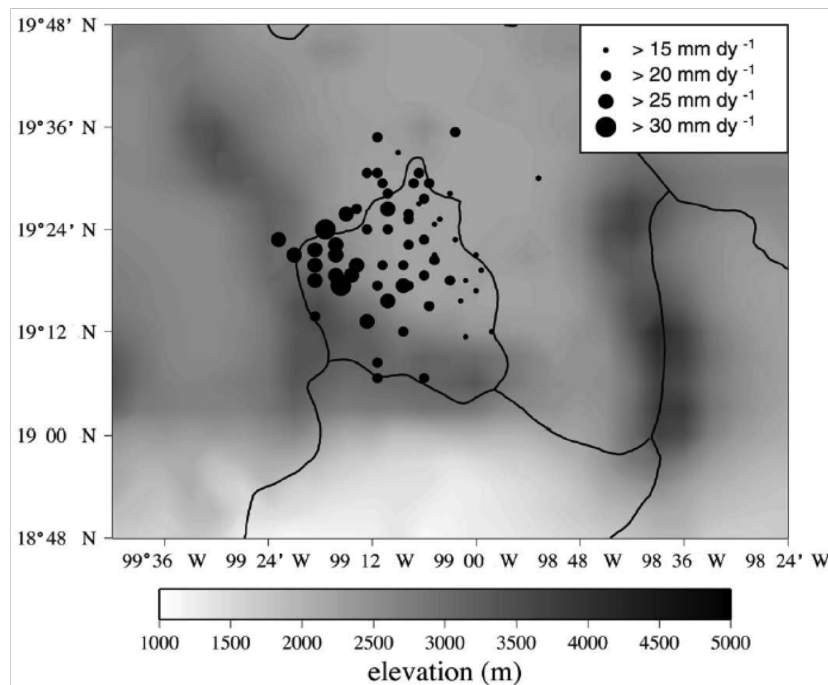


**Fig. 4.24:** Distribución Gamma para lluvia acumulada en 24 hrs. La región sombreada ( $q$ ) corresponde al 10 % de probabilidad y denota un evento de lluvia extrema.

Es decir: un evento de precipitación extrema se define como aquél que es suficientemente improbable de ocurrir y que cuando lo hace, siempre rebasa un cierto umbral de precipitación. Magaña y Méndez (2003) analizaron la función de distribución de probabilidad de la lluvia en 24 hrs registrada en la red pluviométrica de la SACM, mediante una función de distribución Gamma (FDG), y consideraron como un evento extremo a la lluvia en ese rango de tiempo, que estuviera en el percentil correspondiente al área con 10 % de probabilidad (en el extremo derecho de una FDG). Es decir, se considera que el percentil que ocupe el área con 90 % de la probabilidad representa casos de lluvia normal, y sólo el 10 % restante corresponde a eventos de precipitación extrema.



El valor de este umbral estadístico no puede ser homogéneo para un área como la de la Ciudad de México, ya que tiene que reflejar la cantidad de precipitación que es normal en alguna zona en particular. De este modo, el umbral será mayor para zonas donde sean comunes precipitaciones intensas en 24 horas, mientras que en áreas con una precipitación mas débil, el valor del umbral será relativamente bajo. De esta manera, asociar umbrales a datos de precipitación diaria permite tener una idea de la variabilidad espacial de la precipitación extrema. De acuerdo Magaña y Méndez (2003), para que ocurra un evento de precipitación extrema en la parte oeste y suroeste de la ciudad, la precipitación en 24 horas debe exceder un umbral de 25-30mm, mientras que en la parte este de la ciudad, más de 15mm en 24 horas ya constituyen un evento extremo. El cambio sustancial de valores umbrales se debe a la componente orográfica de la precipitación, i.e, a la interacción de las montañas con el flujo del este que ocasiona mucha mas precipitación en el este de la ciudad que en el oeste.

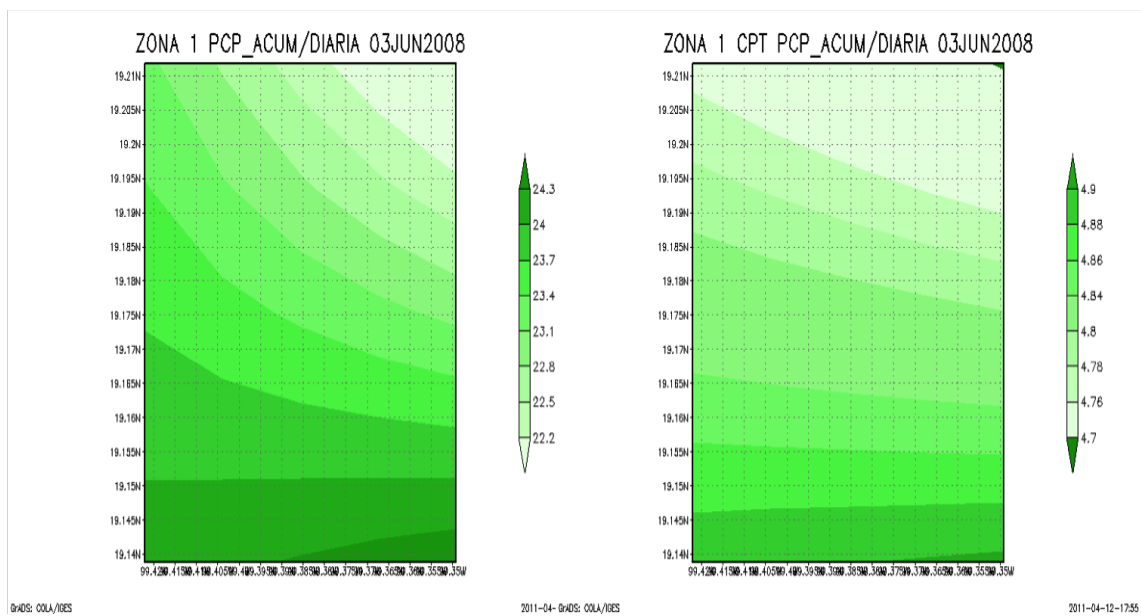


**Fig. 4.25:** Valores umbrales para determinar eventos de lluvia extrema, la topografía se expresa en una escala de grises, (tomado de: *Diagnosis and prognosis...*, Magaña et al, 2003)

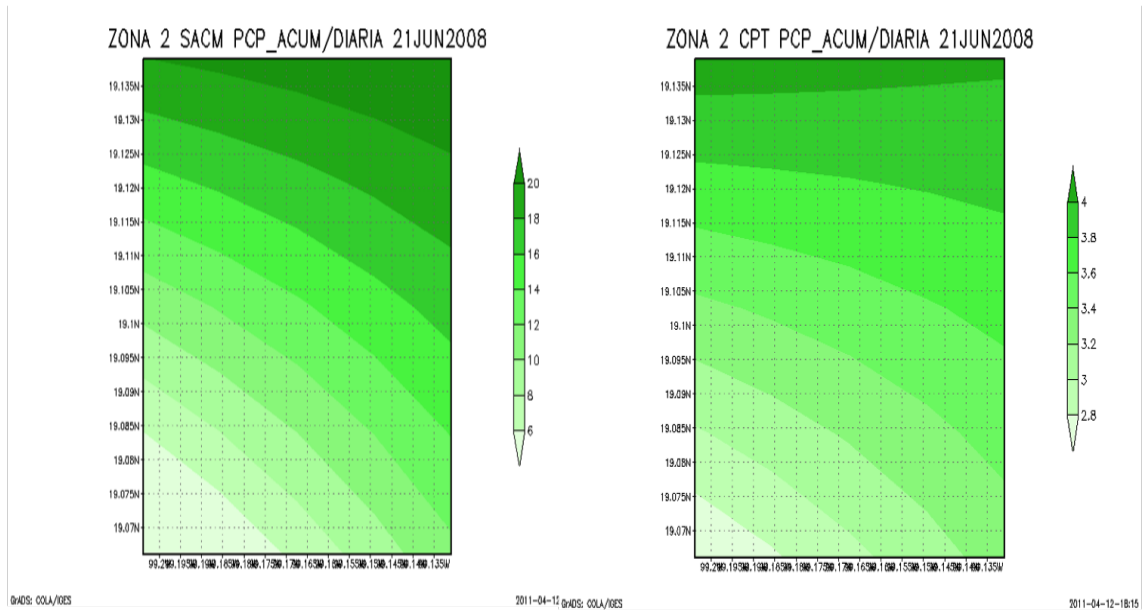
Para las 3 zonas de estudio definidas anteriormente (1,2,3, ver subsección 3.2.2, Criterio de Comparación) y a partir de la base de datos asimilada horaria se construyeron campos de lluvia acumulada diaria en una malla de 16 x16km y se eligieron aquellos que superaron los umbrales zonales definidos por Magaña et al. De acuerdo a esto, los umbrales para cada zona son aproximadamente:

	UMBRAL
Zona 1	>25mm/día
Zona 2	>20mm/día
Zona 3	>15mm/día

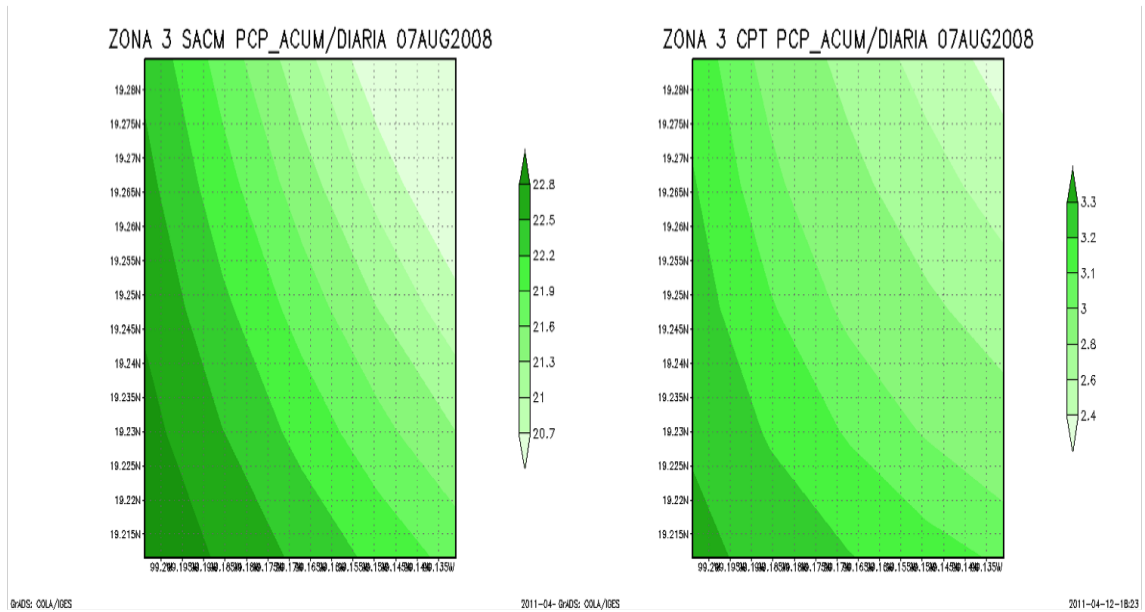
Con los pronósticos de CPT, a la misma resolución espacial, se construyeron campos de lluvia acumulada diaria para las mismas zonas, y después se les comparó con los campos de precipitación *diaria* de la base de datos asimilada que hubieran rebasado algún umbral. De tal manera que se comparó el patrón espacial de un evento extremo, y se analizó el campo pronosticado por CPT para la misma fecha. Para ejemplificar lo anterior se muestran varios casos de campos de lluvia extremos:



**Fig. 4.26:** LLuvia acumulada para el 3 de junio del 2008 en la zona 1 (evento intenso). A la izquierda el campo observado, y a la derecha el campo pronosticado en CPT.



**Fig. 4.27:** LLuvia acumulada para el 21 de junio del 2008 en la zona 2 (evento intenso). A la izquierda el campo observado, y a la derecha el campo pronosticado en CPT.

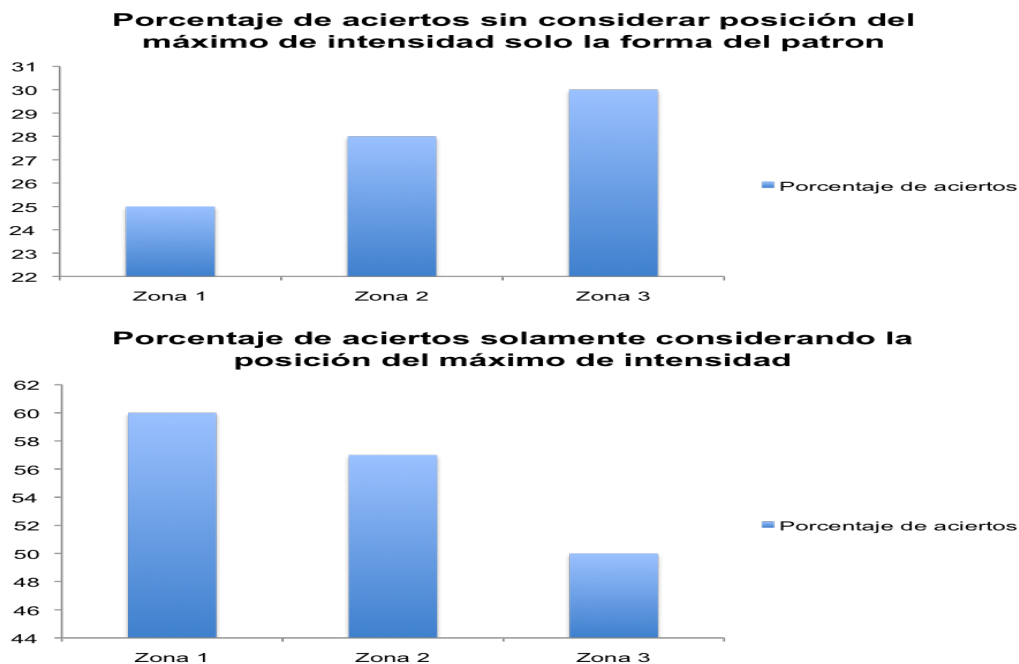


**Fig. 4.28:** Lluvia acumulada para el 7 de agosto del 2008 en la zona 3 (evento intenso). A la izquierda el campo observado, y a la derecha el campo pronosticado en CPT.

Siguiendo el mismo criterio de comparación planteado en la sección 4.2.2 se otorgó la calificación de acierto o no acierto a los campos diarios de precipitación acumulada para cada zona ya definida. En base a esto se obtuvieron porcentajes de aciertos en cuanto ubicación, porcentajes que se muestran en la figura 4.25. Se consideraron dos casos en el criterio de comparación, el primero que corresponde a los casos en los que el patrón de lluvia dentro del área elegida obtuvo un puntaje de 3 (que el patrón corregido fuera similar a su par observado) y el segundo que corresponde a los casos en los que el patrón de lluvia completo no necesariamente mostrase una alta similitud con el patrón observado, sino que se reflejara la posición exacta del máximo de intensidad. Según la figura 4.23, en el primer caso, la representación fiel del patrón de lluvia de un evento extremo se da en la zona de la ciudad donde las lluvias son mas moderadas (zona 3, al centro-norte de la ciudad), siendo menos la capacidad de ajustarse a zonas con mayor cantidad de lluvia diaria. Cuando el criterio de comparación considera solamente la posición del máximo de intensidad de lluvia, este comportamiento se invierte, pudiéndose alcanzar porcentajes de aciertos de hasta 60 % en zonas donde regularmente

llueve mas, debido a la acción de las montañas (zona 1 y 2, oeste y suroeste respectivamente). Esto tiene relación con lo que ya habíamos discutido antes, esto es: las salidas del MM5 ajustadas por el modelo estadístico, tienen la capacidad de pronosticar mejor la lluvia en las épocas donde más llueve, contrario al desempeño que se conocía del modelo MM5 (Pérez 2007).

Es necesario llamar la atención al siguiente hecho: en el patrón espacial de un evento extremo pronosticado por MM5 y ajustado por el modelo estadístico, comparando con los campos observados, el máximo de intensidad siempre está orientado hacia algún lado en la malla, nunca está propagado por toda el área ni ocupa el centro de la misma. Además, siempre tienen una simetría con respecto del centro del campo: un *ojo* de máxima intensidad y un simétrico pero de intensidad mínima.



**Fig. 4.29:** Porcentaje de aciertos para el patrón de eventos extremos en las tres zonas definidas; considerando la forma del patrón completo en cada zona (arriba) y considerando únicamente la posición del máximo de intensidad (abajo)

## 5. CONCLUSIONES

### *5.0.5. Desempeño del Pronóstico Cuantitativo de Precipitación*

El modelo de mesoescala MM5, aunque es una buena herramienta para pronóstico, presenta varios errores sistemáticos en cuanto a precipitación en el Valle de México, entre los que podemos mencionar: que la distribución de la magnitud del error en la predicción de lluvia sobre el área de la Ciudad de México (cuando se compara con un campo observado), es relativamente menor al inicio y al final de la temporada de lluvias, con un aumento a la mitad de la temporada debido a la ocurrencia de eventos extremos de lluvia. De tal modo que la confianza en la simulación de lluvia acumulada disminuye conforme la intensidad de las lluvias aumenta, y hay zonas en donde, aún con asimilación de datos, el resultado no es satisfactorio (Pérez, 2004). En el caso de simulaciones puntuales, y para meses poco lluviosos (mayo, junio), el MM5 llega a tener hasta un 77% de confianza en el pronóstico de lluvia acumulada a 24 horas <sup>1</sup> pero cuando se trata de meses considerados húmedos, como julio y septiembre, la confianza disminuye a menos del 50% en general, principalmente en las zonas poniente, suroeste y sur donde la precipitación es mucho mayor que en otros sitios de la ciudad por el efecto de la orografía. En el caso de la predicción de eventos extremos, resulta visible la incapacidad del MM5 para reproducir la cantidad de lluvia (aún con un margen de error). Sin embargo la aproximación del patrón de lluvia simulado respecto del observado sí da una idea de la distribución del patrón de lluvia en toda la ciudad (más aún cuando el evento de lluvia no es forzado por la componente orográfica). (Pérez 2004).

Se analizan las posibles ganancias de la predicción corregida al utilizar MOS, que podemos dividir en varias categorías:

---

<sup>1</sup> Lo anterior puede reflejar la eficacia del MM5 para simulaciones de lluvias sin presencia de eventos intensos

### 5.0.6. Area de lluvia, desempeño del QPF

-El ajuste que hace el modelo estadístico utilizando las salidas del MM5 corregidas es confiable, ya que siempre se encuentra que el número de casos de aciertos siempre es mayor al número de casos de falsas alarmas.

-Los resultados del QPF indican que, por un lado, se mejora la capacidad de detección de eventos de lluvia para los meses más lluviosos (junio, julio y agosto) aunque en el mes más seco (mayo) se presentan más falsas alarmas y el sesgo crece, por lo que la capacidad de detección disminuye. Cuando no se utiliza un esquema MOS en las salidas del MM5, la magnitud del error es mayor para los meses más húmedos. Al utilizar un esquema MOS, esta dependencia se invierte, siendo mejor el comportamiento del pronóstico precisamente para los meses con mayor cantidad de lluvia.

### 5.0.7. Patrón de lluvia, similitud entre patrones

El MM5 tiene la capacidad para predecir con cierta confiabilidad el patrón de distribución espacial de la lluvia acumulada a 24 para un área como el Valle de México (Pérez, 2004), sin embargo, la eficiencia se reduce cuando se hace presente la componente orográfica, reduciendo la así la confianza en el pronóstico. Utilizando CPT, se mejora esta capacidad y se logra predecir correctamente la distribución espacial en un rango de entre 30% y 40% para eventos de lluvia individuales horarios. Recordemos que este resultado se obtuvo a través de criterio de comparación (entre simulados y observados), completamente subjetivo. En cierto aspecto, la importancia en el diseño y la aplicación de este criterio obedeció a encontrar una similitud del patrón de lluvia sobre toda el área definida, sin importar la posición del máximo de intensidad (ver sección ... Análisis de la distribución del patrón de lluvias). Si el criterio hubiera sido, precisamente: considerar a la posición de este máximo como *acierto*, el índice de aciertos aumentaría dramáticamente, acercándose al 66% de acierto para algunos meses. Cabe destacar que, en un porcentaje significativo, los patrones de lluvia pronosticados/corregidos si son prácticamente iguales a los observados.

#### *Patrón de lluvia de eventos extremos*

Con los campos de lluvia horarios corregidos, es posible adecuar la definición de evento extremo en 24 horas, al utilizar los campos de precipitación diaria



pronosticados por el MM5, de tal modo que se observa una semejanza significativa entre el campo de un evento extremo pronosticado/corregido (definido bajo el criterio anterior) con su semejante observado. Cuando se comparan estos campos diarios, la semejanza cualitativa es significativamente alta si se considera solamente la posición del máximo de intensidad en ambos campos. Uno de los problemas que tiene el MM5 en cuanto a precipitación, es precisamente que el desempeño del pronóstico falla cuando se presentan eventos de lluvia intensos, de hecho, la calidad del pronóstico se ve disminuida desde que comienza la temporada de lluvias (Pérez 2007). Con la corrección hecha por el modelo estadístico, este comportamiento se modifica. No obstante, el índice de aciertos en zonas donde llueve mucho, disminuye si se compara el patrón completo dentro de una malla de 16x16km. Ahora, es necesario desarrollar un esquema para categorizar a los eventos extremos en una escala horaria, generando umbrales zonales al estimar la función de distribución de los datos horarios y sus intensidades asociadas. Que es posible debido a que se posee una base de datos de lluvia a escala horaria, una herramienta para regionalizar y ajustar las salidas de un modelo numérico, y una técnica para verificar este producto.

#### *Intensidad*

Debido a la falta de algoritmos para calcular intensidades en la herramienta CPT, y al relativamente poco tiempo utilizado para calibrar el modelo estadístico, siempre se pronostica una lluvia ligera, que posiblemente influye en la magnitud de la intensidad (pero no en la distribución espacial de la lluvia acumulada) tanto horaria como diaria. Es importante señalar, que aunque esta precipitación era constante en los campos ajustados por el modelo estadístico, para los meses mas secos el índice rondaba los  $10^{-14}$ mm/hr, que en realidad habla de eventos de no lluvia. Sin embargo es necesario utilizar una mayor climatología horaria para realizar la calibración del modelo estadístico, ya que esto mejora las ecuaciones de regresión, al eficientar los EOF's entre predicandos y predicadores.

#### *5.0.8. Trabajo a Futuro*

-Desarrollar nuevos métodos para eficientar el número de EOF's utilizados para calibrar las ecuaciones de regresión, ya que no existe ninguno. Esto podría lograrse utilizando a CPT para corregir los errores de las bases de

---

datos de lluvia antes de ser asimiladas; ya sea para corregir estructuras de tormentas, o corrigiendo la magnitud de un evento extremo medido en estación, encontrando primero que tipo de análisis se debería ejecutar en cada caso (Análisis de componentes principales o Análisis de Correlación Canónica, PCR o CCA por sus siglas en inglés.

-Aprovechar al máximo las últimas versiones de la herramienta de cómputo CPT que construye el modelo estadístico, ya que éstas permiten utilizar una mayor cantidad de predictores en el cálculo de las funciones de transferencia, por lo que podrían incluirse datos de viento y temperatura, además de los de precipitación, para calibrar el modelo estadístico.

-Adaptación y aplicación de una definición de eventos extremos horarios, quizás utilizando una función de probabilidad sesgada, para generar una primera distribución geográfica de umbrales de precipitación.

-Generar una base de datos horaria con al menos 20 años de registros horarios, para verificar que la habilidad del MM5, una vez corregido, para pronosticar la intensidad de lluvia acumulada (ya sea en días o en horas) mejore.

-Diseñar un esquema de Alerta Temprana por eventos de precipitación intensos que sea amable hacia el usuario final de los productos de pronóstico y que complementen las demás herramientas con las que cuenta un tomador de decisión.

## 6. BIBLIOGRAFÍA

Abramovitz, J.N. 1999. *Unnatural Disastres*. World Watch Magazine 12: 30-35.

Aguilar, E. Peterson, P. Ramirez Obando, R. Frutos, J.A. Retana, M. Solera, I. González Santos, R.M. Araujo, A. Rosa Santos, V.E. Valle, M. Brunet India, L. Aguilar, L. Álvarez, M. Bautista, C. Castañón, L. Herrera, E. Ruano, J.J. Siani, F. Obed, G.I. Hernández Oviedo, J.E. Salgado, J.L. Vázquez, M. Baca, M. Gutierrez, C. Centella, J. Espinosa, D. Martinez, B. Olmedo, C.E. Ojeda Espinoza, M. Haylock, R. Núñez, H. 2005. *Changes in precipitation and temperature extremes in Central América and Northern South America*. J Geoph. Res. (110), 2005.

AGROASEMEX, S.A. 2006. *La experiencia mexicana en el desarrollo y operación de seguros paramétricos aplicados a la agricultura*, México. <http://www.agroasemex.gob.mx>

Boissonnade, A; Heitkemper Lawrence, J. y Whitehead David. 2002. *Wheather Data, Cleaning and Henhancement In. Risks Books*. Risk Management Solutions; Earth Satellite Corporation Londres, Inglaterra.

Bushong J.B. 2005. *Quantitative Precipitation Forecast: Its Generation and Verification at the Southeast River Forecast Center*. Proceedings of the 1999 Georgia Water Resources Conference, held March 30-31, 1999, at the University of Georgia. Kathryn J. Hatcher, Editor, Institute of Ecology, The University of Georgia, Athens, Georgia

Brunet-India, M y Lópe Portillo, D. 2001. *Detecting and Modeling Regional Climate Chance*. Springer Berlin-Hidelberg-New York. 2001

CENAPRED. 2001 *Diagnóstico de peligros e identificación de riesgos de*

---

*desastres en México: Atlas Nacional de riesgos de desastres en México.* Centro Nacional de Prevención de Desastres, Secretaría de Gobernación, México. 2005

Cressman, G. P. 1959. *An operational objective analysis system.* Monthly Weather Review, (87), 367-374.

Charney, J.G. Fjortoft, R. y Von Newman, J. 1950. *Numerical integration of the barotropic vorticity equation,* Tellus, (2), 237-254.

Charba, J. P. y Samplatsky, G. 2011. *High-Resolution GFS-Based MOS Quantitative Precipitation Forecasts on a 4-km Grid.* Monthly Weather Review. (139). 39-68.

Calderón, A.F. 2003. *Comentarios sobre la elección de una plataforma para la ejecución del modelo de simulación del tiempo a escala regional MM5 del NCAR.* Simbiosis Ltda. 1-10. Colombia.

Caetano, E. y Magaña, V. 2007. *Pronóstico climático estacional regionalizado para la República Mexicana como elemento para la reducción de riesgo, para la identificación de opciones de adaptación al cambio climático y para la alimentación del sistema: cambio climático por estado y por sector.* Dirección de Investigación sobre Cambio Climático. SEMARNAT. México.

DesInventar 6. 2003. *La Red Sistema de Inventarios y Análisis de Desastres.* México.

Daley, R., 1991 *Atmospheric Data Analysis.* Cambridge University Press, Inglaterra.

Dawit A. Z y Mekonnen G. 2009. *Evaluation of CMORPH Precipitation Products at Fine Space-Time Scales.* Civil and Environmental Engineering Department, University of Connecticut, Storrs, Connecticut. Journal of Hydrometeorology. (10). 300-307.

Dennis, J. Steven, J. Llana, R. Timothy, J. 1994. *An Introduction to Atmospheric and Oceanographic Data.* NCAR Technical Note. Climate and Global Dynamic Division. NCAR. Boulder Colorado. EU.

---

EIRD.2007 *Estrategia Internacional para la prevención de Desastres*. Naciones Unidas. <http://www.eird.org/index-esp.html>

Ebert, E. Brown, B G. Bally, Brooks, J. H E. Casati, B. Jaeneke M, Nurmi, P. Stephenson, D. Wilson, L. 2004. *Forecast Verification Activities for the Sydney 2000 Forecast Demostration Project*. American Meteorological Society. 73–96.

Ebert, E.E. McBride, J.L. 2000. *Verification of precipitation in weather systems: determination of systematic errors*. Journal of Hydrology (239), 179–202.

Giroto, P.O. 2002. *Scaling up. Resilience to hazards and the importance of cross-scale linkages*. UNDP Expert group Meeting on Integrating Disaster Reduction and adaptation to Climate Change, La Habana, Cuba, 17-19 de junio.

Gay, C. 1999. *México: una visión hacia el siglo XXI*. El cambio climático en México. Instituto Nacional de Ecología / Universidad Nacional Autónoma de México / U.S. Country Studies Program. Centro de Ciencias de la Atmósfera, UNAM, México.

Gilchrist, B., y Cressman, G. 1954. *An experiment in objective analysis*. Tellus, (6), pp. 309-318.

Glahn, H. R., y Lowry, D. A. 1972. *The use of Model Output Statistics (MOS) in objective weather forecasting*. J. Appl. Meteor. (11). 1202-1211.

García, S. H. 2003. *Homogeneización e Interpolación de Observaciones de Alta Resolución*. Proyección Regional de Escenarios de Cambio Climático. Técnicas y Proyectos en Curso. Instituto Nacional de Meteorología. Universidad de Cantabria. Santander, España.

Idowu, O.S., Rautenbach, C.J., 2008, *Model Output Statistics to improve severe storms prediction over Western Sahel*. Atmos. Res. (93), 419-425.

Secretaría del Medio Ambiente *Informe Climatológico Ambiental del Va-*

---

lle de México. 2005. Cap. 5. Gobierno del Distrito Federal. México.

INEGI, 2005. *Estadísticas del Medio Ambiente*. Instituto Nacional de Estadística, Geografía e Informática. [www.inegi.gov.mx/est/](http://www.inegi.gov.mx/est/)

Jáuregui, E. (2000). *El Clima de la Ciudad de México*, Instituto de Geografía de la UNAM y Plaza y Valdéz Editores. México

Jáuregui, E., y E. Romales. 1996. *The urban effects on convective precipitation in Mexico City*. Atmospheric Environment, (30), 3383-3389.

Kirtman, B. P y Shukla, J. Balmaseda, M. Graham, N. Penland, C. Xue, Y. Zebiak, S. 2000. *Current Status of ENSO Forecast Skill*. A Report to the Climate Variability and Predictability (CLIVAR) Numerical Experimentation Group (NEG). WCRP Informal Report No 23/01, ICPO Publication No. 56

PNUD. 2004 *La reducción de riesgos de desastres: un desafío para el Desarrollo*. Dirección de prevención de Crisis y de Recuperación. Programa de las Naciones Unidas para el desarrollo, Nueva York. EU. disponible en: [www.undap.org/cpr/disread.rdr.html](http://www.undap.org/cpr/disread.rdr.html)

Lixin, L. Revesz, P. 2004. *Interpolation methods for spatio-temporal geographic data*. Computers, Environment and Urban Systems. (28) 201-227.

Magaña, V. y Neri, C. 2007. *Eventos Hidrometeorológicos Extremos en el Valle de México*. Revista Ciudades. (74). 26-30.

Magaña, V, Pérez, J, Méndez M. 2003, *Diagnosis and Prognosis of extreme precipitation events in the Mexico Basin*. Geofísica Internacional, (42), UNAM, México.

Magaña, V. 1999. *Los Impactos del Niño en México*. Dirección General de Protección Civil. Secretaría de Gobernación. México.

Magaña V. y Pérez J. 1998. *Usos de un modelo de mesoescala en el estudio de la dinámica atmosférica regional de México*. Revista GEOUNAM. (5). Num 1. 33-39.

---

Magaña, V. 2010. *Guía para generar y aplicar escenarios probabilísticos regionales de cambio climático en la toma de decisiones*. Centro de Ciencias de la Atmósfera. UNAM. México.

Martínez, R., Mascarenhas, A., Alvarado, A. 2009. *Guía Técnica para la Implementación de un Sistema Regional de Información Aplicada a la Gestión de Riesgo Agrícola en los Países Andinos*. Centro Internacional para la Investigación del Fenómeno de El Niño –CIIFEN, 1-160.

Mason. S. y Baddour, O. 2007. *Statistical Modeling*. Chapter 7. International Research Institute for Climate and Society, International Research Institute for Climate and Society. EU.

Manson, H. Jardel, E. *et al.* 2009. *Perturbaciones y desastres naturales: impactos sobre las ecoregiones, la biodiversidad y el bienestar económico*. Capital Natural de México. Vol.2. Estados de Conservación y tendencias de cambio. CONABIO. México. 132-184.

Nicholls, N. 1987. *The Use of Canonical Correlations to Study Teleconnections*. Monthly Weather Review. (115), 393-399.

Pérez, J. 2004. *Pronóstico Numérico del tiempo para el Valle de México*. Tesis de Maestría. UNAM, México.

Panofsky, H. 1949. *Objective weather map análisis*. Journal Applied of Meteorology, (6), 386-392.

Richman, Michael B. and Peter J. Lamb. 1985. *Climatic Pattern Analysis of Three and Seven-Day Summer Rainfall in the Central United States: Some Methodological Considerations and Regionalization*. Journal of Climate and Applied Meteorology (24). Num. 12. 1325-1343.

Rivano Gutiérrez, F. 2004. *Análisis de eventos extremos de precipitación y su efecto en el diseño de drenaje superficial de tierras agrícolas del sur de Chile*. Tesis de Lic. Universidad Austral de Chile. Fac. de Ciencias Agrarias. Chile.

Robert J. J. John E. J. Pingping Xie, y Phillip A. A. 2004. NOAA/NWS/NCEP.

---

Climate Prediction Center (CPC), Camp Springs, MD, USA 2 Earth System Science Interdisciplinary Center, Univ. of Maryland, College Park, MD, EU.

Rossa, A. Nurmi, P. Ebert, E. 2007. *Overview of methods for the verification of quantitative precipitation forecasts*. Cap. 16. Weather and Environmental Prediction. The Centre for Australian Weather and Climate Research.

SACM, 2010. Subdirección de Macromedicación y Control de Redes. Sistema de Aguas de la Ciudad de México. Gobierno del Distrito Federal. México

Saucier, W. 1955. *Principles of Meteorological Analysis*. University of Chicago Press. EU.

Silva, V. B. S. Kousky V. E. Shi, W. y Higgins, R. W. 2007. *An Improved Gridded Historical Daily Precipitation Analysis for Brazil*. Journal of Hydro-meteorology. (8). 841-860.

Schneider, T. 2001. *Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values*. Journal of Climate (14), 853-871.

Servicio Meteorológico Nacional. 2000. ERIC II, (Extractor rápido de información climatológica) Base de datos meteorológico. Instituto Mexicano de Tecnología del Agua (IMTA, 1996) México.

Stern, W. Sirutis, J. Ploshay, J. y Miyakoda, K. 1986. *Recent Development in the GFDL Extended-Range Forecasting System, Short and Medium Range Numerical Weather Prediction* (Special Volume of the Journal of the Meteorological Society of Japan), Editado por T. Matsuno, 359- 363.

Shlens, J. 2009. *A Tutorial on Principal Component Analysis*. Center for Neural Science, New York University. New York City, NY. EU.

Vázquez-Aguirre, J. 2006. *Datos climáticos de la República Mexicana: panorama actual y requerimientos inmediatos*. 1er Foro del Medio Ambiente Atmosférico en el Estado de Xalapa, Veracruz. México

Wilks, D. 1995. *Statistical Methods in Atmospheric Sciences*. Chap. 7:



---

Forecast Verification, Academic Press, 233-283. EU.

Warner, T. y Anthes, R. 1978. *Numerical simulations with a three-dimensional mesoscale model*. Monthly Weather Review, (106), 1079–1099. EU.

Wilby, R. L. y Wigley, T. M. L. 2000 *Precipitation predictors for downscaling observed and general circulation model relationships*. International Journal Of Climatology. (20), 641-661

Yong-Fu, Q. y Zhong, Z. 1987. *General Forms of Dynamic Equations for Atmosphere in Numerical Models with Topography, Short and Medium Range Numerical Weather Prediction* (Special Volume of the Journal of the Meteorological Society of Japan), Editado por T. Matsuno. 743-756.

Whittow J, 1980. *Disasters. The anatomy of environmental Hazards*. University of Georgia Press. Atenas, Grecia.

WMO Programa Mundial de Datos y Vigilancia del Clima. 2007. Directrices Sobre la Gestión de Datos Climáticos, Organización Meteorológica Mundial. Ginebra, Suiza.

Zermeño, D. D. M. 2008. *Análisis probabilístico de escenarios escalados de precipitación y temperatura bajo cambio climático en México*. Tesis de Maestría en Ciencias. UNAM 83pp.

Sitios de Internet.

<http://www.iges.org/grads/>

<http://ccr.aos.wisc.edu/model/processing/>

<http://www.ecmwf.int/products/forecasts/guide/>

<http://fluidos.eia.edu.co/hidrologiai/probabilidad/probabilidad.htm>

<http://www.rmmag.com/>

<http://www.eird.org/esp/revista/no162009/art15.html>

## 7. INDICE DE FIGURAS

## 8. ABREVIATURAS

Do it again.