



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA  
INGENIERÍA CIVIL - HIDRÁULICA

LA ASIMILACIÓN DE DATOS COMO PROCEDIMIENTO PARA LA ESTIMACIÓN DE LLUVIA ESPACIALMENTE  
DISTRIBUIDA, Y SU APROVECHAMIENTO EN LA MODELACIÓN DE ESCURRIMIENTOS SUPERFICIALES

**TESIS**

QUE PARA OPTAR POR EL GRADO DE:  
DOCTORA EN INGENIERÍA

PRESENTA:

**MARIA ASUCENA RODRIGUEZ RAMIREZ**

TUTOR PRINCIPAL

DR. ÓSCAR ARTURO FUENTES MARILES, INSTITUTO DE INGENIERÍA - UNAM

MIEMBROS DEL COMITÉ TUTOR

DR. RAFAEL BERNARDO CARMONA PAREDES, INSTITUTO DE INGENIERÍA - UNAM

DR. RAMÓN DOMÍNGUEZ MORA, INSTITUTO DE INGENIERÍA - UNAM

DR. EDGAR GERARDO MENDOZA BALDWIN, INSTITUTO DE INGENIERÍA - UNAM

DRA. MARITZA LILIANA ARGANIS JUÁREZ, INSTITUTO DE INGENIERÍA - UNAM

CIUDAD UNIVERSITARIA, CIUDAD DE MÉXICO, MARZO 2025



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**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.

**INTEGRANTES DEL JURADO:**

Presidente: Dr. Carmona Paredes Rafael Bernardo

Secretario: Dr. Domínguez Mora Ramón

1er. Vocal: Dr. Fuentes Mariles Óscar Arturo

2do. Vocal: Dr. Mendoza Baldwin Edgar Gerardo

3er. Vocal: Dra. Arganis Juárez Maritza Liliana

Lugar donde se realizó la tesis:

Posgrado de Ingeniería UNAM, Ciudad Universitaria, Ciudad de México.

**TUTOR DE TESIS:**

Dr. Óscar Arturo Fuentes Mariles



-----  
**FIRMA**

### ***Agradecimientos:***

*A la Universidad Nacional Autónoma de México, espacio excepcional para el desarrollo integral de nuestras capacidades.*

*A mi asesor, Dr. Óscar Arturo Fuentes Mariles, por su confianza, apoyo y enseñanzas transmitidas durante el desarrollo de nuestro proyecto.*

*A los miembros de mi comité, Doctora Maritza Liliana Arganis Juárez y Doctores Ramón Domínguez Mora, Rafael Bernardo Carmona Paredes y Edgar Gerardo Mendoza Baldwin. Mi gratitud hacia ustedes, por su notable participación en mi formación académica.*

*Un agradecimiento especial al Dr. Iván Vilchis Mata, por toda la información y conocimientos invaluable que tuvo a bien compartirme, sobre la obtención de lluvias estimadas a partir de imágenes de reflectividad de radares meteorológicos.*

*Al M.I. Adam Hernández Hernández, por su asesoría en el manejo del modelo hidrodinámico para análisis de escurrimientos superficiales aplicado. Muchas gracias.*

*A quienes intervinieron en diversas etapas de este proyecto, especialmente: Geógrafa Claudia Elena del Valle Domínguez, M.I. Germán Adrián Aguilar Martínez, Dr. Gerardo Acuña Soto y Dr. Rojacques Mompremier. Les agradezco su importante contribución.*

*Al Consejo Nacional de Ciencia y Tecnología (CONACYT), ahora Secretaría de Ciencia, Humanidades, Tecnología e Innovación, por el apoyo otorgado en mis estudios de posgrado, registro con CVU 554857.*

***Dedicatoria:***

*A mi amado esposo, José Alberto Sánchez Cruz,  
porque gracias a tu compañía, paciencia y cariño,  
juntos logramos cerrar este ciclo en nuestras vidas.  
Gracias por ser la luz y guía en cada instante de mi camino.*

# Índice General

Resumen.....	3
Abstract.....	5
Introducción .....	7
Objetivos .....	9
<b>1. Comportamiento de la precipitación en México.....</b>	<b>10</b>
1.1. El fenómeno de la precipitación .....	10
1.2. La precipitación sobre la República Mexicana.....	13
1.2.1. Características de la precipitación.....	13
1.2.2. Registros históricos de precipitación y el análisis de su comportamiento espacial .....	19
1.2.3. Precipitación estimada con percepción remota.....	26
<b>2. Conceptos y descripción de métodos para la asimilación de datos .....</b>	<b>36</b>
2.1. Aspectos generales.....	36
2.1.1. Esquema de Panofsky.....	40
2.1.2. Esquema de Gilchrist y Cressman .....	40
2.2. Asimilación de datos para sistemas dinámicos no lineales .....	41
2.2.1. Formulación básica de mínimos cuadrados para modelos perfectos.....	41
2.2.2. Mejor estimación lineal de mínimos cuadrados.....	43
2.3. Esquema de asimilación de datos secuencial .....	46
2.3.1. Esquemas de correcciones sucesivas.....	47
2.3.2. Interpolación óptima o estadística .....	52
2.3.3. La interpolación secuencial: Filtros de Kalman .....	55
2.4. Esquema de asimilación de datos variacional .....	57
2.5. La aplicación de modelos para el análisis de campos meteorológicos.....	59
2.5.1. La combinación de lluvias registradas en superficie y estimadas por satélite... 59	

2.5.2. La técnica de suavizado Kernel para la combinación de datos de lluvia .....	60
2.5.3. La combinación de lluvias registradas en superficie y estimadas por radar .....	61
2.5.4. La cuantificación de lluvia con datos de radar, satélite y registros en superficie .....	62
2.6. Aspectos de interés para la selección de un procedimiento de combinación y su empleo.....	63
<b>3. Programación computacional para los procesos de asimilación .....</b>	<b>66</b>
3.1. Aspectos generales de la modelación.....	66
3.2. La estimación de lluvia mediante un modelo de Asimilación de datos .....	71
3.3. Programación del método de Barnes .....	72
3.3.1. Primera etapa.....	73
3.3.2. Segunda etapa.....	73
<b>4. Aplicación del modelo de asimilación de lluvia sobre diferentes regiones del territorio nacional.....</b>	<b>78</b>
4.1. Modelo de asimilación aplicado.....	78
4.1.1. Entradas del modelo .....	78
4.1.2. Calibración del modelo y evaluación de resultados.....	80
4.2. Aplicación del modelo .....	81
4.2.1. Análisis de una región con disponibilidad de pluviómetros alta.....	81
4.2.2. Análisis de regiones con disponibilidad de información escasa .....	94
<b>5. Consideración de lluvia distribuida espacialmente como insumo de un modelo para la simulación de escurrimiento superficial.....</b>	<b>120</b>
5.1. La modelación de escurrimientos en la revisión de la calidad de estimaciones de lluvia distribuida espacialmente.....	120
5.2. Modelación de escurrimientos superficiales.....	120
5.2.1. Escurrimientos superficiales en la subcuenca de la E.H. Teapa.....	124
5.2.2. Escurrimientos superficiales en la subcuenca de la E.H. Cadereyta.....	132
5.2.3. Escurrimientos superficiales en la subcuenca de la E.H. Las Perlas .....	137
<b>6. Conclusiones generales .....</b>	<b>143</b>
Referencias .....	151
Anexo 1. Metodología propuesta para Asimilación de tres bases de datos de lluvia.....	158
Anexo 2. Programa para Asimilación de datos.....	161
Anexo 3. Metodología para estimación de lluvia con imágenes de radar .....	169
Anexo 4. Tablas.....	179

# Resumen

La precipitación y el escurrimiento, como parte del ciclo hidrológico, son fenómenos naturales fundamentales para el desarrollo de la vida en la Tierra, por lo que el análisis del comportamiento espacial y temporal del agua de lluvia es importante para una gestión más eficiente de este recurso, objetivo que históricamente se ha perseguido implementando criterios técnicos para el diseño y operación de infraestructura para el desarrollo social, como es el suministro de agua potable, el riego de cultivos, la generación de energía y el manejo de eventos extremos para controlar avenidas o mitigar sequías, entre otros.

La observación tradicional de la precipitación en México se ha realizado mediante redes terrestres de pluviómetros, y aunque en general los registros diarios que emiten se consideran confiables, su representatividad está limitada a la región próxima a los sitios donde se ubican, con el inconveniente que representa la variabilidad de las tormentas, razón por la cual la disponibilidad de estos valores suele ser insuficiente para ciertos fines. Cabe enfatizar la necesidad de contar con mayor número de datos de eventos de precipitación sobre el territorio nacional, de manera que un conocimiento más integral de la distribución de las tormentas permita mejorar los resultados de diversos estudios que requieren la distribución de la lluvia como insumo prioritario.

Dado que existen antecedentes del empleo de productos de percepción remota para el registro de la distribución de lluvia, en este trabajo se revisó la posibilidad de obtener mejores estimaciones de la distribución espacial de la lluvia diaria. Suponiendo que las mediciones de lluvia con redes de pluviómetros son confiables, y que las estimaciones por percepción remota describen el comportamiento espacial y temporal de los eventos, se optó por utilizar un modelo de asimilación de datos secuencial hasta la etapa de diagnóstico (esquema de Barnes) para obtener una distribución bidimensional de los campos de lluvia en superficie por la combinación de información pluviométrica con los productos de sistemas satelitales y radares meteorológicos, dando especial atención al beneficio que sobre las estimaciones tiene la densidad de las redes de observación.

La calidad del procedimiento de asimilación de las tres bases de datos se revisó con el análisis de cinco eventos de precipitación ocurridos en la Ciudad de México, los que se supusieron representativos al involucrar láminas de acumulación puntual diaria ligeramente superiores a las lluvias medias máximas anuales registradas en la red de pluviómetros de la región. Con un tamaño de celda que coincide con la resolución de los datos satelitales, la evaluación de campos de lluvia sobre una malla regular se realizó con criterios estadísticos en un esquema de validación cruzada, obteniendo resultados satisfactorios de la combinación de las tres bases respecto de las estimaciones obtenidas con cada base de manera independiente.

Como complemento, se realizó el análisis de los mismos eventos en la Ciudad de México, tomando en cuenta una menor cantidad de registros, al considerar conjuntos de pluviómetros de diferentes densidades, con distribución espacial de observaciones relativamente uniforme. Se observó que con la combinación de las tres bases de datos pueden obtenerse aproximaciones adecuadas, independientemente de la configuración de una red, para densidades alta o media, aunque la calidad de las estimaciones se reduce si la densidad es baja, con errores comparables a los obtenidos con la interpolación IDW de los registros pluviométricos. Adicionalmente, para un análisis en condiciones más desfavorables, se resolvieron ejemplos donde se utilizó una sola base de datos de percepción remota en la asimilación, circunstancia de interés especial debido a la disponibilidad limitada de registros de radar meteorológico en México. Se observó que al combinar

registros pluviométricos con datos de satélite o de radar se obtienen errores con un orden de magnitud semejante, y son estimaciones útiles respecto a interpolaciones IDW calculadas con los registros de una red de pluviómetros de densidad media.

Por haber obtenido resultados satisfactorios en la estimación de lluvias diarias distribuidas sobre la Ciudad de México, se aplicó la metodología en otras regiones del país, donde la disponibilidad de registros pluviométricos por unidad de superficie es sensiblemente menor. En vista de que con la asimilación de las tres bases de datos recomendadas se consiguieron las mejores estimaciones de lluvia, y la integración de datos pluviométricos con los de origen satelital resultaron ser buenas aproximaciones, se optó por este tipo de combinación, en el entendido de las dificultades que implican la recopilación por escasez y el manejo de registros de radar. Se identificaron tres regiones de interés por su tamaño, disponibilidad de información hidrométrica y pluviométrica con distribución no uniforme de registros, localización geográfica que influye en la magnitud de láminas de lluvia precipitadas en varios días consecutivos, así como su cercanía con centros urbanos de mayor dimensión. Los sitios se localizan dentro de los estados de Chiapas-Tabasco, Coahuila-Nuevo León y Oaxaca-Veracruz.

De los resultados se concluyó que los mayores errores de estimación de las lluvias asimiladas están relacionados con la falta de exactitud de las lluvias calculadas como productos de satélite, lo que es más recurrente en los campos de lluvia de magnitud reducida. Sin embargo, en caso de que los datos de satélite guarden un orden de magnitud similar a los registros pluviométricos disponibles, la aproximación de la lluvia por asimilación es más apropiada.

El análisis de precipitación en cada una de las regiones estudiadas comenzó con la identificación de las subcuencas de captación, definiéndose como puntos de salida del escurrimiento la ubicación de estaciones hidrométricas con registros de gasto de las fechas analizadas; la revisión de la calidad de la lluvia distribuida estimada se complementó al considerarla como insumo para el análisis de escurrimientos superficiales por lluvia en cuenca propia. En conocimiento de antecedentes satisfactorios del empleo del modelo matemático Iber para el cálculo del hidrograma a la salida de una cuenca [45], los resultados obtenidos con dicho modelo permitieron verificar la calidad de las lluvias estimadas por asimilación respecto a las obtenidas por interpolación IDW de los registros pluviométricos considerados en la asimilación. Asimismo, con la intención de aportar mayor confiabilidad a las estimaciones, la variación de los gastos a la salida de las subcuencas también se obtuvo mediante un procedimiento hidrológico de parámetros concentrados.

En los resultados del análisis de las subcuencas de menor tamaño, se aprecia que existe similitud entre magnitudes y variaciones de gastos obtenidos para un mismo esquema de lluvia analizado con un modelo hidrodinámico o hidrológico, y se considera que las diferencias principales respecto a los registros históricos de gastos son consecuencia de la sobreestimación de las precipitaciones. Los caudales tuvieron una mejor aproximación al utilizar las lluvias estimadas con el método de asimilación propuesto. En atención a los resultados obtenidos para la subcuenca de mayor tamaño, también se concluye que las diferencias entre el registro hidrométrico y los valores obtenidos por simulación de los procesos de lluvia-escurrimiento se deben especialmente a la sobreestimación de las precipitaciones; sin embargo, el resultado que representó una mejor aproximación se obtuvo al emplear las lluvias estimadas con el método de interpolación IDW, dado el tamaño del área analizada y la cantidad reducida de estaciones pluviométricas, localizadas principalmente en el perímetro de la subcuenca de captación, condiciones adversas que provocaron una falta de aproximación en las estimaciones de lluvia sobre un mayor número de celdas, al utilizarse el procedimiento de asimilación propuesto.

# Abstract

Precipitation and runoff, as part of the hydrological cycle, are natural phenomena fundamental to the development of life on Earth, so the analysis of the spatial and temporal behavior of rainfall is important for a more efficient management of this resource. This objective has been pursued through the implementation of technical criteria for the design and operation of infrastructure for social development, such as drinking water supply, crop irrigation, power generation and management of extreme events to control surface runoff or mitigate droughts, among others.

Traditional precipitation observation in Mexico has been carried out by means of terrestrial rain gauge networks, and although the daily records they provide are generally considered reliable, their representativeness is limited to the region near the sites where they are located, with the disadvantage of variability in storms, which is why the availability of these values is often insufficient for certain purposes. It is important to emphasize the need for more data on precipitation events over the national territory, so that a more comprehensive understanding of storm distribution will improve the results of various studies requiring rainfall distribution as a priority input.

Given the history of using remote sensing products for recording rainfall distribution, this work reviewed the possibility of obtaining better estimates of the spatial distribution of daily rainfall. Assuming that rainfall measurements with rain gauge networks are reliable, and that remote sensing estimates describe the spatial and temporal behavior of events, a sequential data assimilation model was chosen to be used up to the diagnostic stage (Barnes scheme) to obtain a two-dimensional distribution of the surface rainfall fields by combining rain gauge information with satellite and weather radar products considering the effects on estimates given the density of the rain gauge network used.

The quality of the assimilation procedure for the three databases was reviewed with the analysis of five precipitation events in Mexico City, which were assumed to be representative by involving some records of daily rainfall slightly higher than the average annual maximum rainfall recorded in the region's rain gauge network. With a cell size that matches the resolution of the satellite data the evaluation of rainfall fields on a regular grid was performed using statistical criteria in a cross validation scheme obtaining satisfactory results from the combination of the three databases in respect of the estimates obtained for each database independently.

As a complement, the analysis of the same events in Mexico City was carried out taking into account a smaller number of records, considering sets of rain gauges networks of different densities, with a relatively uniform spatial distribution of observations. It was observed that the combination of the three databases can provide adequate approximations regardless of network configuration, for high or medium densities, although the quality of estimates is reduced if the density is low, with errors comparable to those obtained by IDW interpolation of rain gauge records. Additionally, for analysis under more unfavorable conditions, examples were resolved where a single remote sensing database was used in the assimilation, circumstance of special interest due to limited availability of weather radar records in Mexico. It was found that combining rain gauge records with satellite or radar data results in errors of a similar order of magnitude, and these are useful estimates for IDW interpolations calculated from the records of a medium density rain gauge network.

Because of the satisfactory results obtained in estimating daily rainfall distributed over Mexico City, the methodology was applied in other regions of the country where the availability of rain gauge records per unit area is significantly lower. Since the best estimates of rainfall were obtained by merging the three recommended databases, and the integration of rain gauge data with satellite data proved to be good approximations, the latter combination was chosen considering the difficulties involved in collecting and handling radar records due to scarcity. Three regions of interest were identified for their size, availability of hydrometric and pluviometric information with uneven distribution of records, geographic location that influences the magnitude of rainfall accumulation on several consecutive days, as well as its proximity to larger urban centers. The sites are located within the states of Chiapas-Tabasco, Coahuila-Nuevo León and Oaxaca-Veracruz.

The results concluded that the greatest errors in estimating assimilated rainfall are related to the lack of accuracy of rainfall calculated as satellite products, which is more common in low-magnitude rainfall fields. However, if the satellite data are of an order of magnitude similar to the available rain gauge records, the rainfall assimilation approach is more appropriate.

The precipitation analysis in each of the regions studied began with the identification of sub-catchment areas, defining as points of outflow the location of hydrometric stations with records of the dates analyzed. The quality review of the estimated distributed rainfall was complemented by considering it as input for the analysis of surface runoff from rainfall in its own basin. By having knowledge of the satisfactory background of the use of the mathematical model Iber for the calculation of the hydrograph at the exit from a basin [45], the results obtained with this model allowed to verify the quality of rainfall estimated by assimilation compared to those obtained by IDW interpolation of the rain gauge records considered in the assimilation. In order to increase the reliability of estimates, the variation of surface runoff at the outlet of the sub-basins was also obtained by a hydrological procedure using concentrated parameters.

The results of the analysis of smaller sub-basins show that there is a similarity between the quantities and variations in surface runoff obtained for the same rainfall pattern analyzed using a hydrodynamic or hydrological model, and the main differences from historical surface runoff records are considered to be due to overestimation of rainfall. The flows were better approximated by using estimated rainfall with the proposed assimilation method. In view of the results obtained for the larger subbasin, it is also concluded that the differences between the hydrometric record and the values obtained by simulation of the rainfall-runoff are mainly due to overestimation of precipitation; however, the result that represented a better approximation was obtained by using estimated rainfall with the IDW interpolation method, given the size of the area analyzed and the small number of rain gauge stations, located mainly on the perimeter of the catchment sub-basin, adverse conditions that caused a lack of approximation in the rainfall estimates over a larger number of cells, when the proposed assimilation procedure is used.

# Introducción

Uno de los propósitos del análisis de información meteorológica es caracterizar el estado de la atmósfera de manera consistente, tanto en el espacio como en el tiempo, por lo que la elección del método de análisis empleado es importante por la exactitud que puede alcanzarse a partir de un conjunto limitado de observaciones de la realidad física reciente.

Respecto a la precipitación, se sabe que su distribución natural es irregular a consecuencia de las características geográficas de la superficie terrestre, y en México se observa que la latitud tiene influencia en el incremento de la lluvia registrada del norte hacia el sur, además de que aproximadamente el 68% de la precipitación normal ocurre entre los meses de junio y septiembre, aspectos que invariablemente condicionan la disponibilidad del agua y el control de avenidas, dependiendo de la región del país y la época del año. La lluvia precipitada sobre la superficie nacional (1 964 375 km<sup>2</sup>) generalmente se ha registrado en intervalos de 24 horas, utilizando un conjunto de alrededor de 5 500 estaciones climatológicas, aunque al 31 de diciembre de 2017 solo 3 079 estaciones se encontraban operando bajo una distribución espacial irregular y dispersa, con menor densidad de observaciones en la parte norte, noroeste y sureste del país [23].

Para una estimación de la estructura espacial de las tormentas suelen considerarse mediciones puntuales localizadas dentro o en la proximidad de cada región de interés [50], aunque tales registros son únicamente representativos de la vecindad de los puntos de medición. Un procedimiento empleado desde la segunda mitad del siglo pasado es la aplicación de relaciones empíricas conocidas como Factores de Reducción por Área, con los cuales se procura la atenuación de lluvias observadas para estimar la precipitación media sobre una superficie, los que suelen considerarse útiles para definir lluvias aproximadas en cuencas de áreas y climatologías similares [54]. No obstante, dado que este concepto supone a la lluvia como un fenómeno estático, su utilidad se reduce cuando necesita tomarse en cuenta la cinemática del evento en relación con el área de captación [11].

La calidad y cantidad de los datos requeridos para el análisis de un sistema son cambiantes, y se tienen mayores exigencias cuando se refieren a de fenómenos de menor escala. Por ello, para el llenado de los vacíos de información conviene propagar la mayor cantidad de datos discretos disponibles.

Debido a que el uso de técnicas de interpolación puramente deterministas para el procesamiento de lluvias puede resultar insuficiente, en la modelación espacial del fenómeno también se han empleado métodos basados en la estructura estadística de los campos de observación, como es el caso del método Kriging, utilizado para la construcción de campos de tormenta en la Ciudad de México [22].

Una alternativa a la dificultad de interpolar en regiones donde la red de observaciones es irregular y dispersa es la aplicación de modelos para la combinación de diferentes bases de datos, los cuales han permitido considerar la información obtenida a través de percepción remota, misma que se considera de utilidad porque toma en cuenta la variación de una tormenta a distintas resoluciones. Por ejemplo, Rozante *et al.* practicaron un procedimiento de combinación de registros pluviométricos y valores estimados con la misión satelital TRMM (Tropical Rainfall Measuring Mission) mediante la aplicación del esquema de Barnes para definir campos de lluvia diaria con correcciones sucesivas [71].

Asimismo, dada la incertidumbre en la estimación de lluvias con registros dispersos, en especial de eventos convectivos por su mayor variabilidad espacial, Li *et al.* estudiaron sobre la incompatibilidad en la combinación de datos con escalas espaciales distintas; para no requerir un comportamiento estacionario de los mismos, aplicaron un método de interpolación de suavizado Kernel [52].

Con igual intención, y debido a que la calidad de estimaciones por interpolación de la distancia inversa o métodos geoestadísticos como Kriging pueden ser insuficientes incluso para redes pluviométricas relativamente densas, especialmente en escalas de tiempo reducidas, Nanding *et al.* evaluaron métodos de combinación de registros pluviométricos y estimaciones con radar meteorológico basados en Kriging, y analizaron diferentes densidades y configuraciones de la red de observaciones [59].

Más recientemente, Calvetti *et al.* insistieron en la importancia de considerar diferentes bases de datos para la estimación cuantitativa de la precipitación, enfatizando cómo las técnicas de integración de lluvia procuran la remoción de sesgos sistemáticos de las estimaciones de radar y satélite. Para combinar las estimaciones de radar, satélite y pluviómetros, utilizó un método numérico multicelda para solucionar una ecuación de Poisson, el cual es un algoritmo similar a las técnicas de asimilación de datos al emplear una condición de frontera como primera suposición [18].

Como el interés principal de esta tesis es proponer una metodología que permita mejorar la estimación del comportamiento histórico de un campo original de lluvia distribuida espacialmente en México, se consideró valioso tomar en cuenta las experiencias compartidas en documentos de investigación orientados a definir el estado de diagnóstico de un sistema como una representación discreta de valores localizados sobre las celdas de una malla regular, optando por el empleo del procedimiento de correcciones sucesivas que considera la función Kernel Gaussiana propuesta por Barnés. Ello se debe a los antecedentes satisfactorios de su aplicación en el análisis de tormentas, para aprovechar la confiabilidad de los registros diarios de la red de pluviómetros nacional, así como el conocimiento de la estructura espacial de las tormentas derivado del empleo de productos de sistemas satelitales y radares meteorológicos.

# Objetivos

## Objetivo principal

Proponer una metodología de asimilación de datos de lluvia diaria para estimar su distribución espacial sobre la superficie continental de la República Mexicana, mediante la combinación de registros pluviométricos con valores obtenidos de sensores satelitales y radares meteorológicos.

## Actividades complementarias

- Se analizó el comportamiento de la lluvia en el territorio nacional, para identificar fechas y sitios en donde ocurrieron precipitaciones importantes mediante la acumulación de láminas diarias y durante varios días consecutivos, tomando en cuenta la disponibilidad de registros pluviométricos y estimaciones como productos de radar y satélite.
- Se revisaron procedimientos de asimilación de datos, para identificar cuáles y bajo qué circunstancias es más conveniente su empleo para la estimación de la lluvia espacialmente distribuida.
- Se trabajó en la programación de rutinas propias de procesos de asimilación, para su aplicación sobre los esquemas de precipitación histórica para su análisis particular.
- Se evaluó la calidad de las lluvias estimadas sobre las celdas de la malla regular que representa el área de captación de lluvia sobre regiones de interés, a partir de índices estadísticos útiles para la medición de la correspondencia entre eventos estimados y observados.
- Se estudió la exactitud de las estimaciones de lluvia asimilada, en función de la densidad de las redes de pluviómetros disponibles para el análisis de eventos históricos. Además, se revisaron diferentes configuraciones espaciales de los pluviómetros de las redes consideradas.
- Se revisó la calidad de las estimaciones de lluvia espacialmente distribuida, para emplearlas como insumo de un modelo matemático de simulación de escurrimientos superficiales en dos dimensiones, y comparar los hidrogramas calculados a la salida de las subcuencas con los registros de gastos de estaciones hidrométricas seleccionadas.

# 1. Comportamiento de la precipitación en México

## 1.1. El fenómeno de la precipitación

Se denomina Ciclo hidrológico al proceso de transporte global y permanente del agua al interior del planeta Tierra [19], por lo que como concepto se ha vinculado al estudio de las causas relacionadas con la recirculación del agua entre la atmósfera y la tierra, desde la evaporación causada por la radiación solar sobre el agua de la superficie terrestre, hasta la condensación que provoca su caída en forma de precipitación [34], Figura 1-1.

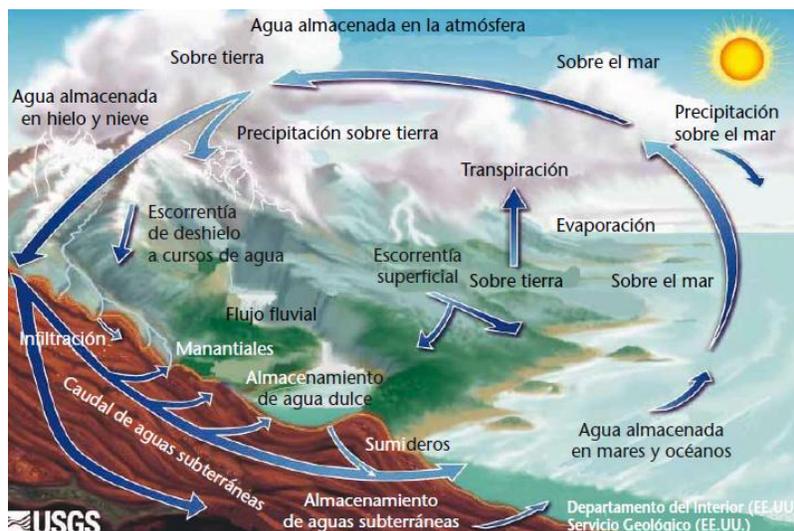


Figura 1-1: Ciclo hidrológico. Fuente: [62].

Como parte de la atmósfera, el agua está presente en cualquiera de sus tres estados: cristales de hielo, minúsculas gotas, o vapor de agua proveniente de la evaporación ocurrida en los océanos y aguas continentales. En este sentido, la cantidad máxima de vapor de agua contenida en un volumen definido de aire es función de la temperatura, y para que el vapor de agua contenido en la atmósfera llegue a condensarse se requiere la saturación del aire; en otros términos, a cada temperatura de la masa de aire le corresponde una cantidad específica de vapor de agua para alcanzar su saturación, lo que generalmente tiene lugar durante un proceso de enfriamiento [19].

Respecto a la precipitación como un fenómeno físico, su ocurrencia está supeditada a la existencia de nubes, entendidas como porciones de aire enturbiado por vapor de agua condensado, en estado líquido o sólido según el proceso que dio origen a su condensación, por lo que pueden estar constituidas por gotas de agua pequeñas conocidas como nube cálida, por gotas y cristales de hielo que conforman nubes mezcladas, o solamente cristales que componen nubes frías [19].

Las nubes se forman por el enfriamiento del aire por debajo de su punto de saturación, lo que ocurre a través de procesos relacionados con el movimiento vertical ascendente del aire, generalmente con la rapidez suficiente para que no ocurra un intercambio de calor eficiente en el sistema, por lo que las presiones a las que estará sometido irán disminuyendo, con la consecuente expansión y enfriamiento del aire. Si el enfriamiento persiste después de la formación de las nubes, el volumen de precipitación y su intensidad dependerán del contenido de humedad del aire y la velocidad de su ascenso vertical, relacionados con el mecanismo de la tormenta [19].

La precipitación de las gotas sucede cuando éstas alcanzan un tamaño (del orden de 0.1 mm) y peso suficiente para vencer el empuje de las corrientes ascendentes de aire. Dentro de este desarrollo, cabe mencionar a los denominados núcleos de condensación, presentes en la atmósfera como resultado de la evaporación de espuma marina y procesos de combustión, los que promueven la formación de gotas por tratarse de partículas diminutas ( $10^{-3}$  a  $10 \mu\text{m}$ ) conocidas como aerosoles, sobre las cuales el vapor de agua puede adherirse y provocar su disolución, y la condensación desencadenarse cuando la humedad relativa es próxima al 75%, hasta formarse gotas mayores que acaban por precipitarse [20], Figura 1-2.

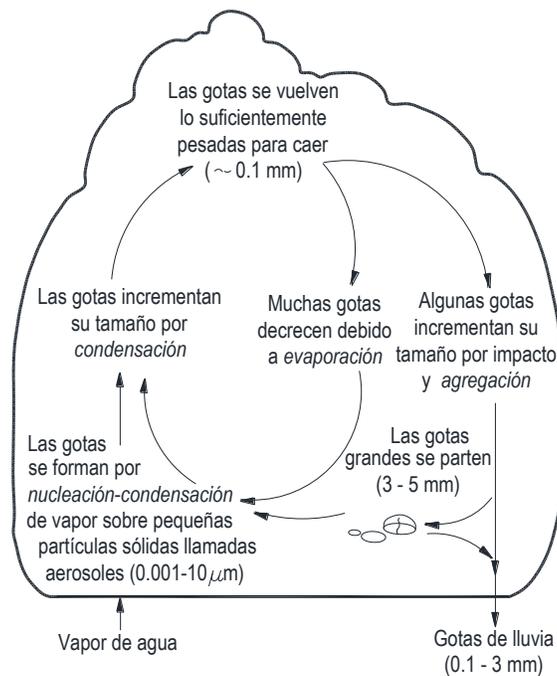


Figura 1-2: Etapas en la formación de las gotas de lluvia. Fuente: [20].

Con relación a las causas que promueven el ascenso de la masa de aire húmedo en la atmósfera, fenómeno necesario para su eventual enfriamiento y condensación, se distinguen tres mecanismos principales que lo favorecen, de naturaleza orográfica, convectiva y frontal (Figura 1-3).

La precipitación orográfica ocurre por el ascenso de masas de aire cargadas de humedad, cuando la trayectoria es obligada por la obstrucción que ejercen las barreras montañosas [19]. En estas condiciones, el tipo de nube a formarse dependerá de la velocidad de ascenso, que afecta a la rapidez en el descenso de la temperatura, de manera que a mayor velocidad de ascenso se promueve la formación de nubes cumuliformes (de aspecto amontonado), y con menor velocidad la generación de nubes estratiformes (delgadas y extendidas) [4], aunque en ambos casos la precipitación siempre será mayor en la ladera de barlovento [19].

La precipitación convectiva se genera por una falta de estabilidad entre masas de aire de distinta temperatura, condición que desencadena un fenómeno de turbulencia que promueve el ascenso del aire caliente (fase inicial en la que el cúmulo de nubes tendrá su base a unos 2 000 m sobre el nivel de la superficie terrestre, y alcanzará una altura y extensión aproximadas de 6 km y 8 km, respectivamente). El movimiento ascendente induce al descenso de la temperatura hasta alcanzar el punto de rocío y su posterior condensación (etapa de madurez, en la que el tope de la nube alcanza niveles próximos a los 12 km, con una extensión de la base de hasta 16 km). Este ascenso da lugar a la formación de nubosidad cumuliforme, y se detiene hasta que la masa de aire entra en equilibrio térmico con el ambiente que la rodea (fase de disipación), por lo que la tormenta logra su desarrollo máximo y comienza a desvanecerse, llegando a su final cuando la masa de aire frío y húmedo desaparece, después de haberse generado una precipitación intensa de corta duración sobre una superficie de magnitud reducida [4].

La precipitación frontal está asociada a los frentes o superficies de contacto entre masas de aire en movimiento, cuyas condiciones de contenido de humedad y temperatura son muy distintas, por lo que en su interacción no se mezclan y el aire más caliente es forzado a elevarse sobre el aire frío, cuya densidad es mayor [19]. Un frente cálido corresponde con el movimiento de aire caliente, que en su avance asciende lentamente por encima de una masa de aire más frío, siguiendo una pendiente suave que promueve la formación de nubes estratiformes, que ocasionan precipitaciones de gran duración y baja intensidad. Por el contrario, en el frente frío ocurre un desplazamiento ascendente y súbito de una masa de aire caliente a consecuencia del empuje que le causa la introducción en forma de cuña de aire frío en movimiento, y la turbulencia creada favorece la formación de nubosidad de naturaleza cumuliforme [4].

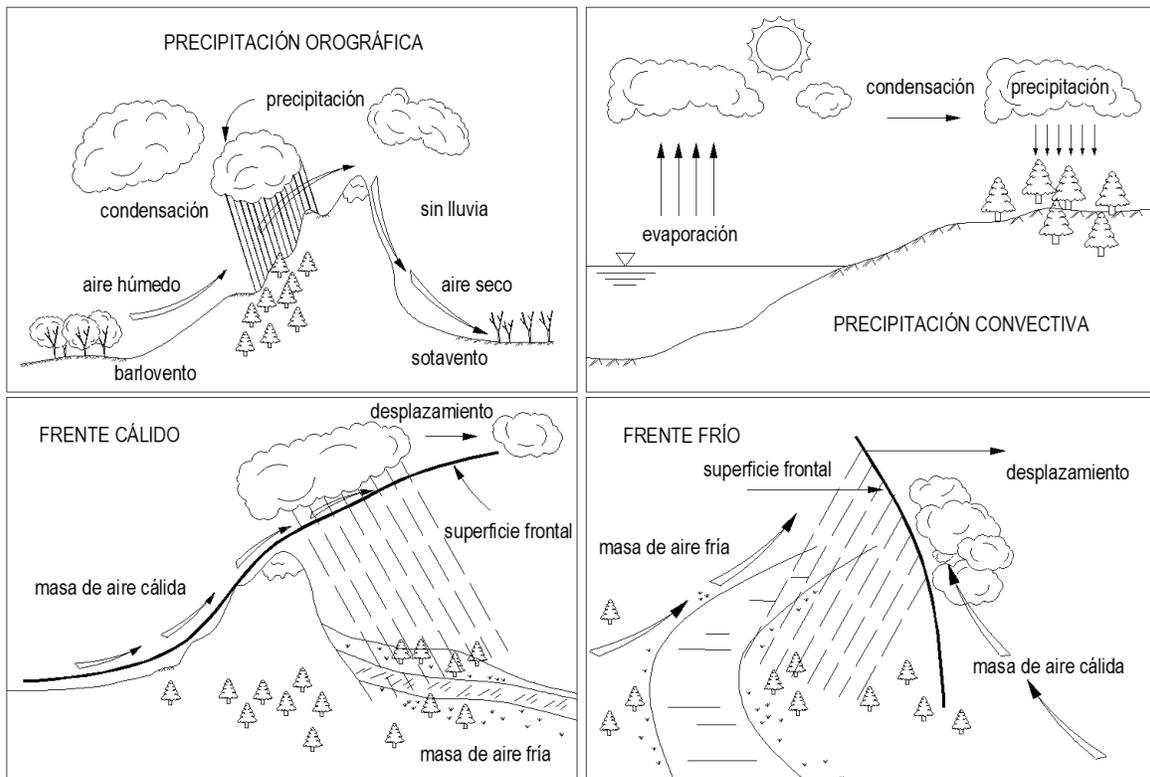


Figura 1-3: Diferentes mecanismos de precipitación. Fuente: Modificado de [1, 19].

## 1.2. La precipitación sobre la República Mexicana

### 1.2.1. Características de la precipitación

De acuerdo con la Comisión Nacional del Agua (CONAGUA), respecto al cálculo de la precipitación normal anual del país (que corresponde con los promedios de precipitación calculados para un periodo climatológico uniforme y de longitud representativa, con un mínimo recomendado de 30 años), en el periodo de 1981 a 2010 México recibió anualmente un volumen de agua en forma de precipitación de aproximadamente  $1\,449\,471\text{ hm}^3$ , del cual se estima que el 72.1% regresó a la atmósfera por evapotranspiración, 21.4% escurrió superficialmente por ríos o arroyos, y el 6.4% se infiltró en el subsuelo y recargó los acuíferos [24].

En función de lo anterior, y tomando en cuenta los flujos de salida y entrada de agua con los países vecinos, se estima que México cuenta con alrededor de  $451\,585\text{ hm}^3$  de agua dulce renovable (cantidad equivalente a cubrir toda la superficie del territorio nacional con una lámina de 23 cm de agua), definida como la cantidad máxima de agua que es factible explotar anualmente sin alterar el ecosistema, pero se sabe que el 67% de ella ocurre en el sur y sureste del país, y únicamente el 33% restante corresponde al norte y noreste del territorio. Esta agua renovable se calcula como el escurrimiento natural medio superficial interno anual, más la recarga total anual de los acuíferos, más los flujos de entrada menos los flujos de salida de agua a otras regiones [24]. Los componentes y valores medios anuales del ciclo hidrológico en México se aprecian en la Figura 1-4.

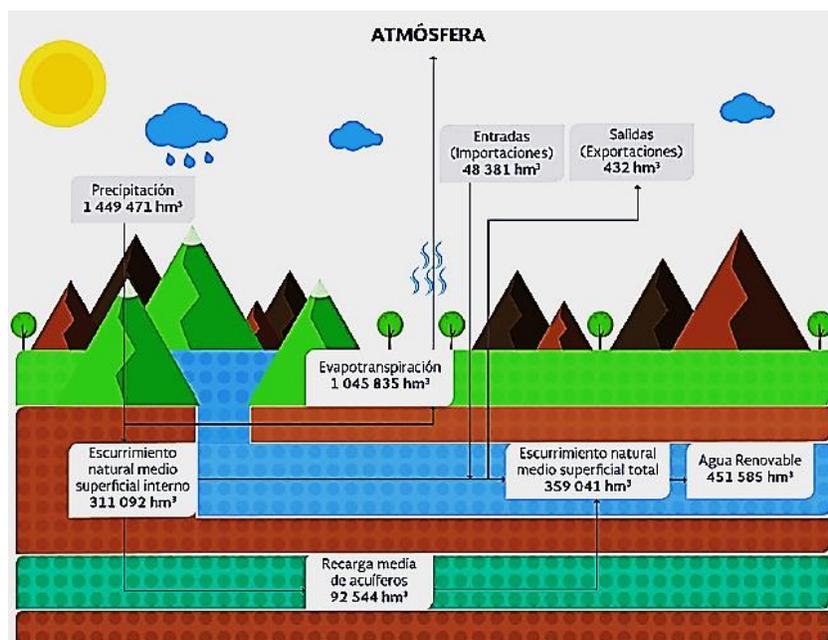


Figura 1-4: Componentes y valores medios anuales del ciclo hidrológico en México. Fuente: [24].

En México, el tipo de precipitación que prevalece es de naturaleza pluvial, pero su distribución es irregular a consecuencia de sus características geográficas [19]. La precipitación normal anual en el periodo 1981-2010 fue de 740 mm, cuya distribución espacial resultó bastante irregular [23], como se aprecia en la variación de las precipitaciones correspondientes a las 37 regiones hidrológicas (RH) en que se encuentran organizadas las cuencas del país, valores consignados en

la Tabla 1.1. De manera similar, en la Tabla 1.2 se observa la distribución temporal de la precipitación normal mensual nacional, correspondiente al mismo periodo de treinta años, sobre cada una de las trece regiones hidrológico-administrativas (RHA) en que está organizado al país [23].

Según se aprecia, la acumulación de la precipitación es mayor en la parte sur del territorio, que comprende las regiones V Pacífico Sur, X Golfo Centro, XI Frontera Sur y XII Península de Yucatán, como consecuencia de las condiciones de humedad atmosférica y los factores climáticos (viento, temperatura y presión atmosférica) que allí prevalecen, muy favorables para la precipitación pluvial característica de la zona, de naturaleza convectiva y ciclónica (movimiento de las masas de aire desde una región de alta hacia una de baja presión). Por el contrario, en la parte norte que comprende las regiones I Península de Baja California, II Noroeste, III Pacífico Norte, VI Río Bravo y VII Cuencas Centrales del Norte, existe una combinación de factores climáticos y de masas de aire continental seco que no favorecen la precipitación pluvial.

Tabla 1.1: Precipitación pluvial normal anual por RH, 1981–2010, en mm, [23].

Clave	RH	Extensión territorial continental (km <sup>2</sup> )	Número de cuencas hidrológicas	Precipitación normal anual 1981-2010 (mm)	Clave	RH	Extensión territorial continental (km <sup>2</sup> )	Número de cuencas hidrológicas	Precipitación normal anual 1981-2010 (mm)
1	B.C. Noroeste	28 492	16	209	20	Costa Chica de Guerrero	39 936	32	1 282
2	B.C. Centro-Oeste	44 314	16	116	21	Costa de Oaxaca	10 514	19	951
3	B.C. Suroeste	29 722	15	200	22	Tehuantepec	16 363	15	884
4	B.C. Noreste	14 418	8	151	23	Costa de Chiapas	12 293	25	2 220
5	B.C. Centro-Este	13 626	15	132	24	Bravo-Conchos	229 740	37	399
6	B.C. Sureste	11 558	14	291	25	San Fernando-Soto la Marina	54 961	45	703
7	Río Colorado	6 911	4	98	26	Pánuco	96 989	77	855
8	Sonora Norte	61 429	9	297	27	Norte de Veracruz (Tuxpan-Nautla)	26 592	12	1 422
9	Sonora Sur	139 370	16	483	28	Papaloapan	57 355	18	1 440
10	Sinaloa	103 483	30	747	29	Coatzacoalcos	30 217	15	2 211
11	Presidio-San Pedro	51 717	26	819	30	Grijalva-Usumacinta	102 465	83	1 703
12	Lerma-Santiago	132 916	58	717	31	Yucatán Oeste	25 443	7	1 175
13	Río Huicicila	5 225	6	1 400	32	Yucatán Norte	58 135	2	1 143
14	Río Ameca	12 255	9	1 063	33	Yucatán Este	38 308	6	1 210
15	Costa de Jalisco	12 967	11	1 144	34	Cuencas Cerradas del Norte	90 829	22	298
16	Armería - Coahuayana	17 628	10	866	35	Mapimí	62 639	6	292
17	Costa de Michoacán	9 205	6	944	36	Nazas-Aguanaval	93 032	16	393
18	Balsas	118 268	15	947	37	El Salado	87 801	8	393
19	Costa Grande de Guerrero	12 132	28	1 215		<b>Total</b>	<b>1 959 248</b>	<b>757</b>	<b>740</b>

Tabla 1.2: Precipitación pluvial normal mensual y anual por RHA, 1981 – 2010, en mm, [23].

Clave	RHA	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Anual
I	Península de Baja California	20	19	14	4	1	1	10	26	32	11	10	20	168
II	Noroeste	24	21	12	6	4	19	108	103	58	25	17	31	428
III	Pacífico Norte	31	16	8	6	9	66	194	188	142	52	26	29	765
IV	Balsas	12	8	6	11	48	179	199	197	194	84	15	6	962
V	Pacífico Sur	8	8	6	15	71	230	200	219	242	113	20	7	1 139
VI	Río Bravo	19	11	11	17	28	40	63	61	64	32	12	15	372
VII	Cuencas Centrales del Norte	18	9	6	12	27	56	79	71	67	29	11	13	398
VIII	Lerma-Santiago-Pacífico	22	11	4	6	23	131	197	180	153	60	13	10	808
IX	Golfo Norte	26	20	19	38	67	120	137	119	166	89	30	23	855
X	Golfo Centro	51	40	30	43	84	222	261	264	293	179	97	64	1 626
XI	Frontera Sur	65	54	36	49	135	276	223	265	331	224	109	76	1 842
XII	Península de Yucatán	45	35	31	39	90	167	153	173	208	147	72	49	1 207
XIII	Aguas del Valle de México	11	11	12	28	51	109	126	115	110	57	13	6	649
	<b>Total</b>	<b>25</b>	<b>17</b>	<b>13</b>	<b>18</b>	<b>42</b>	<b>102</b>	<b>134</b>	<b>134</b>	<b>135</b>	<b>69</b>	<b>27</b>	<b>23</b>	<b>740</b>

Para ilustrar el comportamiento espacial de la precipitación normal sobre el territorio nacional [24], en la Figura 1-5 se muestran tres perfiles que representan la variación de la precipitación a lo largo de tres líneas de corte:

- A-A', de Guaymas a Matamoros
- B-B', entre Puerto Vallarta y Veracruz
- C-C', de Acapulco a Chetumal.

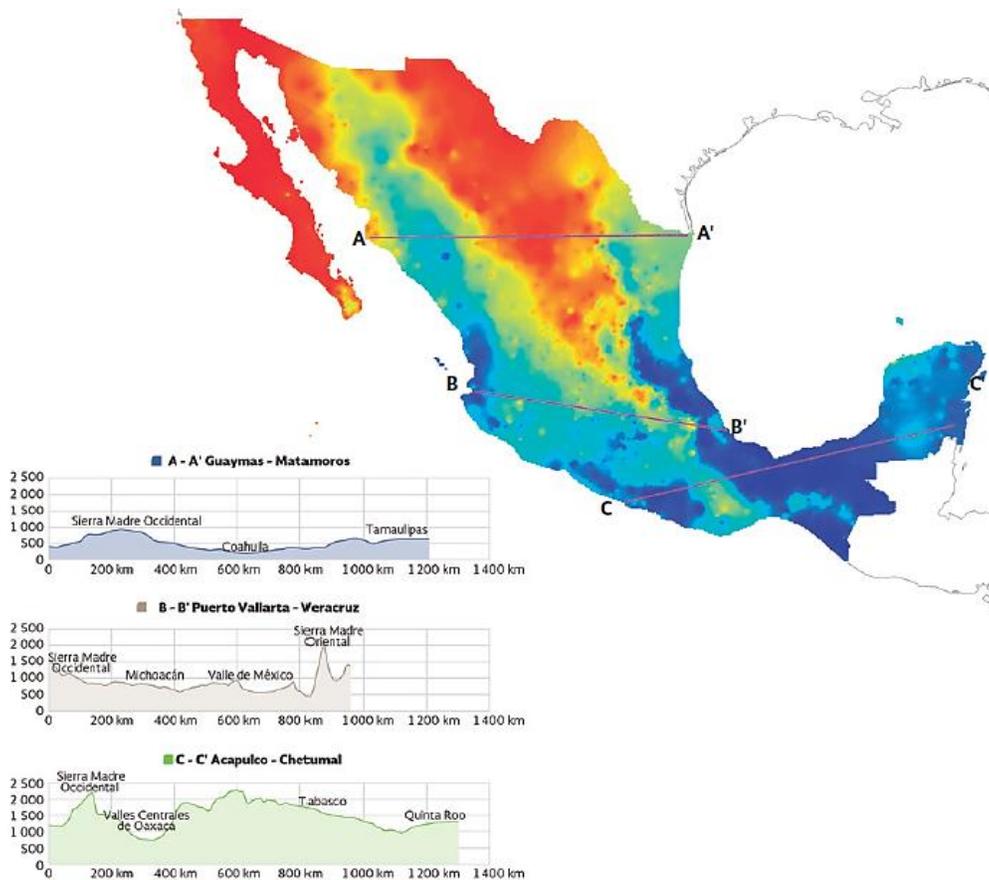


Figura 1-5: Perfiles de precipitación pluvial normal anual, 1981 – 2010, en mm. Fuente: [24].

De manera general, la latitud influye sobre el incremento de la precipitación que se registra del norte hacia el sur del país. La concentración máxima ocurre en el sureste, dada su altitud elevada y la incidencia recurrente de fenómenos meteorológicos; las zonas altas, constituidas por las sierras Madre Oriental, Occidental, del Sur y Volcánica Transversal intervienen al paso del aire húmedo, y su efecto se aprecia con precipitaciones superiores a los 1 000 mm, mientras que dentro de las regiones así limitadas se han registrado precipitaciones mayores a los 500 mm, y la zona con registro de menores precipitaciones corresponde a la Península de Baja California [19].

Respecto a la variación temporal de la precipitación nacional, en la Figura 1-6 se observa que aproximadamente el 68% de la precipitación normal ocurre entre los meses de junio y septiembre [23], distribución que invariablemente condiciona los problemas tanto de disponibilidad de agua como de control de inundaciones, dependiendo de la época del año y la región del país.

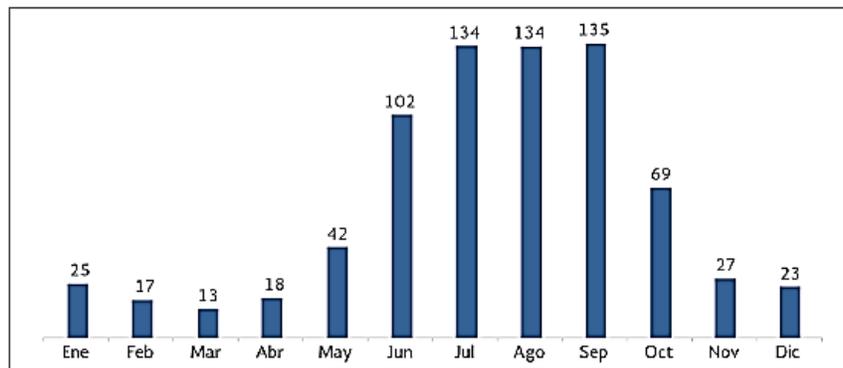


Figura 1-6: Precipitación pluvial normal mensual nacional, 1981 – 2010, en mm. Fuente: [23].

Cuando la disponibilidad de agua sobre un área geográfica es significativamente menor respecto a las condiciones habituales, a dicha anomalía climatológica transitoria se le conoce como sequía, la cual ocasiona graves desequilibrios hidrológicos que perjudican especialmente a los sistemas de producción agrícolas, pues cuando la lluvia es escasa e infrecuente y la temperatura aumenta, la vegetación se desarrolla con dificultad, provocando pérdidas cuantiosas que se consideran como desastres naturales, además de la degradación del suelo y el riesgo de incendios forestales [24].

En general, son los meses de mayo y noviembre los que marcan el inicio y final de la temporada de lluvias en México, justo cuando la evolución de la sequía suele revisarse. De esta manera, la Comisión Nacional del Agua ha identificado alrededor de cien municipios con alta vulnerabilidad de sequía, localizados en las regiones Noroeste, Centro y vertiente del Pacífico, entre los que destacan los municipios de Nogales y Santa Cruz en Sonora, Ascensión y Uruachi en Chihuahua, y Uruapan en Michoacán de Ocampo. En estas circunstancias, México participa en el Monitor de Sequía de América del Norte, en alianza con Estados Unidos de América y Canadá, con la intención de analizar las condiciones climáticas de forma continua y a gran escala, siendo el Índice Estandarizado de Precipitación uno de los parámetros principales para la identificación de periodos húmedos y secos, al considerar únicamente registros de precipitación mensual, con un mínimo de 20 a 30 años, para su ajuste y posterior comparación, lo que permite evaluar la severidad de la sequía en el corto y largo plazo [23].

Como una circunstancia opuesta, pero de importancia semejante, resulta de interés el estudio de los ciclones tropicales en la distribución de la lluvia sobre el territorio nacional. Se sabe por ejemplo que la mayor parte del transporte de humedad del mar hacia las zonas semiáridas del país es

generada por los ciclones tropicales [24]. Así, existen regiones donde las lluvias ciclónicas representan gran parte de la precipitación pluvial anual, como es el caso de la Península de Baja California [44].

Los ciclones tropicales (CT) son perturbaciones atmosféricas que se originan entre las latitudes 5° y 15°, tanto en el hemisferio norte como en el sur, en las aguas oceánicas cálidas de las zonas tropicales, cuando su temperatura es mayor a 26°C. Son grandes masas de aire cálido y húmedo con fuertes vientos que giran alrededor de una zona de baja presión, que en general se trasladan hacia los polos con velocidades comprendidas entre 10 y 40 km/h, con una componente hacia el oeste en latitudes bajas, y hacia el este en latitudes superiores a los 25 grados [70]. La disipación de estos sistemas ocurre cuando en su trayectoria se alejan de su fuente de energía básica, por dirigirse al norte sobre aguas frías, por ejemplo, o cuando se adentran en las masas continentales y a su paso encuentran cadenas montañosas, lo que genera un cambio drástico en los contenidos de vapor de agua, las variaciones verticales del viento y los flujos convectivos [67].

En el caso particular de México, el territorio nacional se ubica en un sitio propicio para recibir el impacto de los CT, dado que sus costas son cercanas a dos de las principales zonas de generación de estos fenómenos, el mar Caribe y el Golfo de Tehuantepec [67]. De manera convencional, la temporada de huracanes inicia oficialmente el 15 de mayo en el Océano Pacífico Oriental, y el 1 de junio en el Pacífico Central y el Océano Atlántico, y finaliza en ambos océanos el 30 de noviembre [70].

Entre las manifestaciones más notables de los CT se tienen sus intensas ráfagas de viento, oleajes, mareas de tormenta y precipitaciones pluviales [64]. De acuerdo con la presión en su centro y la velocidad de sus vientos, los ciclones se clasifican como depresión tropical (de 1 008 a 1 005 mb, o velocidad menor a 63 km/h), tormenta tropical (1 004 a 985 mb, o velocidad entre 63 y 118 km/h) y huracán (presión menor que 984 mb, o velocidad mayor a 119 km/h), clasificados por categorías de la escala Saffir-Simpson [70].

Por otro lado, la gran intensidad, duración y extensión del campo de vientos de los CT promueve la generación del oleaje más severo, que incrementa su rapidez al crecer en longitud y altura, hasta empatar la velocidad del viento cuando alcanza sus condiciones máximas de desarrollo; de esta manera, aun cuando el oleaje máximo en alta mar es mucho mayor que el que incide en la costa, este último no es despreciable por la fuerza que puede ejercer sobre las estructuras, siendo esta la razón por la que es considerado en el diseño de las obras costeras y marítimas [67].

Como efecto adicional, se conoce por marea de tormenta a la sobreelevación del nivel medio del mar cuando un ciclón se aproxima hacia la costa (originada por la acción que el viento ejerce sobre la superficie del mar, así como por el desbalance del momentum de entrada y salida de la franja entre el litoral y la zona de rompientes), para lo cual resulta común la observación de incrementos de entre dos y tres metros sobre el nivel original, que pueden favorecer a la supervivencia de un oleaje mayor (sin rompimientos con un mínimo de pérdidas de energía) y ser más destructivo en las franjas costeras, o producir remansos en las corrientes naturales que descargan al mar [67].

Por otra parte, dado el mecanismo de su desarrollo, los ciclones tropicales elevan enormes cantidades de vapor de agua que, bajo condiciones atmosféricas y orográficas adecuadas, se precipitan inexorablemente, con una intensidad que dependerá del área en planta y la velocidad de desplazamiento del evento particular [70]. En este sentido, existen trabajos en los que se ha estimado la magnitud de la aportación de las lluvias ciclónicas respecto a la precipitación total media anual sobre el territorio nacional. Con la intención de caracterizar la lluvia ciclónica en

México, Jiménez (2016) realizó un análisis de los eventos ciclónicos que generaron lluvia durante el periodo comprendido entre los años 1998 y 2013, considerando estimaciones de precipitación basadas en percepción remota, particularmente los de la misión satelital TRMM. Con la identificación de 102 eventos durante el periodo señalado, observó lo siguiente [44]:

- Quedó de manifiesto que la actividad ciclónica es más intensa en la costa del Pacífico (55% de eventos), además de que las tormentas tropicales ocurren con mayor frecuencia (43%), seguidas de las depresiones tropicales (28%).
- Respecto a las aportaciones a la lluvia anual por ciclones en México, se observaron variaciones importantes pues, por ejemplo, en el año 2004 el tamaño de la contribución fue del 3.5% únicamente, mientras la aportación máxima fue del 24% en el año 2013.
- La distribución de la aportación media de la lluvia por ciclones respecto a la precipitación total sobre el territorio nacional durante el periodo analizado se muestra en la Figura 1-7, donde se aprecia que la contribución tanto en la zona noroeste como en una parte central del país es menor al 10%, en el sureste la contribución es de entre 10 y 20%, en la costa sur del Océano Pacífico y norte del Golfo de México los aportes son del 20 al 30%, y en la Península de Baja California hay aportes del 40 al 50%.

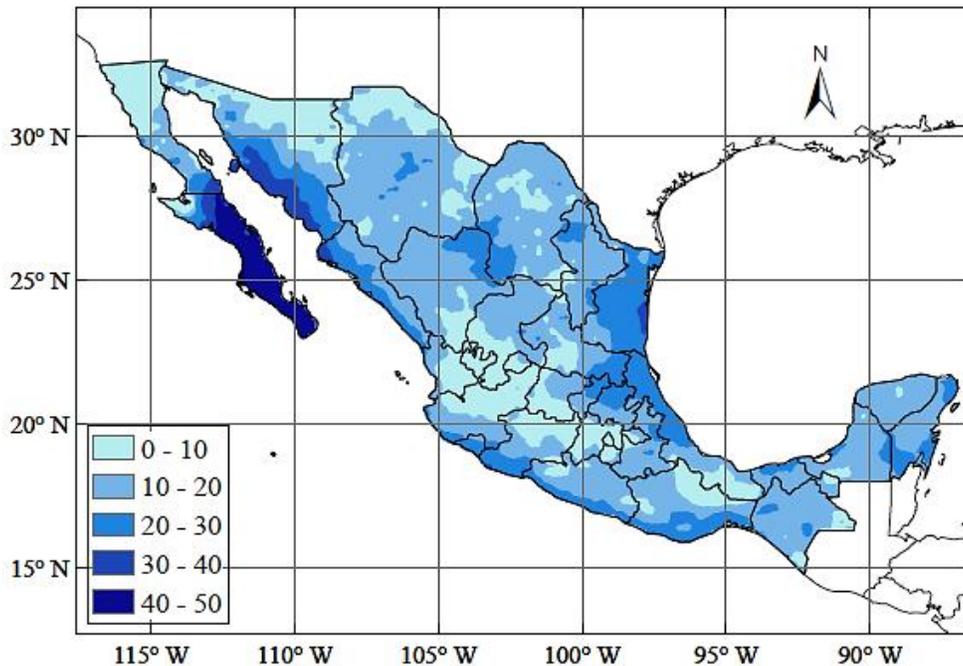


Figura 1-7: Contribución media de la lluvia por ciclones tropicales a la precipitación total en México, durante el periodo de 1998 a 2013, en porcentaje. Fuente: [44].

Adicionalmente, como un elemento importante para la caracterización de las precipitaciones sobre la República Mexicana, resulta de interés la identificación de la magnitud de las tormentas históricas extremas acumuladas en 24 horas.

En la Tabla 1.3 se aprecian los valores de las láminas de lluvia generadas por algunos de los ciclones más destructivos que han afectado al país, información que sirve como una referencia de comparación para los resultados que podrían obtenerse en un estudio sobre las tormentas de diseño de una región particular [54].

Tabla 1.3: Lluvias generadas durante eventos ciclónicos en la república mexicana [54, 89].

Huracán	Fecha	Localidad	Lluvia media anual estatal, mm	Lluvia máxima en 24 h, mm
Gladys	1 sep 1955	Catemaco, Ver		253
Hilda	19 sep 1955	Villa de Reyes, SLP		384
Janet	29 sep 1955	Villa de Reyes, SLP		535
Beulah	22 sep 1967	Monterrey, NL		541
Naomi	13 sep 1968	Río Baluarte, Sin		500
Liza	30 sep 1976	La Paz, BCS	230.0	410
Gilberto	15 sep 1988	Monterrey, NL		310
Roxanne	20 oct 1995	Martínez de la Torre, Ver		297
Paulina	9 oct 1997	Acapulco, Gro	1 400.0	411
Hernán	4 oct 1996	P. La Villita, Mich		422
Emily	18 jul 2005	El Mezquite, Tamps	738.8	350
Wilma	20 oct 2005	Isla Mujeres, QR	1 093.3	1 576
John	1 sep 2006	San Bartolo, BCS	130.9	449
Dean	21 ago 2007	Tecolutla, Ver	1 502.6	391
Jimena	2 sep 2009	Guaymas, Son	444.9	515
Alex	27 jun 2010	Laguna Madre, Tamps	1 011.2	446
Karl	14 sep 2010	Misantla, Ver	1 791.1	355
Beatriz	19 jun 2011	S-SE de La Fortuna, Jal	602.3	222
Jova	5 oct 2011	S de La Fortuna, Jal	602.3	374
Ingrid	17 sep 2013	Xilitla, SLP	875.5	316
Manuel	14 sep 2013	Acapulco, Gro	1 270.5	255
Odile	16 sep 2014	Mulegé, BCS	334.7	127
Patricia	23 oct 2015	Petalcalco, Gro	1 003.3	279
Hanna	26 jul 2020	Cadereyta Jiménez, NL	525.5	474
Otis	25 oct 2023	Acapulco, Gro	1 030.7	

## 1.2.2. Registros históricos de precipitación y el análisis de su comportamiento espacial

- **Registros de precipitación**

Como variable meteorológica, la precipitación es susceptible de medirse respecto a su cantidad, intensidad y regularidad, atendiendo diversas técnicas [19]. La cantidad de precipitación, en cualquiera de sus formas, se mide como la columna vertical de agua acumulada sobre una superficie plana e impermeable, en unidades de milímetro y décimas de milímetro, con una equivalencia en volumen de litros de agua por metro cuadrado de superficie [64].

De manera tradicional, la altura de precipitación total se registra en intervalos preestablecidos de 24 horas, que inician a las ocho de la mañana de cada día, utilizando como instrumento de captación y lectura de precipitación al pluviómetro, el cual consiste de un depósito cilíndrico en cuyo interior se aloja un vaso medidor [19].

En general, se sabe que las mediciones están en función del valor real de la precipitación, de la naturaleza del pluviómetro, de las características del sitio y de las condiciones meteorológicas, lo que significa que dichas mediciones están caracterizadas por errores propios de cada estación de registro, donde regularmente se han observado efectos de subestimación del valor real, desviación identificada principalmente como una influencia del viento [19], por lo cual la condición recomendada de instalación del equipo consiste en rodearlo uniformemente por obstáculos de la misma altura, para contrarrestar las corrientes de cualquier dirección [64].

Respecto a la calidad de los registros de precipitación, se ha reconocido la importancia de los errores introducidos por la técnica de medición del observador, los que se consideran sistemáticos

si tienen su origen básicamente en la instrumentación, o errores aleatorios que pueden caracterizarse por medios estadísticos [62], y su identificación es posible con el análisis de parámetros de posición y dispersión de la variable aleatoria precipitación, correspondientes a una población de interés, entre los que destacan su valor medio, desviación estándar y coeficiente de variación, así como la determinación de los valores máximos y mínimos históricos [30].

A nivel nacional la información de lluvias diarias se ha registrado mediante un conjunto de alrededor de 5 500 estaciones climatológicas, con información originada desde principios del siglo veinte, y actualizada por organismos y direcciones de la Comisión Nacional del Agua [94].

Al 31 de diciembre de 2017, en México se contaba con 3 079 estaciones climatológicas en operación (entre ellas, 81 son observatorios meteorológicos con transmisión en tiempo real), de las cuales únicamente 1 788 fueron consideradas para el cálculo de la precipitación normal relativa al periodo 1981-2010, lo que significa que la disponibilidad de datos es escasa en un gran número de estaciones [23].

Adicionalmente, es importante observar que en México se tiene una densidad diferencial de estaciones climatológicas, donde la menor densidad se presenta en la parte norte, noroeste y sureste del territorio [23], como se aprecia en la Tabla 1.4 y Figura 1-8.

Tabla 1.4: Estaciones climatológicas operadas por CONAGUA, 2017, [23].

Clave	RHA	Número de estaciones
I	Península de Baja California	211
II	Noroeste	150
III	Pacífico Norte	151
IV	Balsas	376
V	Pacífico Sur	123
VI	Río Bravo	237
VII	Cuencas Centrales del Norte	203
VIII	Lerma-Santiago-Pacífico	558
IX	Golfo Norte	332
X	Golfo Centro	175
XI	Frontera Sur	262
XII	Península de Yucatán	165
XIII	Aguas del Valle de México	136
<b>Total</b>		<b>3 079</b>

Debido a su capacidad para producir registros continuos de precipitación, los pluviógrafos son los instrumentos destinados a medir la distribución temporal de la lluvia, por lo que permiten conocer la intensidad de la precipitación ocurrida sobre un punto de observación.

Desafortunadamente, la densidad y amplitud de los registros pluviográficos sobre el territorio nacional ha permanecido muy limitada, por lo que tanto el número de estaciones como la accesibilidad a la información suelen significar una desventaja para su aplicación. En particular, la Ciudad de México supone una excepción, al ser la región más monitoreada del país, y disponer de alrededor de 60 estaciones pluviográficas, operadas por el Sistema de Aguas de la Ciudad de México (SACMEX), con registros desde de la década de los 80 del siglo pasado [93]. En la Tabla 1.5 se indica la cantidad de pluviógrafos operando en las alcaldías de la Ciudad, y algunos municipios del Estado de México.

Al inicio del siglo veintiuno, en México se reconoció la necesidad de incrementar la cantidad de registros de la variación temporal de la lluvia sobre el territorio nacional y disponer de los datos en tiempo cuasi real.

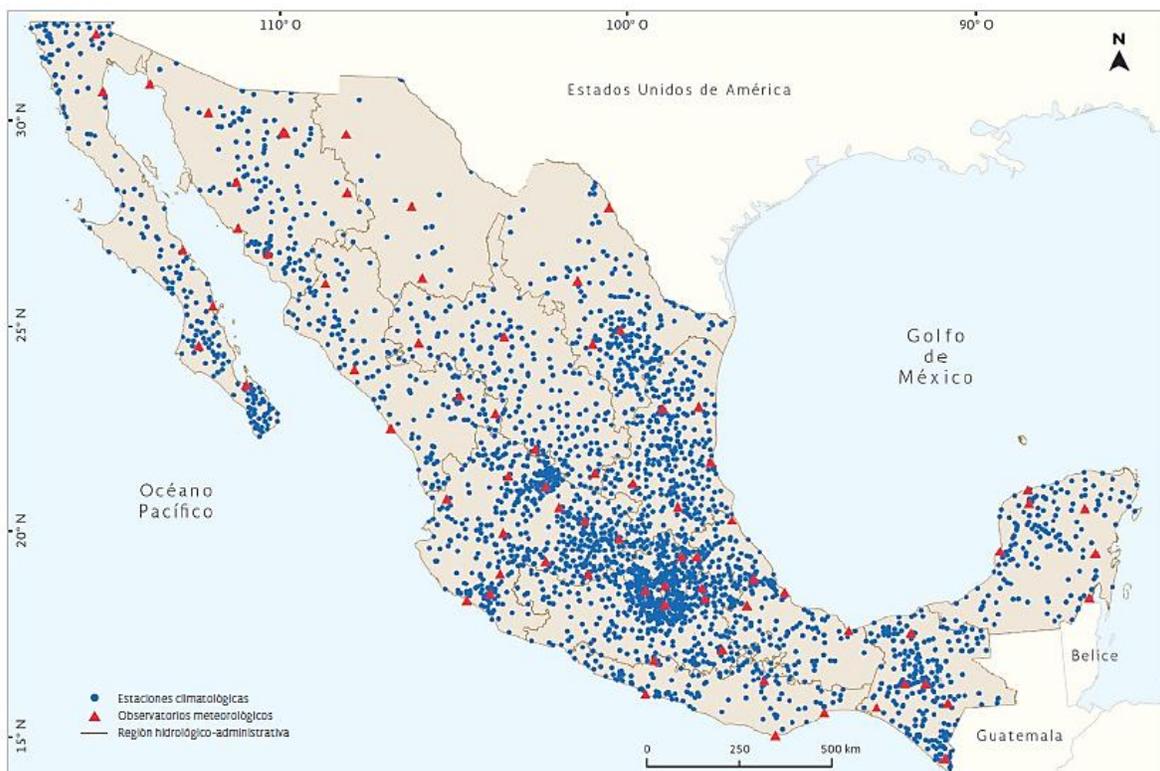


Figura 1-8: Estaciones climatológicas sobre la República Mexicana, 2017. Fuente: [23].

Tabla 1.5: Estaciones pluviográficas operadas por el SACMEX, [93].

Alcaldía (Cd. de Méx.)	Núm. de pluviógrafos	Municipio (Edo. de Méx.)	Núm. de pluviógrafos
Álvaro Obregón	5	Atizapán de Zaragoza	1
Azcapotzalco	3	Cuautitlán Izcalli	1
Benito Juárez	2	Chalco	1
Coyoacán	2	Coyotepec	1
Cuajimalpa	5	Ecatepec	2
Cuauhtémoc	3	Tultitlán	1
Gustavo A. Madero	7	Huixquilucan	2
Iztacalco	1	Ixtapaluca	1
Iztapalapa	8	Naucalpan	1
Magdalena Contreras	4	Ocoyoacac	1
Miguel Hidalgo	3	Temamatla	1
Milpa Alta	3	Texcoco	1
Tláhuac	3	Tezoyuca	1
Tlalpan	7	Tlalnepantla	2
Venustiano Carranza	3		
Xochimilco	2		

Actualmente se encuentran en operación las denominadas estaciones meteorológicas automáticas, EMA's, que de acuerdo con el Servicio Meteorológico Nacional (SMN), están colocadas en sitios estratégicos para el monitoreo y recopilación de algunas variables meteorológicas (velocidad y dirección del viento, presión atmosférica, temperatura y humedad relativa, radiación solar y precipitación), generando archivos cada diez minutos respecto al promedio de todas las variables, y transmitiendo vía satélite la información, en intervalos de una o tres horas [89]. En general, se considera que las EMA's tienen por finalidad el incremento del número y fiabilidad de las observaciones en superficie, y lo consiguen [64]:

- aumentando la densidad de una red existente, en especial en sitios de difícil acceso;
- proporcionando datos fuera del horario de las estaciones atendidas por personal;
- mejorando la fiabilidad de las mediciones con el uso de tecnologías sofisticadas;
- asegurando la homogeneidad de observaciones al normalizar las técnicas de medición;
- reduciendo los errores humanos.

En la Figura 1-9 se aprecia la distribución de las EMA's operadas por el SMN, que representan un conjunto de alrededor de 190 estaciones [89].



Figura 1-9: Estaciones meteorológicas automáticas, SMN. Fuente: [89].

- **Comportamiento espacial de la precipitación**

Como fenómeno natural, la precipitación tiene una importancia fundamental para el desarrollo de la vida en la Tierra, de manera que en el transcurso de la historia se han llevado a cabo diversas actividades orientadas a eficientar su captación, distribución y consumo, bajo criterios científicos y de carácter social, por lo que el análisis de su comportamiento como parte del ciclo hidrológico es indispensable para mejorar el aprovechamiento del agua.

*Se considera que el conocimiento sobre la variación de las precipitaciones resulta esencial para una mejor administración de los recursos hídricos, así como para el diseño y operación adecuados de la infraestructura necesaria para el desarrollo comunitario. Por lo tanto, la densidad y distribución de los sitios de observación (red pluviométrica y pluviográfica) son muy significativas para perfeccionar la caracterización del fenómeno, actividad que en sí misma representa un gran desafío, dada la variabilidad espacial y temporal de la precipitación.*

La representatividad de una observación es el grado de exactitud con el que se describe el valor de la variable, necesaria para una finalidad específica, es decir, no constituye una calidad fija de la observación, sino que se deriva de una evaluación conjunta de la instrumentación, el intervalo de medición y la exposición, teniendo en cuenta las necesidades de una aplicación en particular. Concretamente, las aplicaciones tienen sus propias escalas temporales y espaciales en las

estimaciones de valores medios, así como al trabajar con la densidad de las estaciones y con la resolución de los fenómenos [64].

Las necesidades de datos observados siempre dependen del fin que se persiga, y cambian con el tiempo al mejorar las técnicas, pues en general las exigencias son cada vez mayores al aumentar la capacidad de las computadoras, y porque los modelos utilizados, como los de pronóstico del tiempo, son capaces de representar mejor los fenómenos de escala local [8].

En el caso de la atmósfera, es común la coexistencia de varios fenómenos meteorológicos de distintas escalas; por ejemplo, un núcleo tormentoso puede extenderse unos pocos kilómetros en escala horizontal con duración de varias horas, y un ciclón tropical puede tener 1 000 km de longitud horizontal, con una duración de 10 días o más, mientras muchos núcleos tormentosos pueden aparecer y desaparecer en el tiempo que dura un ciclón. Así, dada la interacción dinámica que existe entre los fenómenos meteorológicos de distintas escalas, quizás no sea posible especificar definitivamente las necesidades de cada una de dichas escalas [63].

En función de lo anterior, *la frecuencia y separación de las observaciones debe ser apropiada para obtener datos que describan los cambios en el tiempo y en el espacio del fenómeno meteorológico, con suficiente resolución para satisfacer las necesidades de los usuarios, es decir, que los fenómenos sean detectados oportunamente.* Por ejemplo, las observaciones sinópticas normalmente deberían ser representativas de un área de hasta 100 km en torno a la estación, mientras que, en el caso de aplicaciones de pequeña escala o local, la dimensión del área a tener en cuenta puede ser de hasta 10 km. La clasificación de las escalas horizontales de los fenómenos meteorológicos puede ser de la manera siguiente [64]:

- microescala (menos de 100 m) para la meteorología agrícola; por ejemplo, evaporación;
- topoescala o escala local (100 m a 3 km); por ejemplo, contaminación del aire y tornados;
- mesoescala (3 km a 100 km); por ejemplo, tormentas y brisa de mar y de montaña;
- gran escala (100 km a 3 000 km); por ejemplo, frentes, ciclones y formaciones de nubes;
- escala planetaria (más de 3 000 km); por ejemplo, ondas largas en la troposfera superior.

Las escalas horizontales están estrechamente relacionadas con las escalas de tiempo de los fenómenos, por lo que cuanto más grandes sean las perturbaciones en la escala horizontal, mayor será la probabilidad de que su período de duración sea más prolongado. En consecuencia, la predicción meteorológica a corto plazo requiere observaciones más frecuentes, procedentes de una red más densa en un área limitada, con el fin de detectar cualquier fenómeno de pequeña escala y su rápido desarrollo. A medida que aumenta la longitud del período previsto, se incrementa también el área de la que se necesitan observaciones [63].

En particular, el pronóstico del tiempo se basa en análisis meteorológicos robustos. Todos los análisis requieren datos de observación altamente fiables que se reciben de forma periódica en centros de análisis, desde una red suficientemente densa u otra fuente de observación. En el caso de los análisis meteorológicos, la exactitud, la resolución temporal y espacial y la oportunidad requeridas de estos datos dependen de los factores siguientes [63]:

- las diferentes escalas de los fenómenos meteorológicos que se deseen analizar,
- la resolución y otras características de las técnicas utilizadas para realizar los análisis y los modelos basados en estos.

En el entendido de que una red de datos hidrológicos es un conjunto de actividades de recopilación de datos diseñadas y utilizadas para cumplir una serie de objetivos compatibles, en las primeras etapas del desarrollo de una red hidrológica se deben establecer las características de una red mínima, con base en la experiencia de los organismos hidrológicos, de manera que la red se integre por el número mínimo de estaciones de registro de lluvia necesarias para evitar deficiencias graves en el desarrollo y gestión de los recursos hídricos [62].

A través de los años se han realizado esfuerzos para proponer recomendaciones de los valores de densidad mínima respecto de diversos tipos de estaciones hidrológicas, para zonas climáticas y geográficas diferentes. De acuerdo con la Organización Meteorológica Mundial (OMM), un criterio simple y conveniente para clasificar las zonas estaría basado en la variación superficial y estacional de las precipitaciones de lluvia [19]. Asimismo, la OMM ha considerado que la densidad de la población es un aspecto importante para tomarse en cuenta en el diseño de una red, pues resulta casi imposible la instalación y operación satisfactoria de varias estaciones localizadas en zonas escasamente pobladas (las que suelen coincidir con regiones de características climáticas extremas, como regiones áridas, polares o de bosques tropicales), a menos que estén muy automatizadas [62].

Por el contrario, las áreas urbanas densamente pobladas (como es el caso de la Ciudad de México) necesitan una red de registros muy densa, que proporcione una resolución temporal y espacial adecuada de las tormentas, de gran utilidad para el diseño, gestión y operación de los sistemas de drenaje en tiempo real, por ejemplo.

La Tabla 1.6 se refiere a los valores mínimos recomendados de densidad de estaciones por superficie de una red hidrológica mínima, y se construyó en consideración del desconocimiento general de la variación espacial y temporal de la lluvia en gran parte de los países con deficiencia de registros. Entonces, se optó por definir un conjunto representativo de zonas extensas con diferentes características fisiográficas, a las que se les asignaron las normas de densidad señaladas. En la Tabla se indican las magnitudes de una superficie en kilómetros cuadrados a la que le corresponde la instalación de una estación pluviométrica (no registradora) y pluviográfica (registradora), donde se recomienda que en las redes en fase de desarrollo es aconsejable contar al menos con un 10 por ciento de pluviógrafos respecto del número de pluviómetros [62].

Tabla 1.6: Densidad de estaciones mínima recomendada (superficie, en km<sup>2</sup> por estación), [62].

Unidad fisiográfica	Precipitación	
	No registradoras	Registradoras
Costa	900	9 000
Montaña	250	2 500
Planicie interior	575	5 750
Montes/ondulaciones	575	5 750
Islas pequeñas (menores a 500 km <sup>2</sup> )	25	250
Áreas urbanas	-	10 a 20
Polos/tierras áridas	10 000	100 000

Respecto al número escaso de estaciones de una red mínima, se ha observado que al respetar ciertos principios de instalación y manejo aún es posible cumplir con las necesidades más inmediatas. En general, los medidores de precipitación deben distribuirse tan uniformemente como corresponda a las necesidades prácticas de datos, incorporándose tres tipos de medidores: estándar, empleado para medir cantidades con periodicidad diaria; pluviógrafos, especialmente en las áreas que reciben lluvias intensas y de corta duración, con prioridad de áreas urbanas, cuencas fluviales, áreas abarcadas inadecuadamente por redes existentes y proyectos de investigación;

medidores de precipitación acumulada, para regiones escasamente pobladas o apartadas, con lecturas de una periodicidad amplia que permita la inspección de las estaciones. Adicionalmente, para disponer de datos de lluvia que permitan ampliar o revisar la calidad de registros de escurrimiento fluvial, los medidores de precipitación deben situándose en las inmediaciones del aforador y en la parte superior de la cuenca de drenaje sujeta a observación, de modo que la precipitación pueda estimarse para cada estación de aforo de caudales [62].

Con relación al comportamiento de la precipitación, se dice que existen dos líneas de investigación de gran interés [11]: el análisis de las series de tiempo de registros puntuales de lluvia, con el objetivo principal de describir los datos disponibles en términos de parámetros estadísticos (con variaciones entre diferentes regiones geográficas), para la estimación de los periodos de retorno correspondientes a los registros individuales; por otro lado, el análisis de datos de lluvia registrados simultáneamente en un conjunto de sitios de medición dentro de un área de estudio, con la intención de realizar estimaciones medias representativas de áreas y de eventos de interés (dada la no uniformidad de la distribución espacial de la lluvia). En ambos casos, el análisis con disponibilidad de datos en diferentes escalas de espacio y tiempo ha sido objeto de estudio.

El conocer la variación de la magnitud de la lluvia y la proporción en que tiende a distribuirse sobre una superficie a lo largo de la duración total de un evento, resulta de gran interés para el manejo y control del agua, en especial dentro de cuencas con una disponibilidad limitada o nula de registros de escurrimiento [65].

Por lo general, para el análisis y solución de problemas se han considerado las mediciones puntuales registradas dentro o en la proximidad de las áreas de estudio, y aunque dicha información pueda considerarse confiable, los valores son representativos únicamente sobre la vecindad de los puntos de medición, y no reflejan la naturaleza del fenómeno, que entendido como la etapa terminal de un conjunto de procesos que ocurren a diferentes escalas, es variable sobre una extensión de la superficie de captación, y en un determinado intervalo de tiempo [19].

Los errores en la estimación de la lluvia sobre un área, considerando los registros de una red pluviométrica, tienen su origen en la naturaleza aleatoria de las tormentas y su paso entre los puntos de medición, aspectos relacionados con las condiciones del terreno mismo, por ejemplo, resulta conveniente instalar una red más densa dentro de una región montañosa y con lluvias convectivas, respecto de una zona plana y con eventos ciclónicos [19]. Además, al distribuir la lluvia utilizando los valores máximos determinados en los puntos de medición, se tiende a sobrestimar su magnitud al suponer que los máximos se presentan simultáneamente en todas las estaciones consideradas, lo cual es poco probable, en especial conforme la superficie de análisis es mayor [65].

Los inconvenientes relacionados con la falta de información sobre la distribución espacial de la lluvia han quedado de manifiesto especialmente con la introducción de técnicas de modelado computacional en la aplicación práctica de la Hidrología; por ejemplo, para la consideración de la lluvia en los modelos de simulación de escurrimientos [11].

Dado que la dificultad principal en la estimación de la lluvia media (areal) representativa de una superficie está en la interpolación o extrapolación de las mediciones puntuales, si se toma en cuenta la física del fenómeno respecto a las limitaciones de escala propias del campo de lluvia registrado [11], una consideración ampliamente utilizada desde la segunda mitad del siglo pasado es el empleo de relaciones empíricas conocidas como Factores de Reducción por Área (FRA), definidas a partir de la relación que existe entre la lluvia media máxima anual asociada a un

periodo de retorno, con la lluvia media de los máximos puntuales asociados al mismo periodo de retorno, los que corresponden necesariamente a regiones homogéneas, pues toda el área de una región seleccionada para el cálculo de una relación FRA debe estar expuesta a la incidencia de los mismos fenómenos meteorológicos [54]. Al proceder así, la lluvia media sobre un área es calculada como el producto entre la lluvia registrada en una estación y el FRA, por lo que se procura la atenuación de los valores puntuales observados (máximos). Cabe señalar que algunos investigadores han enfatizado que la reducción de la lluvia puntual debe realizarse cuando el área de captación exceda los 25 km<sup>2</sup> [11]. En estos términos es como se ha procedido a la obtención de factores propios de distintas regiones climáticas, con el objetivo de poder aplicarlos sobre cuencas de áreas y climatologías similares, y obtener lluvias aproximadas [54].

Por otra parte, debe mencionarse que el concepto de FRA está basado en la consideración de la lluvia como un fenómeno estático, por lo que su utilidad se reduce cuando se toman en cuenta las características cinemáticas de un evento con relación al área de su captación. Así, por ejemplo, al suponer un campo de lluvia de intensidad uniforme que se mueve con velocidad constante sobre una cuenca unidimensional, el patrón de la lluvia observado o hietograma de un registro puntual será de intensidad constante en toda su duración, misma que dependerá de la magnitud y velocidad del campo de lluvia, y ésta será igual en cualquier punto de la cuenca, mientras que el hietograma de lluvia areal será trapezoidal, con duración superior, pero de menor intensidad [11].

Un campo donde las propiedades espaciales y cinemáticas de la lluvia tienen una importancia evidente es el del manejo en tiempo real de sistemas de drenaje, o el del cálculo de avenidas para la prevención de inundaciones, los que suelen requerir de pronósticos de lluvia a corto plazo, como entradas para la simulación de escurrimientos. Bajo estas condiciones, *las técnicas tradicionales del procesamiento de datos no resultan muy apropiadas, por lo que se han empleado modelos que utilizan una descripción estocástica de los procesos de lluvia, pero también se ha visto que el control de las operaciones puede llegar a ser más efectivo al utilizar herramientas de percepción remota*, aunque se ha observado que en términos cuantitativos, la calidad de estimaciones de un radar todavía es superada por una red de estaciones pluviométricas densa [11].

Debido a que aún puede decirse que muchos de los métodos de cálculo hidrológico aplicados de manera práctica están basados en consideraciones simplificadas, pues quienes los ejercen profesionalmente suelen aceptar nuevos métodos únicamente cuando estos mejoran la solución de inconvenientes comunes, o les satisfacen nuevos objetivos, resulta conveniente fortalecer el contacto interdisciplinario de distintas ramas del conocimiento, como son la Hidrología, la Hidráulica, la Meteorología y la Estadística, en la búsqueda de soluciones más aproximadas a complicaciones compartidas, como es la variabilidad de la lluvia [11].

Las líneas de investigación así trazadas permiten el estudio de los procesos de lluvia, contemplando amplios rangos de condiciones fisiográficas y de climatología, con el uso de radares meteorológicos o de satélites para escalas más grandes, de manera que sea posible *obtener mejores estimaciones de la distribución espacial y temporal de la lluvia, en apego a los mecanismos físicos de este fenómeno natural*.

### 1.2.3. Precipitación estimada con percepción remota

- **El radar meteorológico**

Se trata de un sensor de observación atmosférica útil para detectar la presencia de agua en estado líquido o sólido, por lo que facilita el rastreo de tormentas por su capacidad para determinar la

distancia a la que está ocurriendo una precipitación, además de una estimación cuantitativa de la lluvia, de manera discreta en el tiempo y continua en el espacio [84].

La capacidad de medir los campos de precipitación sobre amplias extensiones a partir de un punto se conoce como técnica de teledetección activa, y se sabe que el alcance efectivo de un radar suele ser de 40 a 200 km, en función de las características del equipo, como por ejemplo su haz de antena, su potencia de salida o su sensibilidad de recepción [62].

Para cumplir con su función, el radar meteorológico en espacio abierto envía pulsos electromagnéticos a través de una antena, y en consecuencia recibe el reflejo de los objetos alcanzados, lo que le permite ubicar la posición (distancia) de estos a través de una medición del tiempo transcurrido entre la emisión del pulso y la recepción de su eco (señal reflejada), con una velocidad próxima a la de la luz [68].

Este instrumento está constituido por un conjunto de dispositivos, siendo los principales: un transmisor de pulsos de alta potencia; una guía de onda o conducto para dirigir los pulsos electromagnéticos del transmisor al foco de la antena, y los pulsos recibidos de la antena al receptor; un receptor que detecta, amplifica y transforma la señal recibida, y la envía a la pantalla de despliegue; un procesador de la señal recibida; una consola de despliegue y control; una antena que gira en acimut y en ángulo de elevación, y un radomo casi esférico para proteger la antena contra cargas asimétricas por viento, humedecimiento de su superficie e intemperización [56].

Durante el proceso de teledetección, el haz del radar puede chocar con obstáculos del terreno, produciéndose ecos intensos y persistentes denominados ecos de tierra (clutter), que podrían interpretarse erróneamente como lluvia, aunque pueden eliminarse con la supresión de mapas de eco por medio del procesamiento Doppler [83].

Al contar con capacidad Doppler, el radar puede realizar la estimación de la velocidad y dirección del viento, lo que permite detectar si el fenómeno meteorológico se está acercando o se está alejando del radar. Por otro lado, al estar provistos de doble polaridad, los pulsos u ondas de muy corta duración, que son emitidos a través de una antena parabólica en una determinada dirección, se transmiten en forma simultánea con las polarizaciones vertical y horizontal, por lo que el radar logra medir el tamaño de las gotas o granizos [69], lo que facilita la clasificación de la precipitación pluvial (lluvia, llovizna, granizo, nieve, etc.), además de que permite disminuir la atenuación de la señal por problemas de potencia en fenómenos intensos, y lograr datos más aproximados del tipo de partículas, volumen y distribución espacial [72].

Por otra parte, cuando el haz del radar pierde energía como consecuencia de los efectos de absorción y dispersión que ejercen las gotas de agua y los gases de la atmósfera, la estimación de lluvia se ve afectada, lo que se conoce como atenuación. De esta manera, la atenuación es la pérdida de potencia del haz de radiación electromagnética emitida por el radar al atravesar el medio en el que viaja, y depende de la longitud de onda de la señal que se está propagando en el espacio, por lo que cuanto mayor sea ésta (y menor sea su frecuencia), menor será también la atenuación. De acuerdo con la OMM, la atenuación ocurre por la existencia de gas de la atmósfera (oxígeno molecular y vapor de agua), nubes y precipitaciones, tomando en cuenta que los gases tienen efectos de absorción, mientras que las nubes y las gotas se comportan como elementos de absorción y dispersión [83].

Las longitudes de onda de uso operativo del radar son tres, y la selección de la más apropiada depende de las condiciones climáticas y la finalidad perseguida [83]:

- Banda S, con longitud de onda de 10 cm, resulta adecuada para el análisis de zonas de mayor precipitación, al mostrar efectos menos desfavorables respecto a la atenuación, aunque los radares de onda larga tienen menor capacidad para detectar llovizna y nieve.
- Banda C, de longitud de onda de 5.6 cm, es utilizada en la mayor parte de los radares colocados en zonas no tropicales, aunque pueden requerir la instalación de sistemas adicionales para corregir las características de las señales atenuadas.
- Banda X, con una longitud de onda de 3 cm, es requerida usualmente en radares de onda corta destinados a la práctica de la hidrología urbana.

La detección y caracterización de los hidrometeoros se lleva a cabo mediante la reflexión que sufren las microondas que emite el radar sobre los blancos bajo estudio, en función del tiempo que requiere el viaje de ida y regreso de la señal de las mediciones que se realizan [68]. El factor de reflectividad ( $Z$ ) está relacionado con la energía medida en forma de potencia, ligado a la distribución de los diámetros de las gotas contenidas en un cierto volumen, donde el radar mide la potencia reflejada por las partículas contenidas en dicho volumen, y que a su vez se puede relacionar con la intensidad de lluvia ( $R$ ), [56].

Para determinar la intensidad de lluvia por medio del radar, lo más común es utilizar ecuaciones empíricas que relacionan la reflectividad registrada por el radar con la precipitación medida al suelo [56], es decir, la relación  $Z - R$  que se utiliza para transformar la reflectividad ( $\text{mm}^6/\text{m}^3$ ) en intensidad de lluvia ( $\text{mm}/\text{h}$ ), [72].

Con fines de estimación de la lluvia, es necesaria la calibración de un modelo matemático de la relación que se considera existe entre la potencia retornada o reflectividad, que el radar registra en el espacio durante un evento, respecto a la intensidad de precipitación ocurrida simultáneamente sobre la superficie de captación [56].

De entre las expresiones más utilizadas para analizar eventos de lluvia estratiforme se tiene la propuesta por Marshall y Palmer en 1948, quienes caracterizaron la distribución de tamaños de gotas dentro de un volumen de muestreo como una función simple de la intensidad de la lluvia. En México, Méndez (2005) propuso expresiones considerando un análisis climático de la Ciudad de México, utilizando mediciones del radar Cerro Catedral, que registra información propia de la cuenca del Valle de México, y concluyó que la naturaleza convectiva de las tormentas predomina sobre la estratiforme, y no es necesaria su clasificación mediante un umbral.

Dos de los productos de imágenes digitales de reflectividad que se obtienen del radar son el Indicador de Posición en Planta y el PPI de Altitud Constante (PPI y CAPPI, por sus siglas en inglés), los que comprenden radios de información de hasta 300 km, en resoluciones espaciales de 1 a 2  $\text{km}^2$  (incluso inferiores), y resoluciones temporales de entre 5 y 15 minutos [83].

El producto más común del radar meteorológico es el PPI, siendo una representación en formato polar de una variable, obtenida a partir de una sola rotación completa de antena, que se realiza en un ángulo de elevación constante. Por su parte, el producto CAPPI es la representación de una sección transversal horizontal de una variable a una altitud determinada, obtenida por interpolación a partir de los datos de volumen [83].

Para disponer de una estimación de la lluvia diaria sobre la base de la información obtenida con radares meteorológicos, se requiere la aplicación de algún procedimiento que conste de pasos ordenados para la corrección de los datos, la generación y visualización de la cartografía georeferenciada, así como la organización de la estructura de los datos correspondiente.

En 2011, Vilchis *et al.* presentaron una metodología para estimar precipitación diaria a través de una herramienta geomática, considerando imágenes PPI del radar meteorológico Cerro Catedral, desarrollada como un proceso sistemático en diez pasos, descrito brevemente a continuación.

Como primer paso deben obtenerse y almacenarse las imágenes de radar, por lo que es necesaria la transferencia de archivos crudos PPI (Plan Position Indicator) a un Sistema de Información Geográfica (SIG).

El segundo paso consiste en la integración de las imágenes PPI, indicando el tipo de archivo, las coordenadas extremas del radio de influencia del radar, así como el sistema de referencia geográfico. Como la información original presenta un giro de 180°, debido a la dirección de escaneo del radar, en el tercer paso se requiere invertir el orden de los renglones de la imagen.

En el cuarto paso se realiza la conversión de valores de la imagen PPI, que son cantidades  $N$  en un rango de 0 a 255, y se transforman en un valor de decibelios de  $Z$  (dBZ) mediante la aplicación del modelo propuesto por Environment Canada:  $\text{dBZ} = (0.5N) - 32$ .

Como quinto paso se reducen los efectos de atenuación del haz emitido por el radar debido a lluvia sobre el radomo (metodología de Sempere-Torres), al aplicar un factor de corrección de +3dBZ de reflectividad.

En el sexto paso se realiza el proceso de caracterización de la precipitación, donde se identifican las zonas convectivas y estratiformes dentro de la imagen del radar, aplicando el algoritmo propuesto por Steiner: zonas convectivas  $\geq 40\text{dBZ}$ , y zonas estratiformes  $< 40\text{dBZ}$ .

En el séptimo paso se realiza la conversión de valores dBZ a reflectividad:  $Z = 10^{(\text{dBZ}/10)}$ , en  $\text{mm}^6/\text{m}^3$ .

En el octavo paso se obtienen las estimaciones de intensidad de lluvia  $R$  a partir de la reflectividad  $Z$ , utilizando relaciones empíricas de la forma  $Z = aR^b$ . Para las zonas estratiformes se utilizó el modelo Marshall-Palmer:  $Z = 200R^{1.6}$ , y para las zonas convectivas el modelo NEXRAD:  $Z = 300R^{1.4}$ .

En el noveno paso se generan los productos acumulados, y para ello se realiza una adición de imágenes de radar consecutivas (cada 15 minutos) para un día (24 horas).

Como paso final, a los productos acumulados se les corrigen los errores generados por intercepción del haz con el terreno, y están sujetos a las características orográficas de cada región, por ejemplo:

- los ecos del suelo son aquellas zonas donde el haz intercepta el terreno, y se requiere de una sustitución de zonas contaminadas por interpolación a partir de los datos no contaminados registrados en el entorno del eco a sustituir;
- las zonas apantalladas son aquellas que se encuentran situadas detrás de un eco de suelo, y donde la lluvia es minimizada a causa de la energía perdida en la intercepción;
- los lóbulos secundarios (vecinos del lóbulo principal que describe la distribución de energía alrededor del eje del haz) son causados por la intercepción con el terreno por parte de cantidades de energía no incluidas en el haz principal del radar;
- la eliminación de ruido residual, que excluye zonas de la imagen de tamaño reducido y baja intensidad, menores a 7 dBZ (0.1 mm/h), que generalmente no están asociadas a la precipitación.

Con la intención de dar seguimiento a las tormentas en el territorio nacional, desde el año 1976 el Servicio Meteorológico Nacional dispone de radares meteorológicos de banda C. Se optó por cinco radares Enterprise Electronics Corp. (EEC), instalados en los estados de Tamaulipas, Sinaloa, Durango, Guerrero y Baja California Sur. Entre 1993 y 1996, dichos radares se reacondicionaron, y también se adquirieron siete radares Ericsson para instalarse en los estados de Sonora, Colima, Oaxaca, Quintana Roo, Veracruz, Campeche y Estado de México [85].

El conocimiento sobre la cobertura detallada de los radares meteorológicos es importante para aplicaciones del manejo hidrológico de cuencas, por lo que es conveniente revisar la cobertura particular de cada radar, de manera que se tomen en cuenta los obstáculos orográficos que pueden afectar la estimación de precipitación de cada instrumento.

Rosengaus realizó cálculos de cobertura de radares utilizando los sistemas AUTOTOPO y COBRADAR, que están basados en una simplificación de la topografía digital del terreno del INEGI [90]. De esta forma obtuvo diagramas de cobertura para PPI correspondientes a diferentes altitudes, como la mostrada en la Figura 1-10, donde se aprecian las regiones con cobertura satisfactoria para 3 000 msnm, obtenidas para un alcance máximo de 300 km, condición que parece suficiente si se toma en cuenta que en general los accidentes geográficos que caracterizan el relieve de México no superan los 3 000 msnm, como se observa en la Figura 1-11, [24].

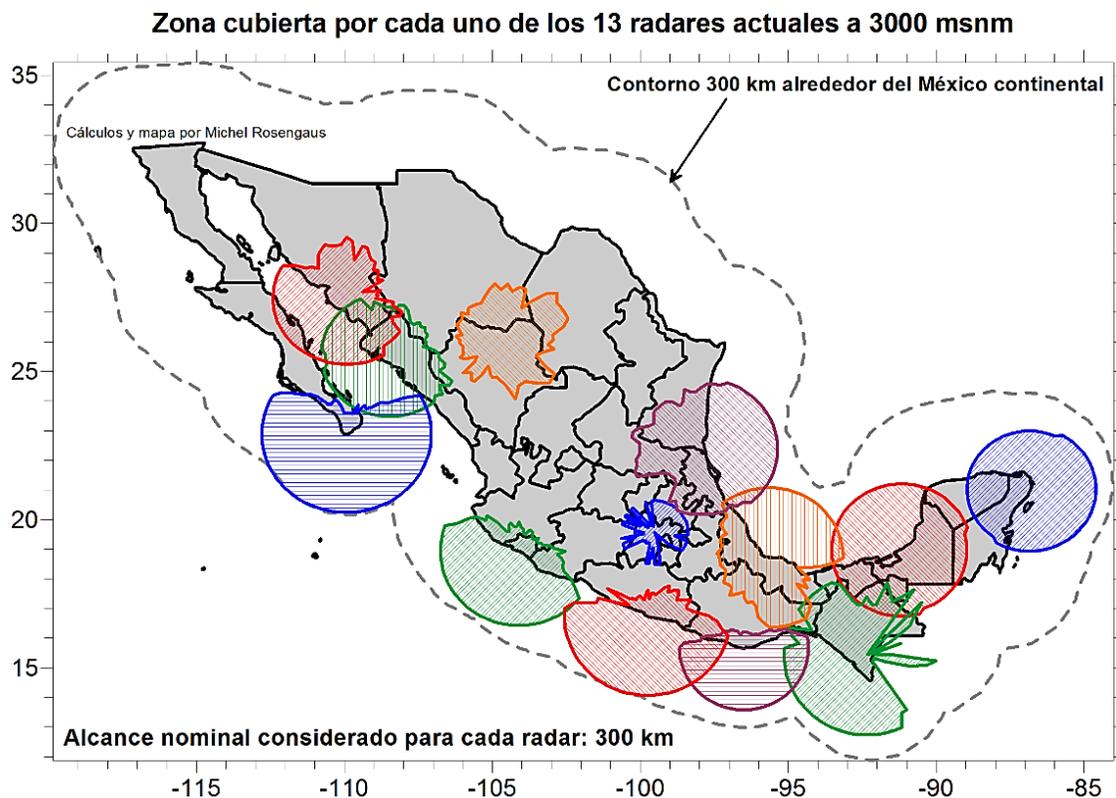


Figura 1-10: Cobertura de cada uno de los trece radares de la red del SMN, a 3 000 msnm. Fuente: [90].

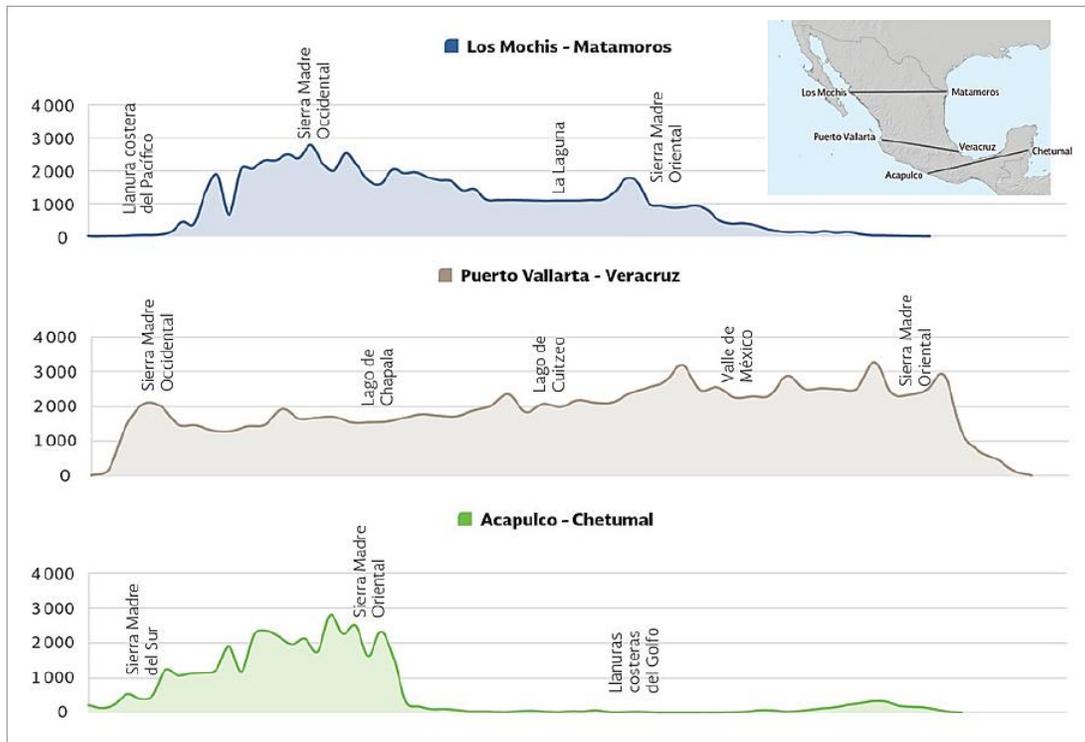


Figura 1-11: Perfiles de elevación representativos de la República Mexicana, en msnm. Fuente: [24].

Para disponer de los productos crudos que se obtienen de los radares meteorológicos en México, particularmente los archivos de imágenes PPI que representan la reflectividad, es necesario establecer contacto de manera electrónica con la Ventanilla Única de la Coordinación General del SMN, [89]. Como respuesta a una solicitud realizada durante el desarrollo de este trabajo de investigación, el Servicio informó respecto a los periodos de registro de reflectividad disponibles de doce de los radares que administra, presentados en la Tabla 1.7.

Tabla 1.7: Periodos de registro de los radares meteorológicos administrados por el SMN, [89].

Radar	Ubicación		Periodo
	Latitud	Longitud	
Acapulco, Gro.	16°45'38"	99°45'	dic 2010 - ago 2020
Altamira, Tamps.	22°23'15"	97°55'31"	dic 2009 - ene 2016
Alvarado, Ver.	18°43'01"	95°37'06"	dic 2007 - ago 2016
Cancún, Q. Roo	21°1'4"	86°51'14"	dic 2009 - feb 2020
Cerro Catedral, Edo. Méx.	19°33'19"	99°31'12"	dic 2007 - dic 2019
Cuytlán, Col.	19°01'35"	104°19'54"	dic 2007 - jun 2010
Guasave, Sin.	25°34'9"	108°27'47"	dic 2011 - sep 2017
Los Cabos, B.C.S.	22°52'44"	109°55'37"	jul 2012 - mar 2016
Obregón, Son.	27°28'01"	109°55'01"	dic 2007 - feb 2018
Palmito, Dgo.	25°45'43"	104°54'18"	jul 2012 - jul 2015
Puerto Ángel, Oax.	15°39'28"	96°30'03"	ene 2008 - dic 2019
Sabancuy, Camp.	18°56'37"	91°10'21"	sep 2011 - dic 2019

Como se aprecia, la cantidad de registros disponibles es muy limitada. De acuerdo con el SMN, la escasez de datos es consecuencia de varios factores, por ejemplo, la capacidad de almacenamiento para el resguardo de la información de archivos crudos PPI tiene un límite, y no existe interés de conservarlos de manera permanente, por lo que no es posible conocer datos

correspondientes a fechas anteriores a diciembre de 2007; por otro lado, la inversión que representa el mantenimiento, rehabilitación o renovación de equipos para el fortalecimiento de la infraestructura de la red nacional de radares es considerable, por lo que históricamente el funcionamiento de los equipos ha sido intermitente y la continuidad de los registros ha sido interrumpida de manera recurrente. Además, como la generación de mayores beneficios a partir del empleo de estas herramientas requiere de la formación, capacitación y contratación de personal aún más especializado [76], la inversión insuficiente ha ocasionado que la información meteorológica producida sea subutilizada.

- **El satélite**

Los satélites artificiales de observación meteorológica son plataformas de masa pequeña que, por localizarse dentro de la órbita terrestre media, permiten una visión global del tiempo atmosférico, proporcionando información de regiones inaccesibles para otros sistemas de observación, muchas veces limitados por aspectos de topografía (registros de pluviómetros) o de accesibilidad y disponibilidad de las bases de datos (de radares). De esta manera, los satélites resultan de gran utilidad para la investigación del comportamiento de algunas variables climáticas, especialmente por la posibilidad de disponer de la información en tiempo real [44].

La observación meteorológica de los satélites se practica por teledetección o percepción remota, la cual se efectúa con sensores instalados en las plataformas espaciales que se encargan de detectar radiación electromagnética emitida o reflejada, y la registran como un valor físico que puede ser grabado o procesado [32].

Entre los campos de aplicación se encuentra la determinación de la altura de las nubes, su temperatura, dimensiones, trayectorias, cambios de forma, y la dinámica de las nubes potencialmente precipitables y áreas cubiertas, lo que hace factible relacionar dichas características con la intensidad instantánea de las precipitaciones. Adicionalmente, permiten determinar la velocidad de los vientos, los perfiles de temperatura, el contenido de vapor de agua en la atmósfera, la extensión del manto de nieve, tipos de vegetación, así como la evapotranspiración, humedad y uso del suelo [85].

De acuerdo con su órbita, los satélites se clasifican en geoestacionarios y de órbita polar, la que se define por la capacidad y el propósito del satélite, que puede variar en términos de la altura, orientación y rotación con respecto a la Tierra, o en función del tiempo que tarda el satélite en realizar su recorrido completo, designado como periodo orbital [44].

En las órbitas geoestacionarias los satélites giran cerca del plano ecuatorial, dando seguimiento a la Tierra en su movimiento de rotación, y se les denomina de órbita geosíncrona si la inclinación orbital es de 0° respecto al plano del Ecuador [44]. La resolución temporal del satélite está sujeta a la resolución temporal del sensor, por lo que sólo depende del tiempo que éste requiere para generar una imagen completa; por el contrario, una desventaja que tienen es su resolución espacial, relativamente baja en comparación de los satélites polares, por encontrarse alejados de la Tierra. En zonas próximas al Ecuador su resolución es máxima, entre tres y cinco kilómetros, pero disminuye al aproximarse a los polos, y entonces para disponer de una cobertura global del planeta se necesitan entre seis y siete satélites [85].

En el caso de satélites de órbita polar, su plano de órbita es casi un plano meridiano, por lo que giran en torno a la Tierra de polo a polo, y se les conoce como de órbita baja por ubicarse a una

distancia próxima de 850 km sobre la superficie, aunque realmente el plano de estos satélites tiene una ligera inclinación respecto al eje de rotación del planeta, porque éste no es esférico [85].

La principal característica de las órbitas sincrónicas es que el satélite pasa sobre un mismo punto de la Tierra siempre a la misma hora, con un periodo entre observaciones de exactamente doce horas, lo que implica una desventaja en cuanto a su resolución temporal (doce horas entre imágenes de un mismo punto), que resulta insuficiente para objetivos meteorológicos porque la evolución de los fenómenos puede ser rápida, lo que evidencia la gran utilidad de los satélites geoestacionarios al captar imágenes cada 15 ó 30 minutos, permitiendo la vigilancia de los sistemas nubosos de una región de interés, por ejemplo [85].

Cabe señalar que los satélites geoestacionarios y de órbitas polares sólo suministran información de la parte superior o tope de las nubes, porque la estimación de la precipitación de lluvia desde el espacio está basada en la medición de la cantidad de radiación reflejada y emitida a través de las cimas de las nubes. Es decir, como la mayor parte de la radiación no penetra profundamente en las regiones nubosas que contienen partículas de tamaño similar o superior a la longitud de onda de la radiación, entonces, con excepción de las longitudes de onda más largas, la mayor parte de la radiación reflejada proviene de las regiones superiores de las nubes precipitantes, y sólo es posible establecer relaciones indirectas respecto a la generación de la lluvia [62].

Existen dos técnicas de teledetección [62]: activa, cuando se emite un haz de radiación artificial hacia el objetivo y se analiza la respuesta de éste, y pasiva cuando se detecta la radiación natural reflejada por el objeto o el área circundante desde la superficie terrestre, y en este caso es la luz solar la más observada en las imágenes infrarrojas. Aunque los satélites de tipo activo son menos numerosos, sus aportaciones son un complemento que enriquece la información.

En cuanto a las características del sensor transportado por las plataformas satelitales, éste capta información por medio de longitudes de onda para las diferentes regiones del espectro electromagnético, y cada región se denomina canal o banda, siendo tres los canales clásicos.

El primer canal es el de las imágenes en canal visible (VIS), donde las reflectividades de los distintos objetos observados son notorias porque cada uno se refleja de manera diferente con la luz que incide sobre su superficie. Así, las superficies se observarán más claras u oscuras en función de la radiación reflejada, por ejemplo, la reflectividad que presenta el mar es muy baja, y en consecuencia la representación de su superficie es prácticamente negra; de manera opuesta, la reflectividad de las nubes es muy alta, y a mayor espesor (cúmulos y cumulonimbos) se apreciarán más blancas. Este primer canal se distingue por su sensibilidad a la radiación solar directa, por lo que dependerá de la estación del año y la hora del día. Bajo estas condiciones, las imágenes en el espectro visible únicamente están disponibles cuando la superficie reflectante tiene luz diurna, por lo que durante la noche permanecen oscuras [85].

El segundo canal contiene el sensor térmico de la banda del espectro infrarrojo (IR), y al igual que los de canal visible, se encuentra principalmente a bordo de los satélites de la órbita geoestacionaria. La radiación que emiten los cuerpos es detectada por dicho sensor, y al estar relacionada con su temperatura, las imágenes infrarrojas obtenidas son una representación de las condiciones térmicas del planeta, lo que incluye el tope de las nubes. De esta manera, los cuerpos más calientes (más oscuros) se relacionan con una radiación de mayor magnitud, y los cuerpos fríos (más brillantes) con radiaciones más bajas, y entonces las tonalidades de las nubes permiten identificar capas a diferentes alturas. En general, la precipitación estimada con datos IR es mejor sobre los trópicos, donde gran cantidad de eventos se originan de sistemas convectivos, pero su

calidad es menor en latitudes medias, con precipitaciones derivadas de nubes estratiformes. Como ventaja, la radiación emitida en el rango infrarrojo es captada permanentemente, y por ello la información está disponible en cualquier momento [85].

El tercer canal está focalizado en la zona espectral donde se produce mayor absorción de vapor de agua (WV), para medir la radiación que llega al sensor del satélite proveniente del sistema tierra-atmosfera. El vapor de agua contenido en la atmosfera absorbe parte de esta radiación, de tal forma que las intensidades de radiación que mide el sensor nos indican la distribución de vapor de agua presente. Debido a la capacidad de absorción que tiene el vapor de agua en esta banda, la radiación emitida por la superficie y por las capas bajas puede no alcanzar al satélite, y la distribución de humedad definida corresponder a las capas medias y altas de la troposfera. Como sucede con la imagen IR, las zonas oscuras se relacionan con temperaturas más altas (que suelen corresponder a regiones de actividad convectiva), y las zonas brillantes con temperaturas más frías [85].

Dada la gran cantidad de información y de algoritmos disponibles para estimar la precipitación con registros satelitales de diversas resoluciones espaciales y temporales, y la evidencia de que las estimaciones generadas con una sola plataforma se caracterizan por una resolución que puede ser insuficiente, para obtener más provecho de las fortalezas de los productos satelitales, en la actualidad existe una variedad de estimaciones de precipitación generadas como productos de algoritmos que posibilitan la combinación de datos registrados por diferentes tipos de sensores, a través de esquemas de calibración [44].

Los productos de análisis de precipitación multi-satelital, TMPA, de la Tropical Rainfall Measuring Mission (TRMM), han estado disponibles como mediciones de las zonas tropicales y subtropicales (50°N – 50°S) desde noviembre de 1997. De esta manera, el TRMM proporcionaba los productos 3B42 en tiempo real (3B42RT), así como el 3B42 versión 7 pre-calibrada (3B42V7), los cuales han sido aplicados extensivamente en simulaciones y predicciones hidrológicas. Sin embargo, se ha observado que los resultados obtenidos con dicha tecnología presentan un sesgo y errores significativos en zonas de topografía compleja, como es el caso del producto 3B42RT que tiende a sobreestimar, por ejemplo [42].

Con la intención de mejorar las resoluciones espacio-tiempo de la precipitación surgió The Global Precipitation Mission (GPM), que es una misión internacional constituida por una red de satélites para la generación de productos de medición de lluvia y nieve, que cubre una latitud de 90°N a 90°S. El GPM ofrece el producto Day-1 multi-satellite precipitation, obtenido mediante el algoritmo Integrated Multisatellite Retrievals for GPM (IMERG). En este algoritmo se toman en cuenta diferentes estimaciones multi-satelitales [42], como son: TMPA de la NASA, el cual considera la calibración entre productos satelitales de las estimaciones de precipitación, así como la combinación a escala mensual del análisis de los satélites y mediciones puntuales; CMORPH de la NOAA, que produce un análisis de precipitación global de alta resolución mediante estimaciones derivadas de observaciones de microondas de satélites de órbita baja, cuyas características se obtienen enteramente de datos IR de satélites geoestacionarios, y el PERSIANN de la Universidad de California, cuyo esquema está basado en una red neuronal para la generación de estimaciones de precipitación de microondas calibradas [85].

Como algoritmo, el IMERG se ejecuta en tres ocasiones (con productos de resoluciones temporales desde media hora, y espaciales de 0.1°): primero, cuatro horas después de ocurrida la observación, IMERG Early, por lo que la aproximación del producto es más limitada, pero su empleo es factible para el análisis de eventos en tiempo real; más tarde, después de doce horas de

la observación, IMERG Late, que es una descripción más completa, por lo que permite el análisis al día siguiente del evento; finalmente, tres y medio meses después, IMERG Final, una vez que se han recibido todos los datos, incluidos los pluviométricos mensuales, y como producto final se destina al análisis científico [42].

Para los propósitos de este trabajo de investigación se tomarán en cuenta las lluvias estimadas como productos de sistemas satelitales, en particular la información que la NASA proporciona a través de la misión satelital GPM, que permite captar la variabilidad espacial de la lluvia en latitudes altas y medias. La descarga electrónica de archivos tipo ráster se realizará a través del sitio web GIOVANNI del Centro de datos GES DISC de la NASA, donde se dispone de información de lluvia distribuida sobre una superficie y fecha de interés.

GIOVANNI (Geospatial Interactive Online Visualization And Analysis Infrastructure) es una aplicación desarrollada por el Centro de Servicios de Información y Datos de Ciencias de la Tierra de Goddard (GES DISC), que permite visualizar grandes cantidades de datos de teledetección de forma sencilla e intuitiva, aun sin la descarga electrónica de los datos, aunque también se admite la descarga de éstos como archivos en los formatos HDF, NetCDF, GeoTIFF y KMZ. La información como mapas, series de tiempo y diagramas está disponible en la aplicación, al indicar la fecha y ubicación específica del área de interés [92].

## 2. Conceptos y descripción de métodos para la asimilación de datos

### 2.1. Aspectos generales

El propósito del análisis de la información meteorológica es la caracterización del estado físico de la atmósfera de manera consistente, desde los puntos de vista espacial y temporal. El conocimiento que se tiene acerca del estado atmosférico se basa en un conjunto limitado de observaciones de la realidad física reciente, que están distribuidas irregularmente en el espacio y en el tiempo, y sin la garantía de que los datos estén exentos de errores [26].

Existen diversos procedimientos para combinar las observaciones y lograr conclusiones acerca del comportamiento de las variables meteorológicas dentro de un área de interés [51]. Con ello se pretende realizar diversos análisis para conseguir la representación regular del estado atmosférico, procurando considerar las principales características de las observaciones, con un efecto mínimo de los errores de los datos. Un ejemplo de esto es el campo de la predicción numérica del clima (NWP, por sus siglas en inglés), que requiere de una definición adecuada de las condiciones iniciales de un sistema inmerso en un espacio dentro de una malla regular [35].

Es importante señalar que existen dos fuentes de información: las mediciones (observaciones) y la comprensión de la evolución temporal y espacial del sistema de interés (modelos). Sin embargo, desde el punto de vista de la información, las estimaciones a partir de las observaciones o de los modelos son similares, con algunas diferencias en el mecanismo para obtener la información [50].

El método de análisis es uno de varios factores que determinan la exactitud que puede alcanzarse a partir del conjunto de observaciones disponibles, pues la exactitud depende de la configuración de las observaciones, de su densidad, su frecuencia, el tipo de variable por analizar y de los errores inherentes a las observaciones [39].

En la primera mitad del siglo XIX comenzaron a realizarse esfuerzos por representar las observaciones meteorológicas de manera coherente, por medio de cartas o mapas. La intención original era la deducción de las leyes que gobiernan a los flujos atmosféricos, pues se consideraba a las cartas como diagnósticos del estado atmosférico del pasado y del presente, para el pronóstico de estados futuros [26].

Los primeros resultados se alcanzaron con el desarrollo de las cartas sinópticas atribuidas a LeVerrier, quien procuró realizar el diagnóstico de una tormenta severa del año 1854, y para ello se basó en observaciones realizadas sobre Europa, en un periodo previo y posterior al evento de interés. Con estas observaciones se construyeron una serie de cartas sinópticas, cada una de las cuales representaba las observaciones meteorológicas correspondientes para un tiempo definido,

de modo que la secuencia completa de las cartas constituía una serie de tiempo, con las que se determinaba el desarrollo y movimiento de las tormentas [26].

Las primeras cartas sinópticas eran construidas a mano, y la ubicación de las observaciones meteorológicas sobre los mapas se trazaba utilizando códigos especiales. De esta manera, el meteorólogo podía dibujar líneas isobaras, isotermas e isotacas; además, las cartas describían las condiciones del aire de capas superiores en elevaciones y presiones constantes, o el potencial de temperaturas sobre la superficie [26].

Con el desarrollo de técnicas de aplicación gráfica fue posible el análisis de los valores leídos de las cartas, para realizar el cálculo de operaciones básicas, junto con la diferenciación e integración de variables escalares y vectoriales. De esta manera pudieron calcularse, entre otros, el transporte a partir de temperaturas horizontales, fenómenos de vorticidad, movimientos verticales y variaciones de la presión sobre la superficie [26].

Dado que los resultados de los primeros diagnósticos del estado de la atmósfera basados en las cartas sinópticas dependían directamente del juicio de los analistas, estos procedimientos ahora son considerados como análisis subjetivos, y aun cuando estas herramientas también fueron utilizadas por meteorólogos para el desarrollo de cartas de pronóstico del clima para uno o dos días futuros, el progreso de estas capacidades en el periodo de 1860 a 1950 fue relativamente lento [26].

Como desde un punto de vista conceptual es deseable utilizar la información disponible sobre el comportamiento de un sistema para extender las observaciones (mediciones que son discretas en el espacio y el tiempo) y llenar los espacios vacíos de manera objetiva (para organizar, propagar y combinar la información), es conveniente señalar la necesidad del equilibrio entre los modelos viables y los modelos objetivos, para optar por modelos manipulables. Nótese que la comprensión adecuada de un fenómeno atmosférico permite un análisis cuantitativo mediante las ecuaciones que describen sus leyes generales, o a través de relaciones empíricas o estadísticas entre parámetros, y en general la cuantificación sobre las bases de leyes es más rigurosa que la hecha sobre las bases de relaciones, las que tienden a aplicarse en casos específicos [50].

Se conoce como Análisis a la obtención de la imagen de la realidad del estado atmosférico en un momento determinado, representado por un modelo a través de una colección de números. De esta manera, dichos análisis pueden ser útiles por sí mismos como un diagnóstico consistente de la realidad atmosférica, aunque también pueden utilizarse como la entrada o el estado inicial de una predicción numérica del clima, o inclusive como un dato (pseudo-observación) dentro de un proceso iterativo, que puede servir como valor de referencia con el cual comparar la calidad de las observaciones [16].

La información principal disponible es de carácter objetivo, y para generar un análisis se requiere de la colección de observaciones realizadas sobre el estado de un sistema, pero si el modelo que lo representa está sobre-determinado respecto a las observaciones, entonces el análisis se reduce a un proceso de interpolación. Sin embargo, en la mayoría de los casos se tiene una condición sub-determinada, pues los datos son escasos y la relación que guardan con las variables del modelo es indirecta. En estos casos, un planteamiento adecuado del estado de la atmósfera puede alcanzarse si se toma en cuenta la información antecedente del sistema y las estimaciones realizadas a priori respecto al estado que guarda el modelo [16].

El conocimiento sobre las limitaciones físicas concernientes a una condición analizada también es de importancia. Los datos relativos a un estado trivial del sistema o de climatología pueden

utilizarse como información antecedente; también pueden generarse a partir de la salida de un análisis previo, utilizando algunas consideraciones respecto a la consistencia del estado del modelo a través del tiempo como la estacionalidad (hipótesis de persistencia), o la evolución estimada mediante un modelo de pronóstico. De esta manera puede esperarse que dichas consideraciones permitan que la información se acumule y propague sobre las variables que describen los estados del modelo a través del tiempo [16].

Para el estudio del estado del tiempo, el medio atmosférico y los fenómenos que allí se generan, es posible obtenerlos de manera robusta mediante modelos matemáticos. En este sentido, en meteorología se utilizan algunas estrategias equivalentes, sin que éstas cambien el análisis de la condición atmosférica y solo se tenga una simplificación práctica de ella (modelación estadística de errores); por ejemplo, los campos pueden representarse mediante valores promedio localizados sobre las celdas de una malla, por componentes espectrales o por discretización mediante el método de elemento finito [16].

Se conoce como mallado a la determinación de un campo  $f(\vec{r}_i)$  sobre un plano regular (malla) definido sobre un conjunto de puntos  $\vec{r}_i$ , con base en observaciones localizadas arbitrariamente. Comúnmente el vector  $\vec{r}_i$  queda definido a través de su longitud y latitud, aunque el punto también puede formar parte de un espacio de tres y cuatro dimensiones (longitud, latitud, profundidad y tiempo). Cabe señalar que el mallado puede complicarse debido a que la distribución espacial y temporal de las observaciones no es homogénea; entonces, la disponibilidad de observaciones podría ser suficiente para detectar las variaciones de gran escala, pero resultar insuficiente para el análisis de procesos de pequeña escala [7].

Los modelos se caracterizan por tener una resolución más baja que las condiciones reales del sistema físico que representan, y es por ello que aún con el mejor de los análisis se manifestarán diferencias con la realidad. En la aplicación de los algoritmos suele hacerse referencia al estado real del modelo, que se refiere a la mejor de las posibles representaciones de un estado a través de un modelo. Por esta razón, aun cuando las observaciones no presenten errores relacionados con las mediciones de la variable (errores de instrumentación), y el análisis coincida con el estado real del modelo, existirán algunas discrepancias inevitables entre los valores observados y sus equivalentes en el análisis debido a errores de representatividad, aunque estos se consideran comúnmente como parte de los errores de observación [50].

Respecto a los errores propios de las observaciones, estos son considerados como errores aleatorios (también conocidos como de precisión) y errores sistemáticos (conocidos como sesgo). En este sentido, los errores aleatorios tienen la propiedad de que son reducidos mediante promedios de valores. En cambio, los errores sistemáticos no son reducidos por el promedio, pero si se dispone de información, estos pueden ser disminuidos con observaciones [50].

Aunado a lo anterior, conviene señalar que tanto los datos iniciales disponibles como los modelos a utilizar contienen imprecisiones y ruido aleatorio, lo que puede provocar diferencias significativas entre el estado pronosticado y el estado real de un sistema. Los errores en los modelos tienen su origen durante su construcción; por ejemplo, pueden estar incompletos por una falta de entendimiento de la condición por analizar, o debido a procesos omitidos para optar por una solución simplificada. La forma más adecuada de combinar la información de las observaciones y del modelo, así como de proporcionar una estimación de los errores, es lo que se conoce como Asimilación de datos [50].

El objetivo de la asimilación de datos meteorológicos es la obtención de una representación regular del estado de la atmosfera, hasta en cuatro dimensiones, la que debe ser físicamente consistente, generada a partir de un arreglo heterogéneo de instrumentos in situ y de percepción remota, como una muestra imperfecta e irregular del espacio y del tiempo. El aspecto regular y consistente del resultado proviene del uso de modelos, por lo que la asimilación de datos es una disciplina que integra de manera natural la teoría (a través de modelos) con las observaciones de la realidad (mediante instrumentos) [80].

La predicción del clima es un problema de valor inicial, siempre que una estimación del estado atmosférico, en términos de las variables del modelo de predicción, permita simular estados posteriores. En este sentido, la determinación del estado inicial del modelo de predicción es la Asimilación de datos, y el resultado del estado se conoce como Análisis. Por esta razón, la calidad del análisis es crucial en la capacidad de predicción del modelo [41]. La dificultad principal que se presenta es que la determinación del estado inicial a partir de observaciones es un problema sub-determinado. Las variables del modelo superan numéricamente a las observaciones, y además, éstas no representan de manera homogénea el volumen simulado por los modelos de predicción. En caso de requerirse información adicional, suelen considerarse valores supuestos, aunque es más común utilizar información derivada del empleo de un pronóstico a corto plazo (campo de antecedentes o primeras estimaciones) [41].

En épocas más recientes, el objetivo principal en la asimilación de datos meteorológicos ha consistido en la combinación de una amplia variedad de observaciones (y de campos antecedentes) para la estimación del estado atmosférico en un momento dado. Desde un punto de vista estadístico, el resultado de la asimilación de datos corresponderá con el estado esperado de la atmósfera dada la información disponible, de manera que pueda evaluarse la magnitud de los errores, y en general existen dos métodos básicos para afrontar este problema. El primero utiliza un “observador dinámico” que proporciona un esquema de asimilación de datos secuencial. Así, las observaciones son una “respuesta” dentro del modelo cada vez que éstas estén disponibles, y la mejor estimación obtenida se utiliza para pronosticar estados futuros. El segundo método se basa en un “observador directo”, el cual proporciona un esquema de asimilación variacional de datos de cuatro dimensiones. En este caso se considera que una trayectoria representativa del estado es el mejor ajuste de los datos observados en un periodo de tiempo, y los estados estimados al final del periodo se utilizan para generar la predicción siguiente. Bajo determinadas hipótesis matemáticas se busca que estos procesos permitan la estimación del estado de manera óptima. En el subcapítulo siguiente se describe la formulación matemática de la asimilación de datos, y más adelante se discuten varias técnicas para realizar la asimilación [50].

Existe una variedad amplia de modelos para la descripción de sistemas orientados al estudio del medioambiente, así como para el análisis de fenómenos dentro de los campos de la física, biología y economía, de manera que se han involucrado desde modelos simples de ecuaciones diferenciales lineales, determinísticos, continuos, hasta modelos de ecuaciones diferenciales parciales no lineales continuos o discretos más sofisticados [50].

Los primeros intentos por definir rutinas para la estimación del clima se caracterizaron por la combinación de trabajo manual extenuante (interpolación de observaciones climatológicas hacia la ubicación de puntos distribuidos espacialmente) y cálculos computacionales de procedimientos numéricos de predicción. Como ejemplos de trabajos tempranos se tienen los métodos desarrollados por Panofsky (1949) y Gilchrist y Cressman (1954), aplicables a la interpolación de datos observados para obtener estimaciones probables sobre los puntos de una malla.

### 2.1.1. Esquema de Panofsky

Uno de los primeros trabajos en los que se hizo mención del término de Análisis Objetivo para el desarrollo de mapas del clima fue un artículo presentado por Panofsky en el año 1949, [66]. El interés principal de ese trabajo fue la representación de campos meteorológicos como modelos polinómicos, utilizando datos obtenidos directamente de estaciones de observación, de manera que los coeficientes se determinaron por medio de la condición de mínimos cuadrados en los puntos observados sobre un área grande.

Con la observación de los vientos a una cierta altura, Panofsky determinó un polinomio de tercer grado, concluyendo que existían pocas diferencias sistemáticas entre los vientos observados y los vientos calculados con el polinomio. Por otro lado, fue el primero en considerar el problema de la combinación de campos meteorológicos, al proponer dos casos para la combinación de los campos de presión y de viento, sobre las bases de las relaciones geostróficas.

Adicionalmente, Panofsky investigó los efectos de los errores aleatorios de las observaciones sobre los resultados de una aproximación polinomial en un campo. Con esta intención realizó un experimento estadístico específico, utilizando el método de “las urnas con errores aleatorios”, y observó que la variación de la presión en las estaciones, al introducir estos errores, solo causaba variaciones pequeñas en los resultados de la aproximación polinomial.

Cabe señalar que los trabajos descritos se realizaron antes de que las predicciones numéricas obtenidas mediante computadoras electrónicas comenzaran a utilizarse en meteorología. Dado que el uso de esta herramienta significó un estímulo poderoso para el desarrollo de métodos de análisis objetivo, desde entonces la atención principal se ha enfocado en obtener valores de los elementos analizados sobre puntos separados regularmente (para utilizarlos después como datos iniciales en la predicción numérica), en lugar de la construcción de todo el campo [35].

### 2.1.2. Esquema de Gilchrist y Cressman

Gilchrist y Cressman presentaron en 1954 los resultados de un experimento relativo a un análisis objetivo en dos dimensiones [37]. Se consideraba un ajuste por mínimos cuadrados de un polinomio de segundo grado, respecto de un conjunto de datos meteorológicos dentro de un área limitada por una malla rectangular. Ellos representaron un campo de presión (alturas de superficies isobáricas) con polinomios, como lo hizo Panofski, pero realizaron aproximaciones separadas por cada punto central de las celdas; es decir, rodearon cada punto con un cuadrado de 1 000 km por lado, y utilizaron los datos de las alturas y de vientos dentro de cada cuadrado para determinar la magnitud del campo de alturas que les correspondían.

Cuando se utilizan datos de alturas de superficies isobáricas junto con datos de viento, se necesita definir cuánto peso se da a los datos del viento en comparación con los datos de altura. Lo anterior se analizó con un experimento numérico en el que los cálculos fueron llevados a cabo para distintos pesos, y después los resultados se compararon con el promedio de tres análisis subjetivos independientes de la misma situación. Se encontró que podían definirse intervalos para los cuales las relaciones de los pesos eran óptimas.

Por otra parte, se tuvieron dificultades en las regiones donde los datos eran escasos para la aproximación de un campo de alturas por medio de un polinomio de segundo grado, ya que se requiere de un mínimo de mediciones iniciales, pues de otra manera se observarían inestabilidades apreciables en las estimaciones. Además, en regiones definidas por una amplia red de puntos

separados, para superficies de 1 000 x 1 000 km<sup>2</sup>, donde generalmente hay menos datos que los necesarios, Gilchrist y Cressman realizaron sus cálculos varias veces e incluyeron como datos valores de alturas en los centros de las celdas, obtenidos de cálculos previos. Adicionalmente, incrementaron el entorno del centro de la celda a un área cuadrada de 1 800 km por lado, y observaron que, para regiones con celdas separadas ampliamente, puede obtenerse información importante de la predicción numérica.

Dispusieron también de los resultados de tres análisis subjetivos independientes, así como de un análisis objetivo. Mediante el cálculo del mínimo de los cuadrados de las diferencias, observaron que las discrepancias entre el análisis objetivo y cada uno de los análisis subjetivos fue aproximadamente el mismo que las discrepancias mutuas entre los tres análisis subjetivos. Finalmente, realizaron predicciones numéricas de una situación particular, de acuerdo con datos analizados subjetiva y objetivamente, y encontraron que la calidad derivada del análisis objetivo fue superior, por lo que concluyeron que éste ofrece mejores resultados [35].

## 2.2. Asimilación de datos para sistemas dinámicos no lineales

### 2.2.1. Formulación básica de mínimos cuadrados para modelos perfectos

En el caso de los esquemas de asimilación de datos, se plantea que para cualquier estado inicial y de entrada de datos, la modelación de sistemas dinámicos es posible mediante ecuaciones para definir exclusivamente los estados del sistema en tiempos futuros, lo que se conoce como hipótesis del modelo perfecto [50].

Un esquema de asimilación de datos se describe mediante un sistema modelado por las ecuaciones discretas no lineales siguientes [50]

$$\mathbf{x}_{k+1} = \mathcal{M}_{k,k+1}(\mathbf{x}_k), \quad k = 0, \dots, N - 1 \quad (2.1)$$

donde:

$\mathbf{x}_k \in \mathbb{R}^n$ , vector de  $n$  estados del modelo en el tiempo  $t_k$ ,

$\mathcal{M}_{k,k+1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , operador no lineal del modelo, que describe la evolución de los estados desde el tiempo  $t_k$  hasta el tiempo  $t_{k+1}$ . Se incluyen las entradas conocidas al sistema, así como funciones y parámetros que lo describen.

*Como un primer paso en el análisis de asimilación de datos, se requiere definir el espacio de trabajo, por lo que el conjunto de números utilizados para representar el estado atmosférico de interés debe organizarse mediante un arreglo matricial denominado vector de estado  $\mathbf{x}_k$  [43]. La manera cómo los componentes del vector se relacionan con la realidad del sistema a representar depende del trabajo de discretización a considerar. Entonces, como la intención principal del análisis es la estimación del estado de un campo meteorológico que varía en el espacio y el tiempo, la aplicación de modelos numéricos es práctica común; sin embargo, este tipo de modelos solo pueden ser representaciones discretas del campo físico, por lo que para la estimación de la realidad del sistema se requiere de su proyección sobre un espacio discreto [14].*

Las estimaciones a priori o antecedentes,  $\mathbf{x}_0^b$ , de los estados iniciales  $\mathbf{x}_0$  en el tiempo  $t_0$  se suponen conocidas, generalmente a través de un pronóstico anterior o por climatología. Se considera que las observaciones están relacionadas con los estados del sistema mediante [50]

$$\mathbf{y}_k = \mathcal{H}_k(\mathbf{x}_k) + \varepsilon_k^o, \quad k = 0, \dots, N \quad (2.2)$$

donde:

$\mathbf{y}_k \in \mathbb{R}^{p_k}$ , vector de observaciones  $p_k$  en el tiempo  $t_k$ ,

$\mathcal{H}_k : \mathbb{R}^n \rightarrow \mathbb{R}^{p_k}$ , operador no lineal de observaciones, que incluye transformaciones e interpolaciones en la malla, del vector de estado hacia el espacio observacional,

$\varepsilon_k^o \in \mathbb{R}^{p_k}$ , errores de tipo observacionales, de instrumentación y por representatividad.

*Para el análisis se requiere un conjunto de valores observados, los cuales se concentran en un vector de observaciones, y para utilizarlos en un procedimiento es necesaria su comparación respecto al vector de estado. Sin embargo, en la práctica el modelo dispone de una menor cantidad de observaciones que de variables, y aunque la falta de observaciones no es una fuente de errores per se, sí es una condición que trae como consecuencia la indeterminación del estado del sistema [15]. La comparación se realiza con el empleo de una función entre el espacio de estados modelados y el espacio de las observaciones, denominado operador de observaciones. Se trata de una colección de operadores de interpolación entre el modelo discretizado y los puntos de observación, y conversiones entre las variables del modelo y los parámetros observados [16].*

*La fuente de errores que regularmente se identifica y caracteriza de manera correcta es la que está relacionada con aspectos de instrumentación o procesos de medición, ya que los valores del vector de observaciones se ven afectados por la existencia de errores de naturaleza aditiva o multiplicativa [15].*

*En la práctica, la física del sistema se considera mediante un modelo determinístico aplicado sobre el vector de estado (operador de observaciones) [80]. Como errores de representatividad se consideran aquellos relacionados con la representación física definida por el operador de observaciones, así como los errores derivados de la proyección del estado real del sistema sobre un espacio discreto [14].*

El análisis óptimo consiste en obtener las mejores estimaciones  $\mathbf{x}_k^a$  de los estados del sistema  $\mathbf{x}_k$ ,  $k = 0, \dots, N$ , para el ajuste de las observaciones  $\mathbf{y}_k$ ,  $k = 0, \dots, N$ , y el estado antecedente  $\mathbf{x}_0^b$  sujeto al esquema de asimilación de las ecuaciones del modelo (2.1). Se trata de un problema pesado no lineal de mínimos cuadrados [50].

*Es importante distinguir entre la realidad en sí misma (mucho más compleja que su versión representada mediante un vector, como una proyección de la realidad sobre un espacio discreto) y la mejor representación de la realidad, definida por un vector de estado  $\mathbf{x}_k$ , o el estado verdadero que corresponde al tiempo analizado (que suele representarse por  $\mathbf{x}_k^t$ ). Si el resultado de un análisis parcial correspondiente a un tiempo presente o anterior se utiliza como punto de inicio o conocimiento previo a un análisis subsecuente, este vector de estado se considera como la estimación a priori, primera estimación o antecedente del estado verdadero,  $\mathbf{x}_k^b$ , antes de completar el análisis. Finalmente, el resultado del procedimiento de asimilación se conoce como Análisis  $\mathbf{x}_k^a$ , y es el valor del estado que se está buscando [16].*

El problema de asimilación de datos no lineal se define por la función objetivo  $J$  [50], a minimizarse con respecto a  $\mathbf{x}_0$ , condicionada a  $\mathbf{x}_k$ ,  $k = 1, \dots, N$

$$J = \frac{1}{2}(\mathbf{x}_0 - \mathbf{x}_0^b)^T \mathbf{B}_0^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) + \frac{1}{2} \sum_{k=0}^N (\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k)^T \mathbf{R}_k^{-1}(\mathcal{H}_k(\mathbf{x}_k) - \mathbf{y}_k) \quad (2.3)$$

que satisface al sistema de ecuaciones (2.1) con el estado inicial  $\mathbf{x}_0$ .

Los estados  $\mathbf{x}_k$  que satisfacen las ecuaciones del modelo (2.1) se determinan únicamente por los estados iniciales, y por lo tanto pueden escribirse de manera explícita en términos de  $\mathbf{x}_0$ . Al realizarse la sustitución en la función objetivo, ecuación (2.3), la optimización del problema se expresa en términos de estados iniciales únicamente. Así, el problema de asimilación se convierte en un problema pesado de mínimos cuadrados sin restricciones, donde los estados iniciales son las variables de control requeridas en la optimización. Las matrices pesadas  $\mathbf{B}_0 \in \mathbb{R}^{n \times n}$  y  $\mathbf{R}_k \in \mathbb{R}^{p_k \times p_k}$ ,  $k = 0, 1, \dots, N$ , son simétricas y positivas, y se consideran para proporcionar una solución suave del problema. Dichas matrices representan la incertidumbre en los estados antecedentes y la incertidumbre en las observaciones, respectivamente [50].

*Para considerar el hecho de que existe incertidumbre respecto a la magnitud de los antecedentes, las observaciones y el análisis, es de interés el estudio de la modelación de errores entre estos vectores y la realidad. Por ejemplo, dado un campo de estimaciones antecedentes, justo antes de completar el análisis, existe uno y solo un vector de errores que lo diferencia del estado verdadero. Si fuera posible repetir el procedimiento de análisis una gran cantidad de ocasiones, exactamente bajo las mismas condiciones, aun entonces se producirían errores generados por distintas causas, y el vector de diferencias sería distinto en cada ocasión. No obstante, en estas condiciones es de esperarse que los estadísticos correspondientes (promedios, varianzas e histogramas de frecuencias) converjan hacia valores que dependen solamente del proceso físico responsable de los errores, y al repetirse el análisis bajo las mismas condiciones, aunque no pueda definirse el vector de diferencias anticipadamente, sus estadísticos sí serán conocidos [16].*

En estas aproximaciones, los estados iniciales se consideran como parámetros para minimizar los errores cuadráticos medios pesados entre las observaciones pronosticadas por el modelo y las observaciones medidas durante el periodo de tiempo. Para obtener el mejor ajuste, el estado inicial puede adaptarse a distintas posiciones mediante un algoritmo de minimización iterativo eficiente [50].

### 2.2.2. Mejor estimación lineal de mínimos cuadrados

Debido a que la solución explícita a problemas de asimilación de datos no lineales (ecuación 2.3) podría no encontrarse, resulta conveniente optar por una solución que corresponda con la mejor estimación lineal del problema no lineal. Así, se supone que la salida del análisis estimado  $\mathbf{x}_0^a$  del estado antecedente  $\mathbf{x}_0^b$  es una combinación lineal de las innovaciones ( $\mathbf{d}_k = \mathbf{y}_k - \mathcal{H}_k(\mathbf{x}_k^b)$ ,  $k = 0, 1, \dots, N$ ), y entonces se procura obtener la mejor estimación  $\mathbf{x}_0^a$  que resuelva el problema de asimilación de datos de mínimos cuadrados [50].

*Un aspecto principal en la asimilación de datos radica en la consideración de las discrepancias que existen entre las observaciones y el vector de estado, lo cual se cuantifica como un incremento de observaciones a través de un vector de diferencias (vector de innovaciones cuando el cálculo se obtiene respecto de valores antecedentes, o vector de residuos cuando las diferencias se calculan respecto al resultado final del análisis) [16].*

El problema de asimilación se linealiza sobre la trayectoria no lineal antecedente ( $\mathbf{x}_k^b = \mathcal{M}_{0,k}(\mathbf{x}_0^b)$ ,  $k = 1, \dots, N$ ). Las matrices  $\mathbf{H}_k$  y  $\mathbf{M}_{0,k}$  representan la linealización de los operadores de observaciones  $\mathcal{H}_k$  y del modelo  $\mathcal{M}_{0,k}$ , sobre la trayectoria antecedente [50].

$$\mathbf{H}_k = \left. \frac{\partial \mathcal{H}_k}{\partial \mathbf{x}} \right|_{\mathbf{x}_k^b}, \mathbf{M}_{0,k} = \left. \frac{\partial \mathcal{M}_{0,k}}{\partial \mathbf{x}} \right|_{\mathbf{x}_0^b}$$

donde  $\mathbf{H}$  es la matriz de interpolación para evolucionar desde la malla regular del análisis (independiente del tiempo) hacia la red de observaciones (dependiente del tiempo) [27], es decir, es el operador observacional que mapea un vector de estado hacia el espacio observacional;  $\mathbf{M}$  es la matriz que representa al modelo de asimilación lineal, que permite la evolución en el tiempo desde un estado de análisis hacia el estado de pronóstico [79].

*La hipótesis del operador lineal de observaciones (operador tangente lineal o matriz Jacobiana) es necesaria para la definición rigurosa de una expresión algebraica de los incrementos asociados a las innovaciones de los vectores de estado. En la práctica el operador de observaciones puede no ser lineal, pero usualmente se considera que su linealización cobra un sentido físico razonable al aplicarse en la vecindad del estado antecedente, y entonces intuitivamente las ecuaciones de mínimos cuadrados para el análisis deberían generar un estado  $\mathbf{x}_0^a$  próximo al óptimo [16]. Respecto al operador lineal del modelo, este describe la evolución de las perturbaciones en un modelo de pronóstico, utilizando aproximaciones lineales respecto a las perturbaciones del campo modelado, de manera que la linealización es próxima a una solución en el tiempo, y los coeficientes del modelo lineal se definen mediante pendientes tangentes a la trayectoria del modelo no lineal en el espacio [31]. Este modelo de pronóstico tangente lineal puede o no ser adecuado en función de la calidad de los componentes del sistema de asimilación, como la cobertura de los datos, calidad de las observaciones y estados antecedentes, y la resolución del modelo y propiedades físicas del sistema [16].*

La linealización de la función objetivo de mínimos cuadrados resulta [50]

$$\tilde{\mathbf{J}} = \frac{1}{2} \delta \mathbf{x}_0^T \mathbf{B}_0^{-1} \delta \mathbf{x}_0 + \frac{1}{2} \sum_{k=0}^N (\mathbf{H}_k \mathbf{M}_{0,k} \delta \mathbf{x}_0 - \mathbf{d}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{M}_{0,k} \delta \mathbf{x}_0 - \mathbf{d}_k) \quad (2.4)$$

donde:

$$\delta \mathbf{x}_0 = (\mathbf{x}_0 - \mathbf{x}_0^b)$$

Al definirse la ecuación del gradiente del problema linealizado, su solución está dada por la estimación del estado óptimo lineal para  $\mathbf{x}_0^a$

$$\mathbf{x}_0^a = \mathbf{x}_0^b + \hat{\mathbf{K}} \hat{\mathbf{d}} \quad (2.5)$$

$$\hat{\mathbf{K}} = \mathbf{B}_0 \hat{\mathbf{H}}^T (\hat{\mathbf{H}} \mathbf{B}_0 \hat{\mathbf{H}}^T + \hat{\mathbf{R}})^{-1} \quad (2.6)$$

donde  $\hat{\mathbf{d}}$  es el vector de innovaciones, y  $\hat{\mathbf{K}}$  se conoce como matriz de incremento.

*El objetivo principal del esquema de optimización es definir una expresión que minimice la varianza de los errores de la combinación del campo de las observaciones y del modelo. El segundo término del lado derecho de la igualdad en la ecuación (2.5) es conocido como incremento de observaciones, y puede verse como la corrección al modelo debido a las observaciones, lo que eventualmente permite que el resultado final del análisis sea una*

*mejor estimación de las condiciones reales del sistema físico que se está representando. Bajo estos términos, el problema de estimación se orienta hacia la búsqueda de la matriz de incremento que satisfaga los requerimientos señalados [27].*

Para los sistemas donde los operadores del modelo y de las observaciones sí son lineales, el análisis (2.5) es un punto fijo, único y exacto del problema de asimilación de datos. Para sistemas no lineales pueden existir múltiples puntos fijos de la función objetivo (2.3), y el análisis (2.5) es solo una primera aproximación a una solución óptima, debido a la linealización de los operadores del modelo no lineal y de las observaciones [50].

En el problema de asimilación de datos se determina un ajuste de mínimos cuadrados de las predicciones del modelo a las observaciones, sujeto a restricciones. Si se realizan hipótesis adicionales sobre la naturaleza estocástica de los errores en las estimaciones del estado inicial y las observaciones, entonces la solución al problema de asimilación puede interpretarse en términos estadísticos, y la incertidumbre en el análisis puede estimarse [50].

*Se conoce como sesgo al promedio de los errores calculados, y puede considerarse como el indicio de algún problema sistemático en el sistema de asimilación, por ejemplo, una desviación del modelo, un sesgo en las observaciones o un error sistemático en la forma de utilizarlos. Los sesgos son vectores de las mismas características que los vectores de las observaciones o de los estados del modelo. Respecto a las covarianzas de los errores como parte de un sistema escalar, éstas quedan definidas como la varianza o el valor cuadrático medio de las diferencias entre los errores. En un sistema multidimensional, las covarianzas se representan por matrices simétricas cuadradas, cuya diagonal principal se compone de las varianzas correspondientes a las variables del modelo, y los demás términos son las covarianzas cruzadas de cada pareja de variables en el modelo [16].*

Para obtener una formulación estadística del problema de asimilación de datos, se supone que los errores  $(\mathbf{x}_0 - \mathbf{x}_0^b)$  entre los estados iniciales verdaderos  $\mathbf{x}_0$  y las estimaciones antecedentes previas  $\mathbf{x}_0^b$  son distribuidos aleatoriamente, con media cero y matriz de covarianza  $\mathbf{B}_0 \in \mathbb{R}^{n \times n}$ . Los errores observacionales  $\mathcal{E}_k^o \in \mathbb{R}^{p^k}, k = 0, \dots, N$  definidos en la ecuación (2.2) se suponen insesgados, en serie no correlacionados, vectores distribuidos aleatoriamente con media cero y matrices de covarianza  $\mathbf{R}_k \in \mathbb{R}^{p^k \times p^k}$ . Los errores observacionales y los errores en las estimaciones antecedentes se consideran no correlacionados [50].

*Los errores estadísticos (sesgos y covarianzas) son funciones de los procesos físicos que gobiernan los aspectos meteorológicos de las observaciones disponibles. En el caso particular de las varianzas, éstas reflejan la incertidumbre que se tiene sobre las estimaciones antecedentes o las observaciones. En la práctica, para realizar estimaciones estadísticas se requiere considerar que son estacionarias durante un intervalo de tiempo y uniformes sobre un dominio, de manera que a partir de una muestra de errores puedan especificarse errores estadísticos de manera empírica [16].*

Según estas hipótesis estadísticas, dadas las estimaciones antecedentes  $\mathbf{x}_0^b$  y las observaciones  $\mathbf{y}_k, k = 0, \dots, N$ , la mejor estimación lineal insesgada (BLUE, por sus siglas en inglés) del estado real  $\mathbf{x}_0$  en el tiempo  $t_0$ , corresponde con la mejor estimación de mínimos cuadrados (2.5) para el análisis  $\mathbf{x}_0^a$ . En consideración del incremento óptimo del error de estimación definido en la ecuación (2.6), la incertidumbre del análisis  $\mathbf{x}_0^a$  se describe por la covarianza del error del análisis [50]

$$\mathbf{A} = (\mathbf{I}_n - \hat{\mathbf{R}}\hat{\mathbf{H}})\mathbf{B}_0 \quad (2.7)$$

De todas las combinaciones lineales de las innovaciones de la forma (2.5), la mejor estimación BLUE minimiza la covarianza del error del análisis, y por lo tanto es la solución al problema de asimilación con varianza mínima. Así, el análisis dado por (2.5) es la estimación lineal óptima.

*La hipótesis de errores insesgados es difícil de sostener en la práctica, pues regularmente existen sesgos en el campo de las estimaciones antecedentes (debido a sesgos en los modelos de pronóstico) y en las observaciones (o en los operadores de observaciones). En caso de que los sesgos sean conocidos, estos pueden sustraerse y los procedimientos descritos aplicarse en consecuencia; en cambio, si los sesgos no son removidos, el análisis no será óptimo, aun cuando el sesgo parezca reducirse en un sistema de asimilación entre estimaciones antecedentes y observaciones. Respecto a la hipótesis de errores no correlacionados, ésta se justifica generalmente porque los orígenes de los errores en las estimaciones antecedentes y en las observaciones se suponen completamente independientes [16].*

En la práctica, para resolver problemas de asimilación de datos se utilizan esquemas de asimilación secuencial y esquemas de asimilación variacional.

### 2.3. Esquema de asimilación de datos secuencial

Se describe el esquema de asimilación secuencial para modelos definidos por la ecuación (2.1), donde la relación que existe entre las observaciones y los estados está dada por la ecuación (2.2), bajo la hipótesis del modelo perfecto. El análisis del estado de un sistema hasta la etapa de diagnóstico comienza al suponer que en algún tiempo  $t_k$  se tiene conocimiento de su estado antecedente, vector  $\mathbf{x}_k^b$ , dado un análisis previo o por climatología. Para obtener estimaciones mejoradas  $\mathbf{x}_k^a$ , el vector antecedente se corrige con un proceso de asignación de pesos a las innovaciones, que son las diferencias entre las observaciones  $\mathbf{y}_k$  de los estados verdaderos y los valores  $\mathcal{H}_k(\mathbf{x}_k^b)$  obtenidos con el mapeo del vector de estado hacia el espacio observacional. En una etapa posterior, de pronóstico, el modelo evoluciona hacia el tiempo siguiente, donde hay observaciones disponibles, y los estados evolucionados se convierten en estados antecedentes,  $\mathbf{x}_{k+1}^b$ , los que después han de corregirse para mejorar el análisis en este tiempo  $t_{k+1}$ , y el proceso puede repetirse. La representación matemática del procedimiento es la siguiente [50]

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k (\mathbf{y}_k - \mathcal{H}_k(\mathbf{x}_k^b)) \quad (2.8)$$

$$\mathbf{x}_{k+1}^b = \mathcal{M}_{k,k+1}(\mathbf{x}_k^a) \quad (2.9)$$

donde la matriz  $\mathbf{K}_k \in \mathbb{R}^{n \times p}$ , conocida como matriz de incremento, se selecciona para procurar que los estados del análisis converjan a los estados verdaderos del sistema. Así, con la selección de las matrices de incremento  $\mathbf{K}_k, k = 0, 1, \dots$ , se determina el comportamiento de los estados analizados en el tiempo, por lo que esta selección caracteriza al esquema de asimilación de datos.

Para un esquema óptimo de asimilación secuencial, el análisis  $\mathbf{x}_k^a$  dado por la ecuación (2.8) es la mejor estimación lineal de la solución al problema de asimilación de mínimos cuadrados en el tiempo  $t_k$ . La matriz de incremento  $\mathbf{K}_k$  está dada por

$$\mathbf{K}_k = \mathbf{B}_k \mathbf{H}_k^T (\mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (2.10)$$

donde:

$$\mathbf{H}_k = \left. \frac{\partial \mathcal{H}_k}{\partial \mathbf{x}} \right|_{\mathbf{x}_k^b}$$

Si se asume que los errores de estados antecedentes se distribuyen aleatoriamente con media cero y matriz de covarianzas

$$\mathbf{B}_k = \mathcal{E} \left( (\mathbf{x} - \mathbf{x}_k^b)(\mathbf{x} - \mathbf{x}_k^b)^T \right) \quad (2.11)$$

entonces, el análisis óptimo es igual a la mejor estimación lineal insesgada, BLUE, y minimiza la varianza del error del análisis, dada por

$$\mathbf{A}_k \equiv \mathcal{E} \left( (\mathbf{x} - \mathbf{x}_k^a)(\mathbf{x} - \mathbf{x}_k^a)^T \right) = (\mathbf{I}_n - \mathbf{K}_k \mathbf{H}_k) \mathbf{B}_k \quad (2.12)$$

Para sistemas lineales, la solución dada por (2.8) y (2.10) proporciona el análisis óptimo exacto, pero para sistemas no lineales, esta solución es solo un primer orden de aproximación al óptimo, debido a la linealización  $\mathbf{H}_k$  del operador no lineal de observaciones que se consideró.

En la secuencia del análisis óptimo BLUE se han identificado dificultades computacionales, como el que las matrices de covarianza antecedente  $\mathbf{B}_k$  son requeridas en cada paso, lo que se aprecia en el cálculo del análisis  $\mathbf{x}_k^a$ . En la práctica la mayoría de los esquemas operacionales de asimilación secuencial evitan estas dificultades mediante el uso de ciertas aproximaciones [50].

Existe una diversidad de esquemas de asimilación de datos de tipo secuencial, desarrollados para su aplicación práctica, los que se diferencian por el nivel de detalle del análisis, de acuerdo con los requerimientos de cada sistema. A continuación, se describen algunos de los esquemas de asimilación secuencial utilizados a nivel operacional.

### 2.3.1. Esquemas de correcciones sucesivas

Para un problema de diagnóstico, la aplicación de los esquemas de correcciones sucesivas resulta práctica porque son procedimientos iterativos donde la asignación de pesos a las innovaciones no se realiza de manera óptima (retroalimentación definida por el incremento  $\mathbf{K}_k$ ), pues la intención es suavizar las observaciones contenidas dentro de un radio de influencia en función de la distancia que existe entre ellas, para establecer los estados correspondientes a todos los puntos definidos por una malla regular [50].

De manera general, la estimación preliminar de un campo de interés puede realizarse mediante un modelo de pronóstico o una primera interpolación obtenida a partir de valores observados. Posteriormente, esta estimación es modificada por una combinación de correcciones calculadas para cada celda del espacio de análisis, las cuales son proporcionales a las diferencias entre lo observado y los valores que se estimaron preliminarmente. El peso o ponderación de estas correcciones se realiza empíricamente, considerando que las observaciones más cercanas a una celda tienen más importancia. Finalmente, es necesario efectuar varias fases de corrección sobre los valores que representan los estados de un sistema, de manera que el resultado de una etapa es la estimación inicial de una fase posterior [39].

Se describen a continuación dos de las técnicas de correcciones sucesivas más utilizadas.

- **Esquema de Cressman**

Los métodos de análisis de Panofsky (1949) y Gilchrist y Cressman (1954) fueron desarrollados originalmente para el estudio de áreas con disponibilidad de datos relativamente grande, por lo que al aplicar el esquema de los polinomios definidos por mínimos cuadrados sobre áreas caracterizadas por un número reducido de observaciones ampliamente separadas, se observaron inestabilidades, como el que dada una variación menor en los datos iniciales, podían presentarse variaciones considerables en los cálculos de los factores de peso para interpolación [35].

Aun cuando se realizaron diferentes esfuerzos para controlar este tipo de dificultades, eventualmente los procedimientos de análisis cambiaron, como es el caso del análisis numérico realizado por Bergthórsson y Döös en 1955, [10]. Ellos concluyeron que con frecuencia resulta imposible obtener análisis razonables solo mediante la interpolación entre observaciones sinópticas, independientemente de los modelos empleados, y en ciertos casos, para un problema de pronóstico sería más útil la consideración de observaciones antecedentes, incluyendo la información climatológica de la región. Así señalaron que no estaban en posibilidades de obtener un análisis verdadero, sino únicamente el más probable, y esto permitía justificar el uso de información climatológica, en combinación con valores observados y de pronóstico [26].

Tomando en cuenta los trabajos y conclusiones anteriores, Cressman presentó en 1959 su análisis objetivo de mapas del clima, el cual es un procedimiento de aplicación de correcciones sobre un campo preliminar definido a partir de valores supuestos [25]. Las correcciones se determinan de una comparación de los datos con los valores interpolados del campo de información supuesta en los puntos de observación.

Este esquema de análisis se diseñó para tener mucha flexibilidad, es decir, no se especificó cuál debe ser la naturaleza del campo preliminar, por lo que podría aplicarse sobre un campo estimado a partir de situaciones de incertidumbre (por pronóstico), como resultado de una extrapolación formal, o por cualquier tipo de cálculo general [35]. En particular, Cressman realizó el análisis de las presiones correspondientes a ciertas alturas medidas sobre la superficie terrestre, y con la información de vientos reportados definió el gradiente lateral de la corrección que debería aplicarse.

Cressman estudió sobre la calidad del análisis objetivo, pues utilizando los valores correspondientes a las celdas de una malla regular obtenidos como resultado de un análisis, practicó un procedimiento inverso de interpolación para determinar los valores en la ubicación de las estaciones con datos, y de esta forma comparar dichas cantidades con los datos observados [35]. La técnica consiste en realizar múltiples pasos para la corrección sucesiva de los estados definidos sobre un conjunto de celdas espaciadas regularmente, considerando la información proporcionada por los datos más próximos, con el objetivo de alcanzar la mayor exactitud posible con la menor cantidad de cálculos.

Si la estimación previa o antecedente del estado de un modelo se denota por  $\mathbf{x}^b$  (obtenida a partir de información climatológica, o como resultado de un pronóstico previo), y  $y(i)$  representa un conjunto  $i = 1, 2, \dots, N$  de observaciones de un mismo parámetro, el análisis de Cressman que define el estado verdadero del modelo  $\mathbf{x}^a$ , en cada celda de una malla, se calcula con la ecuación de actualización siguiente [16]:

$$\mathbf{x}^a(j) = \mathbf{x}^b(j) + \frac{\sum_{i=1}^N w(i,j)\{y(i) - \mathbf{x}^b(i)\}}{\sum_{i=1}^N w(i,j)} \quad (2.13)$$

$$w(i, j) = \max\left(0, \frac{R^2 - d_{i,j}^2}{R^2 + d_{i,j}^2}\right)$$

donde:

$\mathbf{x}^b(i)$  estado antecedente interpolado sobre el punto  $i$ ,

$d_{i,j}^2$  medida de la distancia entre el punto central de una celda y el punto de una observación,

$w(i, j)$  función de peso, igual a la unidad cuando el punto de la celda  $j$  coincide con el punto de observación  $i$ . La función es decreciente con la distancia, e igual a cero para  $d_{i,j} > R$ , donde  $R$  es un radio de influencia definido por el analista, que marca el límite a partir del cual una observación no tiene peso.

Durante el desarrollo de un proceso de estimación, una corrección es calculada en cada celda considerando las observaciones contenidas en el radio  $R$ , por lo que se trata de un parámetro de control que define la escala de longitud sobre la que una observación es utilizada. Esta escala puede hacerse variar en el espacio, dependiendo de la cobertura de los datos disponibles. La característica de este esquema de análisis es que la búsqueda del estado verdadero del modelo se realiza sucesivamente, con el uso de valores decrecientes de  $R$ , lo que permite un análisis de espectros de escalas. Los valores más grandes de  $R$  pueden utilizarse para permitir la corrección de los errores de escala máximos en la primera estimación, mientras que al utilizar valores menores se establece un límite mínimo de la escala que puede analizarse [25].

De acuerdo con el esquema [25], la corrección de los valores puntuales antecedentes se realiza por medio de una combinación lineal de valores residuales entre las observaciones y las primeras estimaciones, asignándoles un peso acorde con la distancia que los separa. El ajuste sucesivo de los campos antecedentes se logra con base en una serie de búsquedas de las observaciones más cercanas (por lo regular, de tres a cuatro procesos de búsqueda). La reducción paulatina del radio de influencia tiene la intención de construir escalas de información más reducidas para el análisis, donde la densidad de los datos lo permita. Sin embargo, el hecho de que la función de peso no se aproxime a cero de manera asintótica con el incremento de la distancia  $d_{i,j}^2$ , sino que abruptamente adopta un valor nulo cuando ésta se iguala con el radio  $R$ , ha representado problemas de discontinuidad de los campos analizados, cuando la información disponible se encuentra distribuida irregularmente [49]. Por último, como parte de las conclusiones mencionadas en su documento, Cressman advirtió sobre la necesidad de una solución más satisfactoria del problema, la cual requeriría métodos con detección automática y corrección de los errores que se transmiten directamente durante los ciclos del procedimiento.

- **Esquema de Barnes**

Como parte de una técnica para el análisis numérico de mapas del clima, Barnes presentó en 1964 un esquema de interpolación convergente de promedios pesados, para obtener un nivel alto de detalle en el análisis de datos espaciados aleatoriamente [5]. Se trata de un esquema de Análisis Objetivo, el cuál utiliza una función de pesos Gaussiana en el ámbito espacial, basado en el supuesto de que la distribución bidimensional de una variable atmosférica puede representarse mediante la sumatoria de un número finito de ondas (representación mediante integrales de Fourier). Posteriormente, en 1973 Barnes modificó su método, disminuyendo el número de cálculos necesarios para lograr la respuesta adecuada en longitudes de onda pequeñas, de manera que solo se requiere de una corrección adicional a partir de un campo interpolado inicial [6].

El esquema de Barnes consta de un sistema de corrección en dos pasos. El primero implica una estimación inicial de la distribución de láminas de lluvia, estado antecedente  $S_g^1$ , considerando un conjunto de observaciones  $S(x_m, y_m)$  de la misma variable, distribuidas irregularmente y localizadas en cualquier ubicación conocida  $m$ . Para determinar la distribución bidimensional de la lluvia sobre las celdas de una malla regular que represente una región de interés, las observaciones se someten a una ponderación a través de una función de peso  $w_m$ .

El estado  $S_g^1$  en la celda "g" de una malla de asimilación con  $N$  observaciones se calcula [49]:

$$S_g^1 = \frac{\sum_{m=1}^N w_m S(x_m, y_m)}{\sum_{m=1}^N w_m} \quad (2.14)$$

$$w_m = \exp\left(\frac{-d_m^2}{\kappa_0}\right)$$

donde:

- $d_m$  distancia entre el centroide de la celda de una malla y una observación  $S$ ,
- $\kappa_0$  parámetro de peso que representa el porcentaje de la amplitud de onda original que será considerada en la función de filtrado,
- $w_m$  función de peso, relacionada explícitamente con la densidad de las observaciones, que tiende a cero asintóticamente con el incremento de la distancia  $d_m$ , por lo que es innecesario limitar el radio de influencia para el proceso de búsqueda de observaciones, y la cantidad de datos a considerar puede incrementarse para procurar que un número suficiente de observaciones influirá en el valor estimado de cada celda, sin cambiar la función de peso, y por lo mismo, las características de la respuesta.

La primera versión del procedimiento de Barnes requería la aplicación sucesiva del mismo filtro (función de peso) sobre las diferencias (innovaciones) entre observaciones de estados verdaderos y observaciones estimadas. Dicho esquema también resultaba convergente, pero su inconveniente principal era el gran número de iteraciones requeridas para alcanzar una respuesta adecuada sobre longitudes de onda pequeñas. En la versión modificada del procedimiento se requiere de una sola corrección, al reducirse el parámetro de la función de peso de la primera estimación, lo que acelera la convergencia, además de considerarse el cálculo de valores suavizados de las diferencias obtenidas, para agregarse al campo definido durante la primera etapa [49].

*La función de respuesta de filtrado que propuso Barnes se basa en un campo idealizado de observaciones, continuo y sin errores, aunque en la realidad solo se disponga de campos de observaciones finitos, con un cierto grado de error de naturaleza aleatoria.*

Por otro lado, como lo describe Koch (1981) en su trabajo de aplicación de la metodología de Barnes [48], respecto a la determinación objetiva del tamaño de una celda existen argumentos que sugieren una relación entre el tamaño de la celda  $\Delta x$  y el espaciamiento entre los datos  $\Delta n$ , tal que  $\Delta x/\Delta n$  está en el intervalo de 0.3 y 0.5. La razón aparente para el límite superior está en que son cinco los puntos de celda requeridos para representar una onda (Figura 2-1), por lo que una onda requiere como mínimo de una escala igual al doble del espaciamiento promedio entre observaciones,  $2\Delta n$ , y entonces  $\Delta x$  no debe ser más grande que la mitad de  $\Delta n$ . En cuanto al límite inferior, como los cálculos de las cantidades derivadas (divergencia y vorticidad, por ejemplo) son muy sensibles a la longitud de la celda, y el empleo de una malla demasiado fina podría resultar en un campo derivado poco realista, si tales campos derivados representan las características de una solución de interés, la longitud entre celdas no debe ser mucho menor que la separación de los

datos. Aun así, respecto a la definición de estos límites, Koch señala que el esquema tiene una versatilidad suficiente que permite el redondeo del tamaño de la malla calculado a números enteros convenientes (1 grado de latitud, 10 km, etc.).

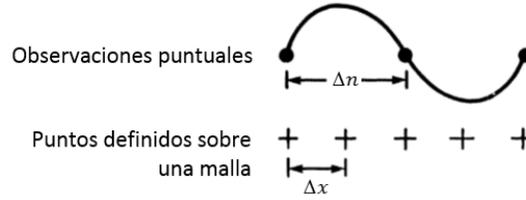


Figura 2-1: Consideración elemental para determinar la relación apropiada entre el tamaño de una celda y el espaciamiento promedio entre datos. Fuente: Modificado de [48].

Una vez que el espaciamiento de datos se define, deben hacerse consideraciones sobre el grado de detalle o nivel de convergencia del campo interpolado en relación con el campo de observaciones, mediante la elección del valor del parámetro de convergencia numérico ( $\gamma$ ). De acuerdo con Koch, para un valor dado de  $\Delta n$ , la máxima respuesta ( $exp^{-1}$ ) del filtro de Barnes se presenta cuando  $\gamma$  es igual a 0.2, y entonces el parámetro de peso  $\kappa_0$  está definido por el espaciamiento de los datos, para una escala donde la longitud de onda es igual a  $2\Delta n$

$$\kappa_0 = 5.052 \left( \frac{2\Delta n}{\pi} \right)^2 \quad (2.15)$$

donde el espaciamiento promedio entre observaciones disponibles sobre el área de estudio es [74]:

$$\Delta n = \sqrt{\text{área de estudio} / \text{número de observaciones}}$$

El segundo paso del procedimiento de Barnes se evalúa mediante la expresión siguiente [49]:

$$S_g^2 = S_g^1 + \frac{\sum_{m=1}^N w'_m [S(x_m, y_m) - S^1(x_m, y_m)]}{\sum_{m=1}^N w'_m} \quad (2.16)$$

donde la función de peso  $w'_m$  para la segunda corrección queda expresada como:

$$w'_m = \exp \left( \frac{-d_m^2}{\gamma \kappa_0} \right)$$

*En la segunda iteración se modifica la aproximación inicial con un parámetro de convergencia numérico  $\gamma$  ( $0 < \gamma < 1$ ), que obliga a un alto grado de convergencia. La función de respuesta (del segundo paso) es la medida del grado de convergencia del análisis, en otras palabras, qué tan aproximados están los valores interpolados con respecto a los datos observados. En el segundo paso el grado de convergencia se incrementa especialmente cuando  $\gamma < 0.5$ . Se observa que el incremento mayor en la respuesta ocurre en la longitud de onda más corta [49]. Tras varias pruebas sobre diferentes datos de distribuciones, Barnes sugiere un valor en el intervalo de 0.2 a 0.4, pues señala que  $\gamma$  menor que 0.2 puede causar desbordamiento al evaluar la función exponencial, y  $\gamma$  mayor que 0.5 no favorece un proceso de convergencia acelerado [6].*

De acuerdo con la ecuación (2.16), el segundo paso del procedimiento consiste en corregir el estado antecedente mediante la asignación de pesos a las innovaciones con la función modificada

$w'_m$ . La innovación se calcula entre cada observación  $S$  y una estimación de la misma,  $S^1$ , la cual se obtiene con el mapeo del vector de estado  $S_g^1$  hacia el espacio observacional, utilizando como operador de observaciones a la interpolación bilineal de los cuatro valores  $S_g^1$  que son adyacentes a la observación  $S$ , o mediante una interpolación más simple cuando la información es insuficiente, como ocurre en el análisis de las celdas periféricas de una malla [48].

Como parte de su investigación, Barnes comenta que los métodos de ajuste polinomial tienen tres desventajas importantes: se requiere de cálculos complicados que implican un costo computacional; los datos respecto a los que se realiza un ajuste se eligen de manera artificial para obtener mejores resultados, y el efecto de los errores en los datos puede ser significativo debido a que cada valor tiene la misma importancia en la determinación del campo analizado [5]. Por otra parte, respecto a la técnica propuesta por Cressman (1959), comenta que la función de peso depende del radio de influencia definido, pues para asegurarse de que una cantidad de información suficiente se ha considerado en la interpolación de datos regionalmente dispersos, es necesario incrementar el radio de búsqueda localmente, lo que produce una respuesta desconocida en el resultado final, obtenido después de aplicar cuatro o más iteraciones sucesiva del esquema. En cambio, el parámetro de peso del esquema de Barnes, útil para establecer el grado de suavizado, se elige anticipadamente al análisis, y los patrones de escala para la distribución de los datos están definidos [39].

### 2.3.2. Interpolación óptima o estadística

Como se describió en los esquemas de correcciones sucesivas, para el planteamiento de tales procedimientos se consideró que el estado que guarda un sistema en la ubicación de un punto de interés puede definirse adecuadamente a partir de los valores observados en la vecindad de dicho punto. Sin embargo, es conveniente señalar que aun cuando la característica principal a considerar en el análisis de un sistema es que todos los datos disponibles son una fuente importante de información, también debe tenerse en cuenta que no es recomendable otorgar una confiabilidad absoluta respecto a la veracidad de todos los datos como una descripción fiel de la realidad. Algunas condiciones bajo las cuales los esquemas de correcciones sucesivas pueden entregar resultados poco satisfactorios son [16]:

- Cuando al disponer de estimaciones de calidad de la variable sobre el espacio definido por el sistema se opte por utilizar datos de aproximación deficiente, solo porque estos últimos se deriven de observaciones realizadas.
- Dada la forma de la función de peso y la importancia que ésta tiene sobre la influencia de las observaciones más distantes a un sitio de interés.
- Si el análisis no respeta las principales propiedades físicas conocidas del sistema, lo que puede tener un efecto en el suavizado y los balances existentes entre variables.

Tomando en consideración la existencia de errores que son propios de la construcción de un modelo que pretende describir las condiciones reales de un sistema, así como los errores en las observaciones realizadas sobre la variable a modelar, no es posible establecer con certeza cuál de ambas representaciones es más confiable en un momento dado. No obstante, resulta conveniente la implementación de estrategias orientadas a minimizar, en promedio, la diferencia entre los valores obtenidos por la aplicación de un esquema y las condiciones reales. Con esta intención se han desarrollado métodos en los que se considera una representación matemática de la incertidumbre sobre los datos, la que puede medirse al asumir o calibrar la estructura estadística de los errores para la interpolación, extrapolación y suavizado de los campos [16].

La estructura estadística de un campo meteorológico se refiere a las leyes que en promedio definen su comportamiento, es decir, las condiciones que regularmente se presentan en conjuntos de elementos meteorológicos de gran dimensión, y no solamente en el comportamiento individual de algunos valores. Sin embargo, en la práctica los valores correspondientes a elementos meteorológicos son conocidos solamente en puntos distribuidos de manera discreta (puntos de observación), pero el número de estos es reducido [35].

La estructura de un campo homogéneo e isotrópico solo depende de la distancia (escalar) entre cada pareja de puntos analizados. Lo anterior se debe a que la homogeneidad de un campo con respecto a la función de estructura de un elemento meteorológico  $f$  se refiere a que la estructura no varía para una pareja de puntos analizados, incluso si estos se transfieren a cualquier lugar dentro del mismo campo, siempre que la distancia entre ellos se mantenga, así como la dirección entre el primero respecto al segundo. Además, un campo es isotrópico si su estructura no se afecta al analizar una pareja de puntos, aun cuando la posición de ambos cambie, siempre que la posición del centro del segmento que los conecta y la longitud de éste no varíen.

Cabe enfatizar que la estructura de los campos meteorológicos no suele ser homogénea e isotrópica, en especial en el análisis de puntos distantes. Así, por ejemplo, en el caso de la temperatura la función de estructura es mayor en la dirección meridional que en la dirección longitudinal, mientras que la función es más grande en la costa que en regiones no costeras. Por ello, los conceptos de homogeneidad estadística e isotropismo no conviene aplicarlos sobre los elementos de los campos meteorológicos directamente, sino únicamente sobre los campos que corresponden a sus desviaciones respecto a un valor medio.

En vista de lo anterior, con la intención de llevar a cabo el análisis objetivo de campos meteorológicos a través de métodos basados en la minimización del error cuadrático medio del análisis, es necesario disponer de información relativa a la estructura estadística de los campos de los elementos por analizar, en particular de la información concerniente a sus funciones de correlación, por ser una medida de la interdependencia entre procesos estocásticos (cero si no están correlacionados) [35].

De acuerdo con Gandin (1965), la Interpolación Óptima debe entenderse como una interpolación media en la cual el error cuadrático medio de interpolación es un mínimo. Este requerimiento es idéntico al que se tiene en la construcción de los modelos de regresión múltiple; la diferencia entre ambos métodos es que los coeficientes de regresión no se determinan cada vez mediante el procesamiento apropiado de una gran cantidad de información. El método de interpolación óptima fue formulado originalmente por Kolmogorov, a quien le atribuyen también resultados importantes respecto a los campos de interpolación y extrapolación. Posteriormente, sus ideas encontraron gran aplicación en varias ramas de las ciencias y la ingeniería, sin embargo, dichos conceptos se aplicaron principalmente para la solución de problemas de extrapolación y discriminación de funciones que representan a todos los puntos dentro de una región, no sobre puntos distribuidos discretamente [35].

Se define a la función de estructura como la diferencia cuadrática media entre los valores de un elemento meteorológico  $f$  en los puntos  $\vec{r}_1$  y  $\vec{r}_2$  de un campo:

$$B_f(\vec{r}_1, \vec{r}_2) = \overline{[f(\vec{r}_1) - f(\vec{r}_2)]^2}$$

donde la barra denota promedio, y la función de estructura depende en general tanto de los vectores  $\vec{r}_1$  y  $\vec{r}_2$  del elemento  $f$ , como de las condiciones promedio.

Junto con la función de estructura del elemento  $f$ , también conviene definir la función de estructura para las desviaciones de  $f$  respecto de su valor medio  $\bar{f}$ , desviación definida como:

$$f'(\vec{r}) = f(\vec{r}) - \bar{f}(\vec{r}) \quad (2.17)$$

Sean conocidos los valores de algún elemento meteorológico  $f$  en la ubicación de  $n$  puntos  $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n$

$$f_i = f(\vec{r}_i) \quad (i = 1, 2, \dots, n)$$

Si se quiere estimar el valor  $f_0 = f(\vec{r}_0)$  que corresponde a la variable en el punto  $\vec{r}_0$ , es conveniente sustraer de cada uno de los valores de la variable  $f$  el valor medio en la ubicación del punto de interés, y posteriormente calcular la desviación  $f'_0$  de la variable  $f$  respecto al valor medio en el punto  $\vec{r}_0$ . Esta desviación puede escribirse como una combinación lineal de las desviaciones respecto al valor medio en los puntos  $\vec{r}_i$ , donde los factores de peso  $p_i$  son desconocidos:

$$f'_0 = \sum_{i=1}^n p_i f'_i \quad (2.18)$$

Aunque no existe un conjunto de factores de peso  $p_i$  que permita alcanzar una situación en la cual la ecuación (2.18) sea satisfecha exactamente en todos los casos de su aplicación, sí puede establecerse que el error (diferencia) cuadrático medio de la ecuación anterior

$$E = \overline{\left( f'_0 - \sum_{i=1}^n p_i f'_i \right)^2} \quad (2.19)$$

es un mínimo, y entonces por tal requerimiento es posible obtener los factores de peso  $p_i$ . La operación de interpolación que utiliza los factores de peso  $p_i$ , de manera que el error cuadrático medio  $E$  se minimiza, se conoce como interpolación óptima.

Al desarrollar la operación indicada en (2.19) se obtiene:

$$E = \overline{f_0'^2} - 2 \sum_{i=1}^n p_i \overline{f_0' f'_i} + \sum_{i=1}^n \sum_{j=1}^n p_i p_j \overline{f'_i f'_j}$$

Una característica importante de la estructura estadística de un campo meteorológico es la función de correlación, la que se define a través del producto de los valores de dos variables meteorológicas en dos puntos diferentes (para el análisis de una sola variable  $f$ , se le denomina función de autocorrelación  $m_{ij}$ ), [58]. Cabe enfatizar sobre la necesidad de considerar a las funciones de correlación respecto de las desviaciones de las variables meteorológicas sobre sus valores medios. En estas condiciones, el error cuadrático medio  $E$  resulta:

$$E = m_{00} - 2 \sum_{i=1}^n p_i m_{0i} + \sum_{i=1}^n \sum_{j=1}^n p_i p_j m_{ij}$$

Como se asume que el campo de estos elementos es estadísticamente homogéneo e isotrópico, entonces las cantidades  $m_{ij}$  dependen únicamente de las distancias escalares correspondientes.

Para disponer de una representación adimensional del error  $E$ , se introduce la variable  $\varepsilon$  (cociente del error cuadrático medio de interpolación y la varianza del elemento  $f$  en el punto  $\vec{r}_0$ ). Entonces, una medida del error de interpolación, a determinarse bajo cualquier método, siempre que la

interpolación pueda reducirse a la determinación de una combinación lineal como en la ecuación (2.18), y los valores  $f'_i (i = 1, 2, \dots, n)$  sean conocidos, es la siguiente

$$\varepsilon = 1 - 2 \sum_{i=1}^n p_i \mu_{0i} + \sum_{i=1}^n \sum_{j=1}^n p_i p_j \mu_{ij}$$

donde  $\mu_{0i}$  y  $\mu_{ij}$  denotan los valores de la función de autocorrelación normalizada, llamados coeficientes de autocorrelación del elemento  $f'$ , en relación a la distancia  $\rho$  que existe entre los puntos respectivos (entre el punto de interpolación 0 y el punto  $i$ , y entre los puntos  $i$  y  $j$ , respectivamente).

Como se requiere de una interpolación óptima, los factores de peso  $p_i$  se determinan con la condición de que la medida del error de interpolación  $\varepsilon$  sea mínima, en comparación con el valor que le corresponda bajo cualquier otra combinación de factores  $p_i$ , es decir

$$\frac{\partial \varepsilon}{\partial p_i} = 0$$

$$\sum_{j=1}^n \mu_{ij} p_j = \mu_{0i} \quad (i = 1, 2, \dots, n) \quad (2.20)$$

que constituye un sistema de  $n$  ecuaciones algebraicas lineales ( $n$  igual al número de estaciones con datos iniciales), para la determinación de los factores de peso  $p_i$ , y posteriormente obtener el valor interpolado del elemento meteorológico mediante las ecuaciones (2.18) y (2.17). Dado que la función de autocorrelación es positiva, el determinante del sistema (2.20) es no nulo, y siempre tiene una solución (mientras los puntos  $\vec{r}_i$  no sean coincidentes), [35].

La aplicación de la condición definida por el sistema de ecuaciones (2.20) implica invertir una matriz de dimensión igual al número de observaciones (o resolver por otro método), cuyos elementos son correlaciones cruzadas entre estaciones de observación, y por lo tanto, las funciones de peso obtenidas son específicas de la configuración de la red de observaciones, ventaja principal del algoritmo. Como una simplificación, puede asumirse un modelo matemático para la dependencia de las correlaciones con la distancia, ajustando una curva a una distribución discreta de correlaciones obtenidas de series históricas de datos, aunque esto implica considerar una forma continua de la distribución de estaciones, cuando en realidad la distribución de datos es heterogénea. Entonces, aunque la aplicabilidad del algoritmo es amplia, éste tiene un costo elevado en términos del proceso de solución, por la necesidad de invertir las matrices de correlación, y también por requerir de series históricas de datos [78].

### 2.3.3. La interpolación secuencial: Filtros de Kalman

En subcapítulos anteriores se ha discutido sobre el proceso de Análisis que forma parte de un esquema de Asimilación de datos, de tal manera que pudo definirse un estimador óptimo bajo cierta consideración del nivel de incertidumbre que le corresponde, a partir de información antecedente y de observaciones disponibles, con la posibilidad de incorporar en la solución del problema el análisis de la dimensión temporal.

Dado que en el campo de la meteorología está presente el interés de contar con mecanismos para la observación y el registro permanente del estado del tiempo con fines de pronóstico, por ejemplo,

se ha procurado que la colección de observaciones y el resultado de los procesos de análisis realizados se apliquen en la práctica de esquemas de asimilación de datos en tiempo real [41].

De acuerdo con Kalman (1960), una variedad importante de problemas en el campo de la comunicación y el control son de naturaleza estadística, como es el caso de la predicción de señales aleatorias, la separación entre señales de interés y señales no deseadas aleatorias (ruido), o la detección de señales con forma conocida en la presencia de ruido aleatorio. Los primeros dos problemas fueron revisados por Wiener en el año de 1949, cuando presentó un trabajo sobre extrapolación, interpolación y suavizado de series de tiempo estacionarias [46]. Estudios posteriores tuvieron por objetivo la especificación de un sistema dinámico lineal (filtro de Wiener) que permitiera la predicción, separación o detección de una señal aleatoria [47]. En su caso, dada una señal de entrada hacia un filtro de Wiener, la respuesta impulsiva es tal que se minimiza el error cuadrático medio definido como la diferencia entre la señal de salida y la respuesta deseada [40].

Como parte de su trabajo, Kalman estudió sobre las características de un sistema dinámico (lineal o no lineal), proponiendo al estado como un concepto fundamental, y lo describe como aquella información cuantitativa (conjunto de números, funciones, etc.) que representa la mínima cantidad de datos que debe conocerse sobre el comportamiento antecedente de un sistema, con la intención de predecir su comportamiento futuro. El aspecto dinámico es descrito entonces en términos de la transición entre estados, es decir, se especifica cómo un estado se transforma en otro a través del tiempo [46].

El filtro de Kalman como procedimiento de interpolación secuencial es un algoritmo para el control óptimo del estado de un sistema dinámico y su correspondiente covarianza, asumiendo un modelo lineal [79]. Bajo estas condiciones, en comparación con el procedimiento de interpolación estadística, donde se pretende la estimación del estado particular de un sistema, ahora se analiza la evolución del modelo para el estado del sistema entre tiempos consecutivos, es decir, se procura la estimación de una serie de estados  $\mathbf{x}_k$ , condicionados a las observaciones pasadas y presentes ( $\mathbf{y}_0, \dots, \mathbf{y}_k$ ). El algoritmo del filtro de Kalman está compuesto de dos etapas: el análisis (actualización de observaciones) y el pronóstico [14].

En un tiempo  $t_k$  se dispone del resultado de una simulación de pronóstico previo, denotado por  $\mathbf{x}_k^f$  (análogo al vector de estado antecedente  $\mathbf{x}_k^b$ , definido anteriormente), y se recolecta un conjunto de observaciones que conforman al vector  $\mathbf{y}_k$ . Dados  $\mathbf{x}_k^f$  y  $\mathbf{y}_k$ , se realiza un análisis para obtener la estimación del estado  $\mathbf{x}_k^a$  (el análisis estadístico que se deriva de estos elementos es similar al descrito en el subcapítulo denominado *Mejor estimación lineal de mínimos cuadrados*). Hasta este punto las observaciones se han utilizado para la actualización del conocimiento previo sobre el estado del sistema (antecedentes). Más adelante se utiliza el modelo dinámico para pronosticar el estado del sistema, del tiempo  $t_k$  al  $t_{k+1}$ , de  $\mathbf{x}_k^f$  a  $\mathbf{x}_{k+1}^f$ , por lo que el resultado del pronóstico servirá como antecedente para un ciclo posterior, y así sucesivamente [15].

Una de las hipótesis más importantes que se admiten en el desarrollo descrito es la linealidad del sistema; sin embargo, en la modelación y el análisis de los fenómenos pueden presentarse condiciones tales como [15]:

- El operador observacional puede ser no lineal,  $\mathcal{H}$ , como suele ocurrir con aquellos que relacionan mediciones ópticas con las variables de estado, por involucrar operadores de transporte radiativo como herramientas para el estudio de la interacción entre la materia

(atmósfera) y la energía (luz). Como ejemplos se tiene a las observaciones de satélite, de dispositivos LiDAR y de radar.

- El modelo dinámico de evolución puede ser no lineal,  $\mathcal{M}$ , como es el caso de las ecuaciones primitivas (diferenciales no lineales) empleadas para el estudio del flujo atmosférico global, incluidas las reacciones químicas.

El algoritmo del filtro de Kalman puede generalizarse para los operadores no lineales  $\mathcal{H}$  y  $\mathcal{M}$ , (Filtro de Kalman Extendido, EKF, por sus siglas en inglés), especialmente si se considera que la naturaleza no lineal del operador observacional y del modelo dinámico es relativamente débil, aunque esto signifique que la característica óptima del análisis se perderá. Si  $\mathcal{H}$  es no lineal, el operador  $\mathbf{H}$  puede definirse como su tangente lineal en la vecindad de  $\mathbf{x}_b$ . Asimismo, si  $\mathcal{M}$  es no lineal, como la mayoría de los modelos meteorológicos y oceanográficos,  $\mathbf{M}$  puede definirse como el modelo de pronóstico tangente lineal en la vecindad de  $\mathbf{x}_a$  [16]. Cabe señalar que, en la aplicación del filtro extendido, cuando el intervalo de tiempo entre dos actualizaciones consecutivas es grande, pueden obtenerse aproximaciones burdas de la matriz de covarianzas del pronóstico [15].

Como principal desventaja del filtro de Kalman se considera que su implementación es impráctica en el análisis de sistemas de gran magnitud, debido al tamaño de las matrices de covarianza que necesitan propagarse. En estas condiciones, Evensen propuso en 1994 una simplificación denominada Filtro de Kalman de Ensamblados (EnKF, por sus siglas en inglés), basado en el concepto de partículas (colección de un número reducido de vectores de estado) como miembros del ensamble, cuya variabilidad está destinada a ser representativa de la incertidumbre del estado del sistema, con el inconveniente de que la calidad de la estimación de los estados dependerá de la dispersión o amplitud del ensamble, que deberá ser suficiente para capturar el comportamiento real del sistema analizado [15].

## 2.4. Esquema de asimilación de datos variacional

Se conoce como asimilación variacional al resultado de la aplicación de algoritmos en los cuales los campos a estimar se determinan explícitamente como los valores que minimizan una función escalar, conocida como función objetivo, útil para el cálculo de la diferencia o desajuste que existe respecto al conjunto de datos disponibles [50].

En la versión original del método variacional, propuesto por Sasaki en 1958, [73], se definió a la función objetivo, la cual consiste en una sumatoria de pesos de las diferencias cuadradas entre las variables de análisis y los datos observados. La estimación óptima de las variables se obtenía como el punto mínimo de dicha función, por lo que se consideró que la mayor ventaja del método era la facilidad con la que podían tomarse en cuenta los balances dinámicos y el suavizado de las estimaciones, si se agregaban adecuadamente ciertos términos de restricción y control [79].

La asimilación variacional, en el caso lineal, minimiza el error de estimación cuadrático de cada elemento que compone al campo de análisis, por lo que puede considerarse como otro algoritmo para resolver el mismo tipo de problema que se planteó al describir el filtro de Kalman [53].

La función objetivo puede representarse como [50]:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T [\mathbf{P}^b]^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y}) \quad (2.21)$$

donde el primer término a la derecha de la igualdad es una medida de la desviación de  $\mathbf{x}$  respecto al estado antecedente, y el segundo término es la desviación respecto a la observación. En este caso,  $\mathbf{x}^b$  es una estimación antecedente del vector de estado  $\mathbf{x}$  (un pronóstico reciente o una estimación climatológica);  $\mathbf{y}$  es un vector de observaciones que dependen del vector de estado a través del operador observacional  $\mathbf{H}$ ;  $\varepsilon^b$  y  $\varepsilon^o$  son errores no correlacionados, con matrices de covarianza  $\mathbf{P}^b$  y  $\mathbf{R}$ , respectivamente.

En la práctica de la meteorología y la oceanografía se ha requerido la formulación de ciertas expresiones más explícitas respecto a la forma general de la función objetivo, ecuación (2.21). Por ejemplo, en el caso de análisis más simple, se dispone de un antecedente  $\mathbf{x}^b$  en un tiempo  $k$ , junto con observaciones y propias del mismo tiempo, y al minimizar la función se produce entonces una estimación del estado del flujo del sistema al tiempo correspondiente. Este caso se conoce como análisis variacional de tres dimensiones, abreviado 3D-Var [50].

Otra situación ocurre cuando se requiere la asimilación de observaciones distribuidas en un periodo de tiempo, por lo que la evolución del flujo debe ser considerada tomando en consideración la disponibilidad de observaciones en tiempos sucesivos  $k = 0, 1, \dots, K$ . Este caso se conoce como asimilación variacional de cuatro dimensiones (4D-Var) de restricción fuerte, lo que significa que la ecuación del modelo debe ser verificada exactamente por los vectores de estado estimados. Adicionalmente, si el error del modelo es tomado en cuenta, puede asumirse que éste tiene una matriz de covarianza  $\mathbf{Q}_k$ , no correlacionada en el tiempo ni tampoco con los errores de las observaciones y de los estados antecedentes. En estas condiciones, a partir de la ecuación (2.21) puede obtenerse una función objetivo definiendo la mejor estimación BLUE para la secuencia completa de estados  $\{\mathbf{x}_k, k = 0, \dots, K\}$ , la cual se conoce como asimilación variacional de cuatro dimensiones (4D-Var) de restricción débil [50].

En la asimilación variacional se busca minimizar la función objetivo en cualquiera de las formas antes definidas, entonces, siendo la variable de control  $\mathbf{x}$  el argumento de la función a minimizar,  $\partial J / \partial \mathbf{x}$  es el gradiente de  $J$  respecto a  $\mathbf{x}$ , y en el mínimo de la función objetivo el gradiente es igual a cero. Una manera de obtener el mínimo puede ser mediante la determinación analítica de expresiones de los componentes del gradiente, para después resolver un sistema de ecuaciones escalares y minimizar los componentes de  $\mathbf{x}$ , lo que puede realizarse principalmente en atención a problemas de dimensiones pequeñas. En caso de no obtener expresiones analíticas para el gradiente, debe procederse de manera numérica, como puede ser la consideración de diferencias finitas con la imposición de perturbaciones  $\Delta x_i$  en todos los componentes del vector de control, y aproximar la derivación parcial  $\partial J / \partial x_i$ . Sin embargo, la aplicación de este procedimiento ha resultado impráctica, especialmente en modelos de grandes dimensiones [50].

El *método de estado adjunto* es un algoritmo eficiente que define una aproximación sistemática para el cálculo del gradiente de una función escalar diferenciable con respecto a sus argumentos. Consiste en el empleo sistemático de la regla de la cadena para la diferenciación de una función compuesta. El propósito es determinar el gradiente  $\partial J / \partial \mathbf{x}_0$  de  $J$  respecto a  $\mathbf{x}_0$ , en consideración de que cualquier perturbación  $\delta \mathbf{x}_0$  de  $\mathbf{x}_0$ . Por otra parte, con la intención de hacer más eficiente el empleo del método variacional, es práctica común utilizar un procedimiento de *aproximación por incrementos* introducido por Courtier en 1994, el cual tiene como idea básica simplificar el modelo dinámico, pues su ventaja fundamental es que permite a voluntad la simplificación de los operadores lineales [50].

Bajo las consideraciones anteriores, la aplicación del método de asimilación variacional se hace factible, y puede decirse que su mayor ventaja es que sí considera la evolución temporal de la

incertidumbre del estado del sistema analizado, pues aun cuando no se calcula explícitamente la incertidumbre en los campos estimados (como en el caso del filtro de Kalman), sí se determina una aproximación de la solución mínima de la función objetivo, que depende de la dinámica del flujo y de la evolución temporal de la incertidumbre. Por el contrario, se considera que la mayor desventaja está en la necesidad de desarrollar el código del estado adjunto, el cual debe validarse cuidadosamente, pues en la práctica se ha observado que incluso errores menores en el cálculo del gradiente pueden mermar significativamente la eficiencia de la minimización [50].

## 2.5. La aplicación de modelos para el análisis de campos meteorológicos

La asimilación de datos como metodología para la estimación del estado (diagnóstico y pronóstico) de un sistema que varía a través del tiempo, a partir de datos observados y de un modelo basado en leyes físicas conocidas del sistema, se ha convertido en una herramienta indispensable para la investigación meteorológica y la predicción numérica del clima, dado que su implementación operacional comenzó en la década de los 50 del siglo pasado [79].

En párrafos posteriores se mencionan algunos procedimientos utilizados en épocas recientes para el análisis de datos, cuyo sustento teórico está basado en conceptos descritos con anterioridad, los cuales se han aplicado especialmente para la estimación del estado de diagnóstico de campos de precipitación, que es el objetivo principal de este trabajo de investigación.

### 2.5.1. La combinación de lluvias registradas en superficie y estimadas por satélite

La evaluación de modelos numéricos de precipitación en regiones donde existe una dispersión irregular y baja densidad de registros de lluvia en superficie, como es el caso de la región sur y centro del continente americano, representa un desafío debido al bajo nivel de confianza que puede asociarse con la interpolación directa de la información disponible. Por su parte, las estimaciones de precipitación por percepción remota, que se basan en la temperatura del tope de las nubes, por ejemplo, se han aprovechado extensamente para cubrir los espacios de información donde no se dispone de datos observados, como es el caso de la Misión de Medición de Lluvias Tropicales (TRMM, por sus siglas en inglés). Sin embargo, a pesar de lo valioso de este tipo de productos en la validación numérica de modelos, se tiene conocimiento de que las estimaciones que proporcionan están asociadas con ciertos errores sistemáticos, lo que suele provocar subestimaciones de lluvia durante el análisis de temperaturas del tope de nubes cálidas, o sobreestimación de lluvia por el análisis del tope de nubes frías [71].

Para hacer frente a los problemas de interpolación en regiones con baja disponibilidad de registros de precipitación, así como a la sobreestimación y subestimación de esta variable en los productos de percepción remota, se han realizado trabajos de combinación de estas bases de datos, los que han permitido la determinación de distribuciones espaciales de precipitación de mayor calidad. Como ejemplo, Rozante *et al.* 2010 realizaron un procedimiento de combinación (MERGE), considerando un aproximado de 1 500 registros de lluvia diaria, correspondientes a los trimestres de verano e invierno del 2007, así como los valores estimados con la misión satelital TRMM de resolución espacial de 0.25°, sobre toda la extensión territorial del sur de América. Las estimaciones de láminas acumuladas de 5 días consecutivos se obtuvieron con el esquema de Barnes, para la interpolación de los datos por correcciones sucesivas. La evaluación de los campos original y combinado se realizó en la ubicación de un conjunto de observaciones inicialmente removidas para fines de comparación (validación cruzada), tomando en cuenta

algunos índices de precipitación como el sesgo, la probabilidad de detección, el índice de falsas alarmas, la raíz del error cuadrático medio y el índice de amenaza corregido.

De acuerdo con los resultados, la evaluación de modelos sobre áreas de gran extensión debe realizarse con especial cuidado, en particular para regiones donde la red de observaciones es irregular y dispersa. En zonas con alta densidad de registros en superficie, no se apreció un reforzamiento significativo en la calidad de las estimaciones por combinación, respecto a los cálculos obtenidos por interpolación simple de registros, lo que implica que en casos como éste la contribución de los valores agregados por la TRMM pierde importancia. Por el contrario, los resultados de análisis sobre regiones con baja densidad de observaciones mostraron una mejora sustancial por la combinación de las bases de datos consideradas.

### 2.5.2. La técnica de suavizado Kernel para la combinación de datos de lluvia

Debido a la incertidumbre asociada al cálculo de la precipitación media sobre una superficie al emplear solamente registros puntuales dispersos, en especial para eventos convectivos locales por tener una variabilidad espacial grande, así como al hecho de que las estimaciones de lluvia basadas en mediciones satelitales de resolución espacial razonable generalmente se caracterizan por una exactitud menor, se han desarrollado procedimientos para la síntesis (combinación) de este tipo de información, basados en algunos conceptos sobre la asimilación de datos, como es el trabajo de Li *et al.* 2010, quienes estudiaron el problema de la incompatibilidad debida a la combinación de datos con escalas espaciales distintas. El procedimiento propuesto se enfocó en la diferencia (residuo) entre la magnitud de las observaciones de lluvia registradas puntualmente en superficie y los valores estimados por la misión TRMM con resolución espacial de  $0.25^\circ$ , empleando un método de interpolación de suavizado Kernel para no depender de un comportamiento espacial estacionario de los datos [52].

Entre las consideraciones principales se menciona que la mayoría de los métodos proceden de manera directa a la combinación de las bases de datos de lluvias puntuales registradas y medias de origen satelital, sin tratamiento alguno sobre estas últimas, lo que genera sesgos alrededor de los límites entre celdas contiguas de una malla, debido a la discontinuidad de los campos analizados por aspectos de discretización respecto de los campos verdaderos. Se describe un procedimiento para reducir los sesgos, que consiste en el suavizado de las lluvias de origen satelital estimadas en los sitios de registro pluviométrico, de manera que el estado antecedente se integre de valores pesados dada la magnitud de las áreas de influencia correspondientes a las celdas más próximas a los pluviómetros.

El método requiere la estimación de un campo de residuos en los puntos de una malla regular, obtenido por la combinación pesada de los residuos en los puntos observados (utilizando lluvia satelital suavizada) y un conjunto de pseudo-residuos sobre la malla definida (calculados a partir de pseudoobservaciones, derivadas de una interpolación simple de las observaciones), en función de las distancias entre datos. El campo combinado de interés está dado por la diferencia entre las estimaciones del satélite (estado antecedente) y el campo de residuos calculado. Cabe señalar que la consideración de las pseudoobservaciones es parte de una estrategia de doble suavizado, recomendada para afrontar el inconveniente de trabajar con escasez de pluviómetros.

Como un aspecto atractivo del procedimiento, se señala que el campo de residuos está representado por un modelo de regresión no paramétrico, porque no se consideran suposiciones relativas a la distribución de los errores, ni la función de regresión requiere de una forma específica. Respecto a la calidad de estimaciones del campo combinado, ésta depende de la

función Kernel empleada, así como de cierto ancho de banda [61]. Por ejemplo, existen esquemas de asimilación de datos que aplican funciones Kernel particulares, como los análisis de Cressman y de Barnes, este último mediante una función Gaussiana; sin embargo, también se menciona que la selección del tipo de función puede no ser tan crítica, y representar hasta un 5% de diferencia en el cálculo del error cuadrático medio. En cuanto al ancho de banda, un valor reducido genera sesgos menores, pero varianzas mayores, y viceversa.

Con la intención de verificar la calidad de los resultados obtenidos con su modelo de combinación, Li *et al.* 2010 realizaron un estudio piloto al considerar un total de 120 precipitaciones diarias ocurridas en el año 2001 sobre la superficie de Australia, implementando una validación cruzada mediante la remoción de algunas observaciones, actividad previa a la combinación, para su comparación posterior con los valores estimados. Como elementos de comparación adicionales, también se consideraron los resultados de aplicación de procedimientos Kriging Ordinario y Co-Kriging, el primero entendido como la mejor predicción lineal insesgada para la estimación de un valor en un punto, dada la combinación lineal óptima de los datos en su vecindad (siempre que su variograma sea conocido), mientras que Co-Kriging es una generalización de Kriging para trabajar con datos y variogramas multivariados, siempre que exista dependencia espacial entre la variable objetivo y las variables auxiliares o covariables. Se observó que los métodos Kriging generaron resultados con mayor sesgo, con una tendencia a la subestimación, lo que pudo ocurrir debido al hecho de que la estacionalidad espacial asumida por estos métodos no necesariamente se cumple en los eventos de lluvia, cuya distribución es ampliamente sesgada a la derecha. Por su parte, al utilizar el doble suavizado Kernel se obtuvieron sesgos menores, así como resultados comparables o mejores en otros estadísticos, como la raíz del error cuadrático medio, además de mostrar un mejor desempeño visual en las zonas con baja densidad de observaciones.

### 2.5.3. La combinación de lluvias registradas en superficie y estimadas por radar

Como consecuencia del espaciamiento entre estaciones de registro de precipitación en superficie, el análisis del campo de la distribución espacial de la lluvia a partir únicamente de los registros puntuales disponibles generalmente se ha practicado utilizando métodos de interpolación como el de la distancia inversa, o métodos geoestadísticos como Kriging. Sin embargo, aun cuando la red de registros puntuales sea relativamente densa, ésta no será suficiente para capturar la naturaleza real del campo de la lluvia, en especial para escalas de tiempo reducidas. Por esta razón, el uso de radares meteorológicos como instrumentos de percepción remota se ha considerado para el estudio de la distribución espacial y temporal de la lluvia con alta resolución en regiones amplias, aunque con resultados limitados debido a ciertos errores identificados en sus estimaciones.

Para obtener provecho de las fortalezas de ambos tipos de mediciones, algunos procedimientos de combinación se han implementado, como son las técnicas estadísticas de análisis uni y multivariados, Kriging Ordinario y Co-Kriging donde el variograma es la función más importante al ser una representación de las características espaciales de la lluvia, influenciado por el tipo de tormenta, la densidad de la red de registros y el periodo de acumulación de lluvia. Se ha observado que el desempeño de cualquier método de interpolación depende de la densidad de la red de registros, mientras que los de ajuste simple, como la corrección del sesgo medio del error [86], muestran menor sensibilidad respecto a la densidad. También se sabe que la combinación de lluvias registradas puntualmente y estimadas con radar mejora si los intervalos de acumulación son mayores, además de que la calidad de las estimaciones puede variar entre estaciones del año [59].

Bajo estas condiciones, Nanding *et al.* 2015 realizan una evaluación del desempeño de diferentes métodos de combinación de lluvias, observando las consecuencias de considerar varias

densidades y configuraciones de la red de registros, pues combinaron datos de 161 pluviómetros y lluvias de radar meteorológico de 1 km de resolución, particularmente 20 días del 2007 sobre una porción del norte de Inglaterra, en escalas de una hora para clasificar tormentas entre convectivas y estratiformes. Para la calibración de las técnicas de combinación, en el área de estudio se definieron regiones para la selección aleatoria de registros a utilizar, lo que permitió conformar diferentes problemas en función de la densidad de los registros considerados, así como la posibilidad de revisar su comportamiento al modificar la configuración de su distribución espacial.

La combinación de las bases de datos se efectuó mediante la aplicación de diferentes métodos. La corrección del sesgo medio se utilizó por considerar que las estimaciones de radar están afectadas por un error multiplicativo uniforme simple, debido por ejemplo a una calibración electrónica pobre del instrumento, o por una relación deficiente entre la reflectividad de una señal emitida y la cantidad de lluvia ocurrida sobre una superficie (ecuación del radar). El método Kriging Ordinario se utilizó únicamente para la interpolación de las observaciones de lluvia sobre puntos específicos, mientras que una modificación de este método se aplicó de tal manera que el campo combinado no solo diera seguimiento a la estructura espacial de las estimaciones originales del radar, sino también preservara el valor medio del campo de las observaciones consideradas. Los resultados mostraron que las combinaciones basadas en Kriging reducen los errores medio absoluto y cuadrático medio de la lluvia del radar original, para los eventos convectivos y estratiformes estudiados, aunque su calidad depende de la densidad de registros, mientras que existe una menor dependencia respecto a la configuración de las observaciones cuando la densidad de registros es mayor.

#### 2.5.4. La cuantificación de lluvia con datos de radar, satélite y registros en superficie

Debido a que la información de la lluvia que puede aprovecharse para analizar la distribución de un evento generalmente está disponible bajo distintos esquemas de resolución espacial y temporal, la integración de métodos para combinar distintas bases de datos no es una labor trivial, pero es de gran interés especialmente para mejorar los resultados de los modelos hidrológicos basados únicamente en mediciones puntuales, al reducir la dependencia sobre parámetros estimados con las características generales de una cuenca, o por la utilidad que estas estimaciones tienen para la vigilancia meteorológica y la predicción numérica del clima (simulación de algunas horas y hasta una o dos semanas de anticipación), por ejemplo [50].

Las estimaciones con radares meteorológicos son útiles debido a la calidad de la resolución que brindan, pero su exactitud depende de una buena calibración, ajustada a los registros propios del área de estudio, la distancia entre radares disponibles, la orografía y el tipo de precipitación, entre otros aspectos. Por su parte, las estimaciones proporcionadas por misiones satelitales tienen una gran cobertura y su resolución es conveniente, pero se obtienen a partir de una relación supuesta entre la temperatura en el tope de las nubes y la intensidad de la lluvia, lo que puede generar errores por la naturaleza de los fenómenos analizados (estratiformes o convectivos). No obstante que las estimaciones por satélite suelen representar subestimaciones o falsas lluvias, en comparación con datos registrados o los estimados con radar, dichos valores constituyen un mecanismo importante para recuperar información sobre eventos, especialmente en áreas sin cobertura de radar y baja densidad de observaciones [18].

De acuerdo con Calvetti *et al.* 2017, en general las técnicas de integración de lluvia procuran la remoción de sesgos sistemáticos de las estimaciones de radares y satélites, ya sea con la

combinación de correcciones para el sesgo por adición y multiplicación (Vila *et al.* 2009), utilizando redes de pluviómetros como referencia y correcciones mediante funciones de peso (Rozante *et al.* 2010), o el empleo de interpolaciones geoestadísticas para combinar información (Nanding *et al.* 2015).

En su trabajo, Calvetti *et al.* 2017 describen un procedimiento para integrar las estimaciones de radar y de satélite con los registros de una red de pluviómetros mediante el uso del algoritmo Siprec (Sistema de precipitación integrada, en portugués), que utiliza un método numérico multicelda para solucionar una ecuación de Poisson [38]. El algoritmo de combinación es similar a las técnicas de asimilación de datos al utilizar una condición de frontera como primera suposición para combinar los valores. Para incluir las tres bases de datos, el proceso se realiza en dos etapas: primero la red del radar (función de origen) es combinada con los datos de lluvia puntual (condición de frontera), suponiendo que la forma satisface la ecuación de Poisson; después, la estimación del satélite se agrega con la ecuación de Poisson, teniendo como condición de frontera al campo definido en la primera combinación. De esta manera, la solución de la ecuación actúa directamente sobre los datos, lo que implica una ventaja por evitar así el cálculo recurrente de coeficientes de peso.

La evaluación de la metodología se realizó considerando el análisis de eventos de lluvia ocurridos entre los años 2013 y 2014 sobre la superficie de subcuencas localizadas al sureste de Brasil y próximas a Paraguay. Se obtuvieron resultados de comparación y validación cruzada, y los valores estimados con el Siprec mostraron mejoría en el valor de la raíz del error cuadrático medio, coeficiente de correlación más alto y desviaciones estándar más cercanas a los valores observados. La mayoría de los errores calculados corresponden a eventos con intensidad superior a 10 mm/h, asociados más probablemente a la gran variabilidad espacial propia de eventos convectivos. Por ejemplo, en general se observó que el radar subestima la lluvia, con algunos picos respecto a los valores registrados; el satélite también produjo varios picos, e inclusive falló en la detección de lluvias, mientras que el Siprec siguió de manera más consistente la variación de los registros, aunque con errores respecto a las observaciones. Entonces, al considerar los datos pluviométricos como condiciones de frontera, el análisis de sistemas que consideran redes de baja densidad es muy sensible a la calidad de los datos de percepción remota.

## 2.6. Aspectos de interés para la selección de un procedimiento de combinación y su empleo

Respecto a la naturaleza de la precipitación que prevalece en México, se sabe que es de tipo pluvial, pero tiene una distribución irregular como consecuencia de las características geográficas de su superficie, de modo que, con respecto a la latitud, puede apreciarse un incremento de la lluvia registrada, del norte hacia el sur del país.

De manera tradicional, la observación de la lluvia sobre el territorio nacional se realiza mediante una red de aproximadamente 5 500 pluviómetros, además de 189 estaciones meteorológicas automáticas [89]. Por esta razón, para analizar el comportamiento de la lluvia sobre una superficie generalmente se han considerado este tipo de registros, localizados dentro o en la proximidad de una zona de interés, y aunque dicha información se considera confiable, estos valores son representativos únicamente sobre la vecindad del punto de registro, y son insuficientes para conocer la variabilidad espacial de la lluvia sobre toda la superficie de captación.

Dada la baja densidad de sitios de medición, es necesario un aprovechamiento más eficiente de los avances tecnológicos desarrollados a nivel mundial. La implementación de procesos de percepción remota, como es el caso de los productos derivados de radares meteorológicos y sistemas satelitales, permiten conocer la variación espacial y temporal de una tormenta, con resoluciones adecuadas al tamaño de una zona de estudio.

En este trabajo de investigación se considera la posibilidad de realizar mejores estimaciones de la distribución espacial de la lluvia mediante el uso de datos disponibles en diversas plataformas, accesibles a través de distintos medios, que incluyan las observaciones diarias de estaciones pluviométricas y pluviográficas, además de las lluvias estimadas por medio de radares meteorológicos y las obtenidas como productos de sistemas satelitales. Con esta información se estudia sobre la combinación de los datos que permita determinar una distribución bidimensional de valores de lluvia sobre las celdas de una malla regular que representa una superficie de estudio.

De acuerdo con lo expuesto en este capítulo, la asimilación de datos se ha convertido en una herramienta para la estimación del estado de un sistema que varía a través del tiempo, tanto a nivel de diagnóstico como de pronóstico. De entre los esquemas numéricos de primera generación se distingue al Ajuste de funciones (Panofsky 1949; Gilchrist y Cressman 1954), al método de Correcciones sucesivas (Bergthorsson y Doos 1955; Cressman 1959; Barnes 1964, 1973), y a la Interpolación óptima (Eliassen 1954; Gandin 1963), los cuales son métodos objetivos de análisis espacial bidimensional de observaciones, para el diagnóstico de un estado sobre los puntos internos de una malla regular [79].

En la práctica, la aplicación simple del ajuste a superficies polinómicas para la representación de un campo en cualquier punto, tiene como principal inconveniente que un ajuste mal condicionado de polinomios de orden superior suele causar oscilaciones ilegítimas en la presencia de errores observacionales, así como por escasez de información en regiones con baja densidad de observaciones, y por ello conviene limitar su empleo a polinomios de orden bajo en sistemas de dimensiones reducidas [78].

Respecto a los procedimientos de correcciones sucesivas, su aplicación se recomienda para el análisis de regiones de gran dimensión con numerosos registros de información y densidad no heterogénea. Asimismo, presentan una sensibilidad reducida a errores observacionales cuando éstos no están correlacionados horizontalmente, como sucede con los datos de superficie y de radio sondeo. Como desventajas, estos métodos tratan a las observaciones como independientes una de otra, y la influencia que se otorga a cada observación en el cálculo de los valores en la malla analizada depende únicamente del uso de coeficientes de peso como función de la distancia entre el punto de malla y la observación, por lo que la magnitud de los datos más próximos puede ser determinante, en especial si el espaciamento entre observaciones es irregular [78].

En cuanto a la interpolación óptima, también permite la estimación de estados mediante funciones de peso, pero éstas son específicas de la configuración de la red de observaciones, es decir, considera la posición de cada dato respecto a los demás, al basarse en las propiedades espaciales estadísticas de las diferencias entre las observaciones y un campo básico, lo que es conveniente al analizar zonas con distribución heterogénea de la información, pero su aplicación requiere la disponibilidad de series históricas de datos, así como la solución de sistemas de ecuaciones lineales mediante la inversión de matrices de correlaciones cruzadas entre estaciones (de dimensión igual al número de observaciones), lo que puede ser una limitante [78].

Debido a que con el paso del tiempo algunas redes de observaciones meteorológicas se han incrementado, la información disponible ha dado lugar al surgimiento de una nueva forma de análisis objetivo y procedimientos de inicialización, conocido como Asimilación de datos continua. Los procesos de análisis que consideran observaciones en instantes sucesivos están orientados a la inicialización de modelos numéricos de predicción. Entonces, a partir de la década de los 80 comenzó la implementación más general de procedimientos como el Filtro de Kalman (1960) y la Asimilación variacional (Sasaki 1958), que han alcanzado gran aceptación para la predicción numérica del clima, especialmente para la generación de campos analizados en espacios continuos, utilizando observaciones en escalas no sinópticas [79].

Como el interés principal de este trabajo de tesis es proponer una metodología que permita mejorar la estimación del comportamiento histórico de un campo original de lluvia distribuida espacialmente en México, para su aprovechamiento en la modelación de escurrimientos superficiales, se consideró valioso tomar en cuenta las experiencias compartidas en trabajos de investigación orientados a definir el estado de diagnóstico de un sistema como una representación discreta de valores localizados sobre las celdas de una malla regular; por ejemplo:

- el hecho de que la contribución de los valores agregados por percepción remota pierde importancia en zonas con alta densidad de observaciones [71].
- que generalmente existe una mayor incertidumbre en la estimación de la lluvia asociada a eventos de naturaleza convectiva, por su variabilidad espacial [52].
- que una manera de enfrentar el problema de incompatibilidad en la combinación de datos con escalas espaciales distintas es mediante el suavizado de las estimaciones de percepción remota, al considerar la influencia pesada de las observaciones próximas [52].
- que el tipo de función Kernel empleada en el modelo de análisis para establecer el peso de las observaciones puede no ser tan crítica, pues funcionan de manera similar y las diferencias en las estimaciones representan hasta un 5% en el cálculo del error cuadrático medio [52].
- que la calibración electrónica inadecuada de un radar o la determinación equivocada de su ecuación generan estimaciones deficientes, las que pueden mejorar si se identifica la magnitud del error multiplicativo que las afecta [59].
- que aun cuando las estimaciones de satélite pueden ser de menor calidad, su utilidad es excepcional cuando la cobertura de radares está limitada [18].
- que la dependencia respecto a la configuración de las observaciones es menor cuando la densidad de registros es más grande [59].

## 3. Programación computacional para los procesos de asimilación

### 3.1. Aspectos generales de la modelación

Se conoce por modelación al mecanismo que permite la interacción entre un objeto de estudio y el sujeto que lo analiza, donde el primero es reemplazado por un modelo que cumple condiciones de semejanza con el objeto real, y por lo regular esto se practica en aquellas situaciones en las que el análisis del objeto es inviable, por razones de costo (inversión en infraestructura, equipo, mano de obra y tiempo) o riesgo, por ejemplo [17]. La alternativa a la investigación convencional son los modelos de simulación, como representaciones simplificadas de los procesos que ocurren en un sistema real, donde interactúan variables que evolucionan mostrando un comportamiento dinámico [77].

Para la construcción del modelo equivalente al objeto se requiere establecer la existencia de una causa que promueva un cambio de estado, dentro de un dominio abstracto o sistema (flujo de materia, energía o información, conocido como proceso) que puede expresarse como un conjunto de elementos interrelacionados entre sí, compuesto por la entrada, la función de transferencia y la salida [17].

En particular, la modelación matemática consiste en el reemplazo de un objeto por su imagen matemática (modelo matemático), con la implementación de algoritmos lógicos que permiten analizar las cualidades del proceso original, y consta de lo siguiente: el modelo, que expresado con formulaciones matemáticas, describe el proceso en base a conocimientos antecedentes, sin distorsionar sus características básicas originales; el algoritmo de cálculo, que permite el análisis y la solución del modelo en un ordenador, y el programa computacional, que traduce al modelo y algoritmo en un lenguaje de ordenador, para obtener resultados satisfactorios con la exactitud requerida, mediante métodos numéricos [17].

La estructura de los modelos matemáticos es tal que posibilita la transformación de las entradas (son las causas representadas como variables, las cuales activan al sistema) en salidas (representadas por un vector que caracteriza al estado del sistema).

De manera general, el modelo está conformado por: un operador que establece las reglas de correspondencia entre las entradas y salidas, por un conjunto de parámetros cuyas magnitudes describen las características físicas y funcionales del sistema, además de las variables de estado que representan las condiciones espaciales del sistema a través del tiempo [17].

De acuerdo con sus características, los modelos matemáticos pueden clasificarse en: estocásticos o deterministas, si toman en cuenta o no la incertidumbre del proceso que se analiza; dinámicos o estáticos, si consideran o no la evolución del proceso a través del tiempo, y distribuidos o

concentrados, de acuerdo con la cantidad de dimensiones en las que se pueden modificar tanto los parámetros del modelo como las variables de estado [17].

En la modelación matemática se busca la solución de problemas de tipo directo e inverso. Un problema se considera directo cuando lo único desconocido son las salidas, por lo que las entradas, las variables de estado, los parámetros y el operador son conocidos. Por su parte, el problema es inverso cuando se presenta una de las circunstancias siguientes: calibración, para encontrar los parámetros que permitan, mediante las entradas conocidas y el operador matemático, obtener las salidas observadas; reconstrucción, para establecer las causas (entradas desconocidas) que produjeron ciertos efectos (salidas), y la propuesta del operador matemático, para establecer las reglas de correspondencia (ecuaciones) que permitan la transformación de las entradas en salidas [17].

Debido a la gran utilidad que la modelación puede tener en la gestión de los recursos hídricos, por ejemplo, es muy importante que durante la toma de decisiones el modelador pueda ejercer control en el manejo de la información, lo que permita reducir la magnitud de los errores de estimación [17].

En el caso del análisis de la transformación de la precipitación en escurrimiento, lo apropiado es el desarrollo de una secuencia de pasos que inicia con la preparación de los insumos requeridos. La primera etapa consiste en la recopilación y almacenamiento de datos pluviométricos, derivados de las principales fuentes de información: registros de estaciones pluviométricas, así como estimaciones de lluvia obtenidas con el empleo de mediciones de radares meteorológicos y sensores satelitales, en consideración de la disponibilidad y calidad de dicha información. Posteriormente, en función de las características de las bases de datos, puede requerirse un pre procesamiento de los mismos, que implique su importación y adecuación, para estar en disposición de eventos acumulados diarios de precipitación, por ejemplo [85].

Para caracterizar el fenómeno de la precipitación se requiere del análisis de un proceso que por naturaleza es espacialmente distribuido; por esta razón, la metodología a seguir recomienda la revisión preliminar de los datos, para observar el cumplimiento de ciertos requisitos, principalmente que su distribución espacial sea homogénea, y que el conjunto no se vea afectado por valores atípicos [9].

Como consecuencia de la distribución espacial irregular y dispersa de las observaciones de la red de registros pluviométricos sobre el territorio de la República Mexicana, la evaluación numérica de la precipitación representa un desafío, pues comúnmente la disponibilidad de la información requerida tiende a concentrarse sobre los centros de mayor densidad de población. Entonces, una distribución no homogénea de registros puede tener como consecuencia que el cálculo de lluvia mediante procedimientos de interpolación sobre mallas regulares sea de calidad insuficiente, y el nivel de confianza correspondiente a estos valores comprometa su aplicación [71].

Con la intención de mejorar la calidad de las estimaciones de lluvia distribuida espacialmente, en este trabajo se ha optado por considerar también los valores de lluvia obtenidos mediante instrumentos de percepción remota, cuyas resoluciones espaciales y temporales permitan la construcción de campos de lluvia con la aplicación de técnicas para la combinación de bases de datos.

La identificación de los valores atípicos en una muestra puede influir de manera positiva en los resultados del análisis del comportamiento de la lluvia, especialmente si en la revisión logra detectarse algún valor discordante, relacionado con el mecanismo de medición, o inclusive con un

error humano. Cabe señalar que los valores atípicos pueden clasificarse en distribucionales, a identificarse porque su magnitud se aleja significativamente respecto al valor medio de la muestra, y los espaciales, cuya magnitud difiere sensiblemente con la de sus vecinos más próximos [9].

Una manera de proceder es mediante el análisis estadístico de los eventos históricos de tormenta registrados por las estaciones pluviométricas consideradas en el área de estudio, al identificar por ejemplo sus valores diarios máximos anuales, y con ello obtener su correspondiente valor medio, desviación estándar, y sus extremos máximo y mínimo. Con estos parámetros estadísticos podría realizarse una depuración de lluvias de eventos particulares, como pueden ser valores muy pequeños (inclusive nulos) de las precipitaciones máximas en todo un año. Para valores muy grandes (tres a cuatro veces superiores a la media), generalmente es de gran utilidad verificar la precipitación simultánea de láminas grandes en las estaciones más cercanas (del orden de dos veces la media), [30].

La verificación de la calidad de las lluvias obtenidas por percepción remota respecto de los datos registrados con pluviómetros durante un mismo evento y área de captación común, puede practicarse a través de su coeficiente de correlación, por tratarse de una medida adimensional del nivel de asociación lineal que existe entre ambos conjuntos [85].

Así mismo, dado que la construcción del diagrama de dispersión entre las lluvias de pluviómetro y radar (así como de satélite) permite conocer la relación entre los conjuntos alrededor de la línea de error nulo, también es posible la identificación de los intervalos de lluvia en donde existe una menor o mayor correlación, y por lo tanto pueden asociarse directamente con condiciones de subestimación o sobreestimación, por lo que en caso necesario es factible la eliminación de datos.

Adicionalmente, este mismo análisis puede ayudar a establecer la conveniencia de un procedimiento de combinación de diferentes bases de datos, pues si la correlación entre conjuntos es pobre, es posible que la variable secundaria no aporte información significativa a la estimación [9].

Después de revisar las características de la información de lluvia disponible, se procede a una etapa de calibración del modelo, que consiste en la asignación de valores a sus parámetros conceptuales, a través de una serie de escenarios de simulación hasta obtener un ajuste aceptable entre valores simulados y observados, proceso que puede llevarse a cabo por mecanismos de prueba y error, optimización, o la combinación de ambos [77].

En la solución de problemas hidrológicos y climáticos, generalmente se considera el uso de modelos determinísticos, paramétricos o estocásticos. Un modelo determinístico es aquel que se obtiene mediante relaciones físicas, sin el requerimiento de datos experimentales. El modelo es estocástico si el empleo repetido de un grupo de datos de entrada genera salidas predecibles solamente en un sentido estadístico. Por lo anterior, un modelo paramétrico puede considerarse determinístico si definidos sus parámetros, el empleo repetido de un grupo de datos de entrada genera la misma salida, mientras que un modelo paramétrico es estocástico cuando cada estimador de un parámetro es función de los valores muestrales [81].

Como un procedimiento para valorar la influencia de cada parámetro de un modelo en algún aspecto de su salida, el análisis de sensibilidad permite evaluar su respuesta a través de la modificación ordenada de los parámetros, dentro de un rango determinado de valores, según las necesidades de la calibración, siendo una práctica factible de realizar independientemente de la complejidad del modelo [81]. Mediante un conjunto de entradas conocidas y el operador matemático (que establece las reglas de correspondencia), en la calibración se procura la

obtención de resultados comparables con las correspondientes salidas, también conocidas, es decir, se busca la puesta a punto del modelo, para evitar una eventual obtención de resultados muy alejados de la realidad [17].

Con la aplicación del modelo y el cálculo de resultados, se procede a la etapa de validación, para evaluar la calidad del modelo al contrastar los valores simulados de una determinada variable, respecto de las magnitudes reales observadas [77]. Existen diferentes criterios para la evaluación del desempeño, los que pueden ser de naturaleza subjetiva, como los derivados de una apreciación visual, o los de evaluación objetiva, donde la calidad de los ajustes es mensurable cuantitativamente [17].

La validación de modelos más sencilla suele practicarse mediante el control visual de los resultados, a partir del cual puedan formularse evaluaciones subjetivas respecto al comportamiento sistemático y dinámico de los procesos involucrados, aunque por su naturaleza el nivel de detalle de la inspección dependerá principalmente de la escala de representación de los resultados. Tal es el caso de la identificación de las tendencias en el comportamiento de una variable, o la detección de desfases temporales de los valores extremos [17].

Para una evaluación más objetiva se recomienda el empleo de criterios de eficiencia para la medición matemática del error que existe entre los valores observados y los estimados de las variables modeladas [81]. En la validación del modelo propuesto para la estimación de lluvia distribuida espacialmente, se consideró conveniente utilizar diferentes criterios estadísticos, en particular el error medio (*ME*), el error medio absoluto (*MAE*), la raíz del error cuadrático medio (*RMSE*), el coeficiente de eficiencia de Nash-Sutcliffe (*NSE*), así como el coeficiente de correlación (*CC*).

El cálculo de los coeficientes para la validación de resultados de las lluvias estimadas ( $P_e$ ) respecto de las observadas ( $P_o$ ), se realizó para  $N$  observaciones de validación, de acuerdo con:

$$ME = \frac{\sum_{i=1}^N (P_e - P_o)}{N} \quad (3.1)$$

$$MAE = \frac{\sum_{i=1}^N |P_e - P_o|}{N} \quad (3.2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (P_e - P_o)^2}{N}} \quad (3.3)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (P_e - P_o)^2}{\sum_{i=1}^N (P_o - \bar{P}_o)^2} \quad (3.4)$$

$$CC = \frac{\sum_{i=1}^N (P_e - \bar{P}_e)(P_o - \bar{P}_o)}{\sqrt{\sum_{i=1}^N (P_e - \bar{P}_e)^2} \sqrt{\sum_{i=1}^N (P_o - \bar{P}_o)^2}} \quad (3.5)$$

El concepto de sesgo se utiliza para hacer referencia de ciertos aspectos estadísticos de un conjunto de mediciones realizadas, de muestras obtenidas o de estimaciones calculadas. En el caso de mediciones, el sesgo es la diferencia entre una media poblacional de las mediciones y un valor de referencia aceptado como verdadero, razón por la cual el sesgo implica una sub o sobreestimación de la realidad.

El sesgo en las mediciones se debe principalmente al uso incorrecto de aparatos o procedimientos, por lo que no desaparece al incrementarse el tamaño de la muestra. El sesgo de una muestra está

asociado a una porción de la misma que no es representativa de la población de estudio, y tampoco desaparece al incrementarse la muestra [86].

El sesgo de una estimación, conocido también como error sistemático, se refiere a un método en el cual el promedio de estimaciones repetidas en un proceso se desvía respecto al valor verdadero, aunque ha de reducirse al aumentar el tamaño de la muestra, lo que representa una característica conveniente de cualquier estimador. Entonces, un buen estimador debería ser insesgado, de modo que incluso bajo una distribución que sub o sobreestime, el sesgo pueda reducirse hacia un valor nulo.

Una medida común para la medición del sesgo es el error medio  $ME$ , que se obtiene como la media de todas las diferencias entre los valores estimados y los verdaderos, por lo que indica si dentro del proceso el estimador genera una sub o sobreestimación de manera consistente [86].

El error medio absoluto  $MAE$  también se considera como un indicador principal del desempeño del modelo, especialmente porque es menos sensible en la penalización de valores atípicos o errores de mayor magnitud [59].

Cabe señalar que un método simple para ajustar un campo de lluvias obtenido a partir de percepción remota es la corrección del sesgo medio ( $MFB$ ), tomando en cuenta que la magnitud de estas lluvias puede estar afectada por un error multiplicativo uniforme, originado por la calibración electrónica deficiente de un aparato como el radar meteorológico, por ejemplo. El factor de corrección  $MFB$  se obtiene como el cociente de la suma de las lluvias registradas por los pluviómetros disponibles en el área de interés, y la suma de las lluvias obtenidas con radar (en la ubicación correspondiente a los pluviómetros considerados), para un mismo evento de lluvia [59].

Por otra parte, una estimación adecuada también se caracteriza por su exactitud, de manera que su magnitud debe ser tan cercana al valor verdadero de la variable simulada como sea posible. En este sentido, los estadísticos típicos para medir la exactitud generalmente toman en cuenta la diferencia entre las estimaciones y el conjunto de valores observados, considerando el cuadrado de esta diferencia o su valor absoluto, para darle importancia solo a su magnitud, al eliminar la dirección de la diferencia.

Una medida común de la exactitud es el error cuadrático medio ( $MSE$ ), calculado como la media del cuadrado de las diferencias, por lo que representa la suma de la varianza de las estimaciones más el cuadrado de su error medio  $ME$ . Debido a que el  $MSE$  se obtiene mediante el cuadrado de las diferencias, este resultado no está expresado a la misma escala que las mediciones originales, por lo que se recomienda el cálculo de su raíz cuadrada, para obtener el  $RMSE$  y representar entonces la desviación standard de los residuos, que penaliza especialmente los errores de mayor magnitud [86].

El criterio propuesto por Nash y Sutcliffe sugiere el cálculo del coeficiente  $NSE$ , que a diferencia del  $RMSE$  es un indicador adimensional de la bondad de ajuste, y se refiere a la fracción de la varianza que puede ser explicada por el modelo, sin que resulte afectada por un cambio de unidades [52]. Aunque este coeficiente varía entre  $-\infty$  y 1, la estimación perfecta tiene un  $NSE$  igual a 1, y puede clasificarse como [55]: muy bueno si es igual a 0.90, bueno de 0.80 a 0.90, aceptable de 0.65 a 0.80, e insatisfactorio si es menor que 0.65, además de que una eficiencia igual a cero implica que la exactitud del estimador es equivalente al cálculo del valor medio de las observaciones registradas por la red pluviométrica de un área de estudio, en el caso de un procedimiento de asimilación de datos de lluvia, por ejemplo [59].

Adicionalmente, el coeficiente de correlación  $CC$ , como una medida adimensional del nivel de asociación lineal entre estimaciones y observaciones, se calculó para observar qué tanta dispersión puede existir entre las magnitudes de lluvia reales registradas y las estimadas mediante procedimientos de interpolación y asimilación de datos, sin perder de vista el hecho de que un modelo que sobre o subestime sistemáticamente puede proporcionar valores satisfactorios del  $CC$ , alejados de cero y próximos a la unidad, aunque todas las estimaciones estuvieran erradas. Lo anterior justifica la necesidad de considerar diferentes criterios de rendimiento para la validación de un modelo que permita reproducir el comportamiento histórico o futuro de una variable [81].

En la evaluación de la calidad del modelo propuesto para la estimación de lluvia distribuida espacialmente, calculada a partir de la combinación de las observaciones disponibles (láminas diarias registradas por pluviómetros y las obtenidas mediante percepción remota), es posible realizar un procedimiento de validación cruzada, que consiste en evaluar el error de la estimación calculada en cada punto donde exista un registro histórico de lluvia, y con el error en cada punto obtener un estadístico global que caracterice la calidad del ajuste, utilizando los criterios de eficiencia antes descritos. El procedimiento debe aplicarse sobre cada uno de los sitios con registro, lo que implica realizar tantos procesos de interpolación como puntos de registro se disponga, retirando en cada proceso un solo valor real del grupo, interpolando y calculando el error correspondiente al punto de registro que no participó en el proceso [2].

Una opción más económica en términos de tiempo de proceso consiste en calcular los términos de error con una sola interpolación, al retirar de entre los registros disponibles un porcentaje de datos seleccionados aleatoriamente (del orden del 10%), o un conjunto escogido que cumpla una condición de interés [71].

En la evaluación del modelo por validación cruzada se recomienda considerar que la magnitud de los errores está relacionada con el grado de suavización alcanzado, como consecuencia de la cantidad de puntos utilizados en el proceso de interpolación, es decir, a mayor número de puntos el suavizado es más significativo y se reduce el error. Así mismo, la distribución espacial de los puntos tiene importancia en la valoración de los términos de error debido a aglomeraciones y redundancias, por lo que previamente podría aplicarse algún procedimiento de ponderación que disminuya la influencia de las aglomeraciones [2].

### 3.2. La estimación de lluvia mediante un modelo de Asimilación de datos

Como se ha manifestado, la baja densidad de redes de medición de lluvia sobre la superficie nacional es una condición desfavorable para la práctica satisfactoria de procedimientos que requieran del conocimiento de la distribución espacial de esta variable, como es el caso de los métodos de análisis de la transformación de la lluvia en escurrimiento. Por tal motivo, en este trabajo de tesis se sugiere la consideración de diferentes bases de datos de lluvia para estar en posibilidad de aprovechar la mayor cantidad de información disponible, y realizar con ello mejores estimaciones de su distribución espacial mediante el uso de un modelo de Asimilación.

El objetivo de los esquemas de asimilación secuencial o variacional de datos meteorológicos es la representación regular del estado de la atmósfera a una resolución adecuada, hasta en cuatro dimensiones, a partir de una muestra irregular e imperfecta de observaciones en el espacio y el tiempo, originada de un arreglo heterogéneo de instrumentos in situ y de percepción remota, de

manera que para cualquier estado inicial y de entrada de datos, la modelación del sistema haga posible definir sus estados en el futuro.

Para el estudio del estado del tiempo y los fenómenos generados como la lluvia se utiliza el concepto de Análisis, siendo éste la obtención de la imagen correspondiente a la realidad del estado atmosférico de un sistema en un momento determinado, a través de una recopilación de números, es decir, se considera como un diagnóstico consistente de la realidad atmosférica porque se basa en una colección de observaciones realizadas sobre el estado del sistema, y puede obtenerse mediante modelos matemáticos donde el campo esté representado por valores de la variable localizados sobre las celdas de una malla.

Existe una diversidad de esquemas secuenciales útiles a nivel operacional, los que se diferencian por el nivel de detalle del análisis. Para un problema de diagnóstico, la aplicación de los esquemas de correcciones sucesivas resulta práctica porque son procedimientos iterativos donde la asignación de pesos a las innovaciones no se realiza de manera óptima, pues el suavizado de las observaciones consideradas es solo función de la distancia que existe entre ellas.

Uno de los métodos de correcciones sucesivas más utilizado es el propuesto por Barnes, quien desarrolló un esquema de Análisis objetivo de interpolación convergente, que utiliza una función de pesos Gaussiana en el ámbito espacial para el análisis de una variable atmosférica aleatoriamente distribuida.

Debido a los antecedentes satisfactorios de su aplicación en el análisis de tormentas, para la estimación de campos de lluvia distribuida espacialmente se optó por el empleo del procedimiento de correcciones sucesivas propuesto por Barnes, para aprovechar la confiabilidad de los registros diarios de la red de pluviómetros nacional, así como el conocimiento de la estructura espacial de las tormentas derivada del empleo de sistemas satelitales y radares meteorológicos, tomando en cuenta que las principales ventajas de su empleo son las siguientes [26, 78]:

- Algoritmo simple para ponderar gran número de observaciones con distribución espacial no necesariamente uniforme,
- Procedimiento de interpolación en dos pasos, por modificación de Barnes (1973),
- Adecuado para el análisis de tormentas, como fenómenos en escalas sub sinóptica y mesoescala, ampliamente utilizado con datos de radar y satélite,
- No requiere de un radio de influencia para la consideración de observaciones,
- Posible modificación de parámetros de escala para la ponderación,
- Presenta sensibilidad reducida a errores observacionales,
- Opción económica en términos de tiempo de proceso, con resultados comparables con métodos más sofisticados, si la densidad de observaciones no es muy baja.

En el Anexo 1 se describe la metodología propuesta para Asimilación de tres bases de datos.

### 3.3. Programación del método de Barnes

Para acelerar el cálculo de resultados correspondientes a la aplicación de una metodología de asimilación de datos, se consideró necesario el desarrollo de una herramienta basada en el método de Barnes. De esta manera se busca realizar procesos de interpolación de forma rápida y expedita, para la combinación de tres bases de datos de precipitación disponibles (conjunto de observaciones que incluye a los registros pluviométricos y las estimaciones de satélite y radar meteorológico), y con ello obtener valores de lluvia en cada celda de una malla regular, los cuales

sean aproximaciones de las láminas que se precipitaron durante un evento ocurrido sobre una superficie de captación. El lenguaje computacional utilizado para realizar la programación del método es MATLAB R2017a.

Los datos requeridos como insumos del programa son los siguientes:

- Conjunto de observaciones de precipitación acumulada diaria (registros de pluviómetros y estimaciones como productos de satélite y radar), asociadas a ubicaciones puntuales dentro de la malla regular de interés,
- Dimensiones del área de estudio, definida por el tamaño de las celdas cuadradas que constituyen a la malla,
- Parámetro de potencia  $\beta$ ,
- Parámetro de convergencia numérico  $\gamma$ .

### 3.3.1. Primera etapa

En esta parte se realiza una primera estimación de lluvia  $S_g^1$  sobre cada una de las celdas de la malla, con el empleo de la ecuación (2.14).

Las variables que intervienen en el cálculo de la primera estimación son:  $S(x_m, y_m)$ , que es la variable observada en la posición  $m$  (deben considerarse todas las observaciones  $N$  contenidas en la celda que se está analizando);  $w_m$ , que es la función que permite el cálculo del peso que tiene cada una de las observaciones con respecto al centroide de la celda de análisis; y  $d_m$ , que corresponde a la distancia que existe entre el centroide y las observaciones contenidas en la celda de análisis, calculada mediante la ecuación de la distancia entre dos puntos:  $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$ .

Como se observa en la Figura 3-1, se cumple la primera etapa del procedimiento cuando se obtiene el valor de la primera estimación  $S_g^1$  correspondiente a cada una de las celdas de la malla.

$S_g^1$	$S_g^1$	$S_g^1$
$S_g^1$	$S_g^1$	$S_g^1$
$S_g^1$	$S_g^1$	$S_g^1$

Figura 3-1: Valores de lluvia  $S_g^1$  obtenidos en cada celda de la malla regular, como resultado de la primera etapa en el método de Barnes.

### 3.3.2. Segunda etapa

Se realiza una segunda estimación de la lluvia, donde para cada una de las celdas de la malla analizada se modifica la estimación inicial  $S_g^1$  de la primera etapa.

Las variables que permiten llevar a cabo la segunda estimación están contenidas en la ecuación (2.16). Cada estimación  $S_g^1$  de la primera etapa se modifica con la adición de un cociente, el cual depende tanto de una función de peso suavizada ( $w'_m$ ) con un parámetro de convergencia ( $\gamma$ ), como de las diferencias entre cada observación ( $S(x_m, y_m)$ ) y el valor ( $S^1(x_m, y_m)$ ) que le

corresponde como una estimación calculada mediante la interpolación de las lluvias  $S_g^1$  que le son adyacentes.

De acuerdo con el procedimiento, en el cálculo de  $S^1(x_m, y_m)$  se requiere de una interpolación Bilineal para el análisis de las celdas interiores de la malla (Figura 3-2). Sin embargo, cuando la información es insuficiente, como ocurre en el análisis de las celdas periféricas (Figura 3-3), para el cálculo de  $S^1(x_m, y_m)$  conviene aplicar una interpolación más simple, como lo sugiere Koch (1981), y en este trabajo se optó por utilizar el método de interpolación IDW (Ponderación de la Distancia Inversa).

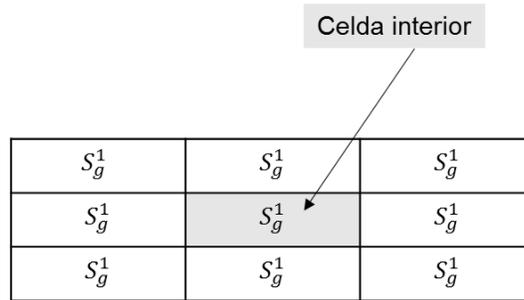


Figura 3-2: Celda interior de una malla.

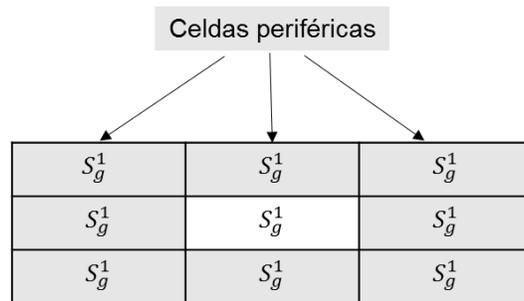


Figura 3-3: Celdas periféricas de una malla.

Para aplicar el método de interpolación IDW en celdas periféricas, se utilizan las expresiones que aparecen a continuación:

$$S^1 = \sum_{m=1}^N \lambda_m S_g^1 \tag{3.6}$$

$$\lambda_m = \frac{\left(\frac{1}{D_m}\right)^\beta}{\sum_{m=1}^N \left(\frac{1}{D_m}\right)^\beta}$$

donde  $D_m$  representa las distancias que existen entre las observaciones contenidas en la celda de análisis y los centroides de las celdas que le son adyacentes;  $\beta$  es el parámetro de potencia que controla la caída del peso  $\lambda_m$  respecto a la distancia, donde valores más grandes del parámetro incrementan la influencia de los puntos más cercanos [87].

La Figura 3-4 es una representación matricial de los valores  $S_g^1$  considerados para estimar la lluvia correspondiente a la ubicación de cada observación que esté localizada dentro de una celda periférica de análisis, aplicando el método de interpolación IDW.

Para el método IDW se tomarán en cuenta todas las celdas adyacentes a la celda que se esté analizando (Figura 3-4), y el peso  $\lambda_m$  debe calcularse para cada observación contenida en la celda de análisis. Por esta razón, son tres valores  $\lambda_m$  por cada observación contenida en cada una de las celdas de las cuatro esquinas de una malla regular. Así mismo, son cinco valores  $\lambda_m$  por cada observación contenida en cada una de las celdas periféricas restantes.

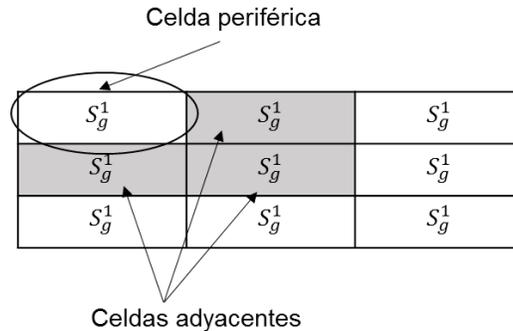


Figura 3-4: Identificación de una celda periférica y las celdas que le son adyacentes para realizar una interpolación IDW.

Para el cálculo de  $S^1(x_m, y_m)$  en las celdas interiores de la malla (no periféricas, Figura 3-5), se aplica el método de interpolación Bilineal, útil especialmente para el análisis de funciones de dos variables sobre un plano bidimensional, porque para cada posición de interés se utilizan los 4 valores más próximos situados en direcciones diagonales, lo que implica interpolaciones lineales tanto en la dirección horizontal (eje  $x$ ) como en la vertical (eje  $y$ ), siendo el resultado independiente del eje que se interpole primero.

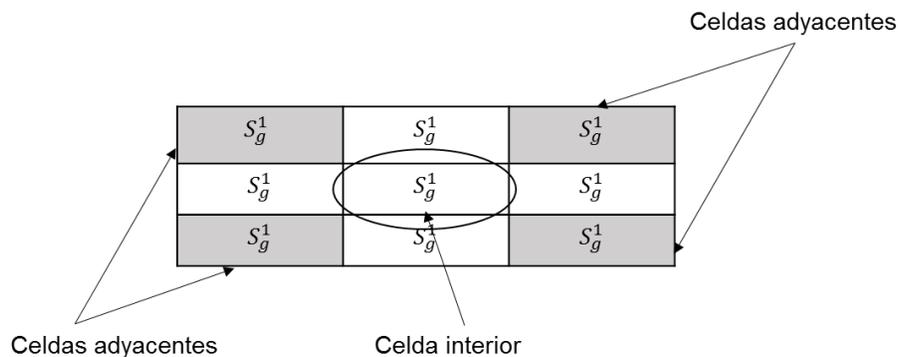


Figura 3-5: Identificación de una celda interior y las celdas adyacentes consideradas para realizar una interpolación Bilineal.

Las expresiones que se involucran en el cálculo de  $S^1(x_m, y_m)$ , considerando la distribución espacial de las observaciones mostradas en la Figura 3-6, son las siguientes:

$$I_{x_i, y_2} = \frac{x_2 - x_i}{x_2 - x_1} (V_{x_1, y_2}) + \frac{x_i - x_1}{x_2 - x_1} (V_{x_2, y_2}) \quad (3.7)$$

$$I_{x_i, y_1} = \frac{x_2 - x_i}{x_2 - x_1} (V_{x_1, y_1}) + \frac{x_i - x_1}{x_2 - x_1} (V_{x_2, y_1}) \quad (3.8)$$

$$I_{x_i, y_i} = \frac{y_2 - y_i}{y_2 - y_1} (I_{x_i, y_2}) + \frac{y_i - y_1}{y_2 - y_1} (I_{x_i, y_1}) = S^1(x_m, y_m) \quad (3.9)$$

El valor estimado  $S^1(x_m, y_m)$ , se obtiene para cada observación contenida en la celda interior analizada ( $I_{x_i, y_i}$ ). Los valores  $V_{x, y}$  son las lluvias  $S_g^1$  calculadas en la primera etapa, justamente las que se encuentran alrededor de la celda que se está analizando (Figura 3-6). Por su parte, las coordenadas de los centroides de las celdas adyacentes están representadas por  $x_1, x_2, y_1, y_2$ , mientras que  $x_i, y_i$  son las coordenadas de cada una de las observaciones contenidas en la celda de análisis.

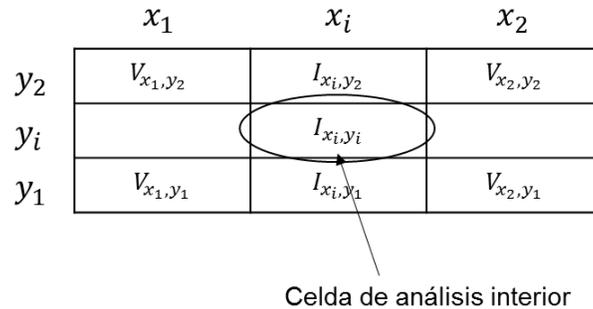


Figura 3-6: Representación de las variables que intervienen en las expresiones de cálculo de interpolación Bilineal, en una malla regular.

La programación del método de Barnes se realizó como una secuencia de ciclos en lenguaje de MATLAB R2017a. Previo a la ejecución del programa, es necesario importar un conjunto de archivos (hojas de cálculo con extensión .xls) que contienen las coordenadas de la ubicación de los centroides de las celdas de la malla regular, así como las coordenadas y láminas de lluvia de todas las observaciones que se utilizarán en el proceso de asimilación. Las coordenadas se encuentran en el sistema UTM (Universal Transverse Mercator, por sus siglas en inglés), para facilitar operaciones durante los procesos, por representar longitudes en metros.

Al ejecutarse el programa (Figura 3-7), éste solicita algunos datos como la longitud horizontal y vertical que cubre la superficie definida por la malla, además de la longitud de cada celda cuadrada. Así se determina el número de renglones y columnas que componen la estructura matricial de la malla de estudio, además de calcularse la magnitud de su superficie. También solicita el número de registros pluviométricos a utilizarse en el análisis de asimilación, por lo que en conocimiento del número de celdas de la malla se define el número de observaciones totales de lluvia disponibles para el análisis. Con esta información se obtiene el espaciamiento entre observaciones, y con ello se determina el valor del parámetro de escala de longitud [74] (raíz cuadrada del parámetro de peso para la respuesta de filtrado, es decir,  $\sqrt{\kappa_0}$ ), el cual interviene en el decaimiento del peso que tiene cada observación, reducción que se advierte principalmente para las observaciones más apartadas (alejadas del umbral definido con el parámetro de escala).

```

Command Window
New to MATLAB? See resources for Getting Started.

Método para calcular precipitación pluvial por método de Barnes
Elaboró: María Asucena Rodríguez Ramírez, Semestre: 2022-1, Programa de Doctorado en Ingeniería, UNAM
Teclea el nombre de la cuenca a estudiar:CDMX
Teclea la longitud de la malla en la dirección horizontal en km:55.65
Teclea la longitud de la malla en la dirección vertical en km:66.78
Teclea la longitud de la celda en km:11.13
El cálculo se va a hacer a tantas celdas horizontales por verticales
6      5

Desea continuar el cálculo con los datos previos o teclear otros:s
Área de estudio km2:
3716.3

Teclea la cantidad de estaciones pluviométricas:72
Número de observaciones:
132
|
Espaciamiento entre observaciones:
5.306

Parámetro de escala de longitud:
7.5924

f> Número de observaciones en cada una de las celdas

```

Figura 3-7: Datos requeridos y cálculos iniciales durante la ejecución del programa.

En un proceso posterior, el programa identifica e informa al usuario respecto del número de observaciones contenidas en cada celda de la malla, y previo al cálculo de la lluvia de la segunda etapa del método, solicita los valores de los parámetros de potencia (para la interpolación IDW) y convergencia (de la función de peso modificada). Finalmente, el programa muestra en un acomodo matricial la lluvia estimada en cada celda, como resultado de la asimilación de las tres bases de datos consideradas, a través del método del Análisis Objetivo de Barnes (Figura 3-8).

```

Command Window
New to MATLAB? See resources for Getting Started.

Número de observaciones en cada una de las celdas
2      3      7      2      3
2      5     11      9      3
5     10     11      6      3
2      5      8      5      3
2      2      5      5      3
2      2      2      2      2

Teclea el valor de beta:2
Teclea el valor de gamma:0.3
SG2 RESULTADO FINAL de lluvia, en mm
14.693      33.888      4.5088      6.8366      1.0417
32.175      24.937      5.8234      7.0644      3.706
35.471      7.7893      10.668      28.58      13.969
27.932      21.591      5.132      18.555      8.9216
16.306      24.983      4.1179      10.467      0.64191
15.525      12.721      2.1459      7.008      3.7437

f> >>

```

Figura 3-8: Lluvia distribuida espacialmente sobre las celdas de la malla de análisis, como resultado final de la ejecución del programa.

La estructura del programa se presenta en el Anexo 2. Cabe señalar que la secuencia de cálculos programada no incluye la medición matemática de errores entre las láminas de lluvia estimadas y los registros considerados para validación.

## **4. Aplicación del modelo de asimilación de lluvia sobre diferentes regiones del territorio nacional**

### **4.1. Modelo de asimilación aplicado**

Como se ha descrito, el comportamiento espacial y temporal de la lluvia generalmente se ha analizado bajo la consideración de registros obtenidos sobre la superficie terrestre, dentro o en la proximidad de las zonas de interés, por lo que las estimaciones obtenidas en regiones con información escasa tienen una componente de incertidumbre fuerte asociada con la distancia entre los puntos de medición.

Por otro lado, la implementación de procesos de percepción remota, como son los productos derivados de radares meteorológicos y sistemas satelitales, ha permitido el estudio de la variación de las tormentas con resoluciones más apropiadas, pero existen inconvenientes respecto a las estimaciones de lluvia así definidas. Por ejemplo, en el caso de los radares se ha identificado falta de exactitud debida a errores como la propagación o recepción inadecuada de la señal por mal funcionamiento del equipo, el bloqueo orográfico, el empleo de longitudes de onda que causan atenuación de la señal, la consideración de lluvia que se evapora durante su caída, o una calibración incorrecta de la relación entre la reflectividad y la intensidad de precipitación [84]. Respecto a los satélites artificiales como plataformas para determinar las trayectorias y características de las nubes precipitables y su relación con la intensidad de lluvia, su empleo supone errores al depender de la medición de ciertos parámetros físicos, y es por ello que dadas las circunstancias de un análisis particular puede resultar más conveniente la estimación derivada del uso de alguna tecnología específica, como el empleo de sensores infrarrojos para analizar sistemas convectivos [72].

Dado que las mediciones de precipitación a nivel de terreno pueden considerarse confiables, pero son representativas solo en la vecindad de los puntos de registro, y los procesos de percepción remota permiten conocer el comportamiento espacial de las tormentas pero pueden tener un nivel de error significativo respecto a los valores registrados simultáneamente, el objetivo principal de este trabajo de tesis es aprovechar las bondades particulares de las bases de datos consideradas, de manera que pueda obtenerse una mejor estimación del comportamiento espacial de la lluvia histórica mediante la combinación adecuada de dicha información.

#### **4.1.1. Entradas del modelo**

Con la intención de discutir las ventajas de aplicar el modelo de asimilación de datos propuesto, considerando tres bases de datos distintas para la estimación del estado de diagnóstico de algún campo histórico de lluvia diaria registrada sobre el territorio mexicano, se realizó una identificación

de eventos de interés dada su magnitud y su distribución espacial, tomando en cuenta que la modelación exige la recopilación de lluvias correspondientes a eventos simultáneos. Para la identificación de eventos representativos se revisó el comportamiento de lluvia registrada con estaciones convencionales y automáticas. Se consideró lo siguiente:

- Realizar el análisis de una región monitoreada extensamente, de manera que la disponibilidad de registros pluviométricos sobre su superficie sea amplia y de longitud conveniente, para estar en posibilidad de practicar ejemplos que consideren distintos esquemas de densidad y configuración de su red de pluviómetros.
- Realizar el análisis sobre diferentes regiones donde la disponibilidad de datos pluviométricos por unidad de superficie sea sensiblemente menor, pero sean de interés por el tamaño de su superficie, localización geográfica que influya en la magnitud de la lluvia prevalente, su cercanía con centros urbanos y la disponibilidad de datos hidrométricos.

Por otra parte, también se recopilaron observaciones meteorológicas de sensores satelitales, considerando lo siguiente:

- Aun cuando el movimiento de rotación de los satélites geoestacionarios respecto a la Tierra resulta apropiado para la vigilancia de los sistemas nubosos, como gran parte de la radiación no penetra profundamente en las nubes la estimación de lluvia desde el espacio se basa en condiciones representativas en la cima o tope de las nubes, y la lluvia desarrollada en zonas más próximas a la superficie solo se calcula mediante relaciones indirectas (OMM, 2011).
- Dado que la lluvia puede estimarse con diferentes algoritmos, y los productos derivados de un solo registro pueden caracterizarse por una resolución que resulte insuficiente, actualmente hay disponibilidad de lluvias estimadas mediante algoritmos que combinan información registrada por diferentes tipos de sensores.
- Desde finales de 1997 empezó a operar la Misión de Medición de Lluvias Tropicales (TRMM, por sus siglas en inglés), constituida por una red de satélites de las agencias NASA y JAXA, que cubrieron latitudes de 50°N al 50°S, con resoluciones de 0.25° por 0.25° cada tres horas. Los productos obtenidos de su algoritmo (TMPA) se utilizaron extensamente hasta 2014, cuando comenzaron las Mediciones de Precipitación Global (GPM, por sus siglas en inglés), con resoluciones de 0.1° por 0.1° cada 30 minutos, para latitudes de 90°N a 90°S. Con su algoritmo IMERG se han conseguido aproximaciones de calidades distintas, obteniéndose un producto final a 3.5 meses de ocurrido cada evento, y en la versión más reciente del IMERG, V06, se dispone del análisis de lluvias diarias registradas desde el año 2000 (Huffman, 2020).
- Debido a los resultados satisfactorios reportados al emplear los productos de las misiones TRMM y GPM, comparables con los productos CMORPH y PERSIANN, como lo describen respectivamente Gebregiorgis *et al.* (2011) y Yu *et al.* (2021), para los propósitos de este trabajo se consideraron las lluvias estimadas con el algoritmo IMERG. Los archivos tipo ráster para las fechas y ubicación geográfica de interés son productos del sistema de datos en línea GIOVANNI del Centro de Servicios de Información y Datos de Ciencias de la Tierra Goddard de la NASA (GES DISC, por sus siglas en inglés).

Adicionalmente, se realizó la recopilación de imágenes de reflectividad de radar meteorológico, así como la aplicación de un procedimiento para estimar la precipitación correspondiente, tomando en cuenta lo siguiente:

- Una limitante para efectuar la asimilación de lluvias registradas con estimadas es la disponibilidad de imágenes de reflectividad de radares meteorológicos en México, pues de acuerdo con el Servicio Meteorológico Nacional, encargado de operar los instrumentos y resguardar la información de los radares que conforman la red nacional (Tabla 1.7), existe una disponibilidad limitada de esta información, cuyo registro histórico en general no es superior a quince años.
- Mediante una solicitud de datos a la ventanilla única del SMN, para cada fecha de interés pueden recibirse hasta 96 imágenes PPI, con resolución temporal de 15 minutos y resolución espacial de 833 m.
- En disposición de las imágenes PPI, se requiere practicar una metodología para la estimación de lluvia. En este trabajo de tesis se siguieron los pasos del procedimiento recomendado por Vilchis *et al.* (2011), los cuales se describen en el Anexo 3. Con el empleo de la plataforma SIG-Idrisi (2016) pueden realizarse procedimientos de importación de imágenes, georreferenciación, conversión de la información numérica de las imágenes a reflectividad, reducción de efectos de atenuación del haz emitido por el radar por lluvia sobre el radomo, caracterización de campos de lluvia y cálculo de intensidad de eventos convectivos y estratiformes (utilizando los modelos de NEXRAD y de Marshall-Palmer, respectivamente), acumulación de los productos de 96 imágenes consecutivas en un día, además de corrección de errores del producto acumulado debido a bloqueos orográficos.
- Para fines de la asimilación de lluvias diarias en un área de estudio, la combinación de las tres bases de datos se practicó considerando la resolución espacial de las estimaciones satelitales, es decir, la distribución de la variable se calculó sobre los centroides de las celdas de  $0.1^\circ$  por  $0.1^\circ$  de resolución que conforman la malla que define la región de estudio. Por esta razón, a partir de la lluvia de radar con resolución de 833 m, para cada celda de resolución satelital se obtuvo una lluvia representativa calculada como el promedio de las nueve láminas de resolución de 833 m más cercanas a cada centroide (Méndez *et al.*, 2006), considerando que el viento puede tener un efecto en la caída de la lluvia, y la reflectividad capturada por el radar puede corresponder con la lluvia depositada sobre la región próxima a cada centroide, en una superficie de  $6 \text{ km}^2$  aproximadamente.

#### 4.1.2. Calibración del modelo y evaluación de resultados

Con la información de eventos históricos distribuidos sobre cada una de las regiones de interés, se buscó la puesta a punto del modelo de asimilación de datos propuesto con la asignación de valores apropiados a sus parámetros conceptuales. En la calibración se evaluó la respuesta del modelo con el ajuste de cada parámetro, dentro de un rango de valores.

Con respecto a la ecuación (2.16), se practicó la variación del parámetro de convergencia numérico  $\gamma$ , útil en el suavizado de la función de peso modificada  $w'_m$ , considerando  $0 < \gamma < 1$ . Además, en la etapa del modelo que corrige el estado antecedente, para las celdas periféricas de la malla se utilizó una interpolación IDW como operador de observaciones, donde la caída del peso respecto a la distancia se controla con un parámetro de potencia,  $\beta$ , tomando en cuenta que valores más grandes incrementan la influencia de las observaciones más cercanas (Yang *et al.*, 2021). Como el parámetro de convergencia influye directamente en la estimación de cada celda de la malla, y el de potencia influye únicamente sobre las celdas periféricas, la calibración se inició suponiendo un valor fijo  $\beta$  igual a 2, de aplicación común en la literatura (Yang *et al.*, 2021), para observar el comportamiento de los criterios estadísticos descritos, según la variación de  $\gamma$  entre 0.1 y 0.9.

Para la medición matemática del error de los valores estimados por asimilación  $P_e$  respecto de los observados  $P_o$  en un campo de lluvias mediante pluviómetros, se emplearon los criterios estadísticos siguientes: el error medio  $ME$ , para identificar si la estimación representa una sub o sobreestimación; el error medio absoluto  $MAE$ , como indicador menos sensible en la penalización de valores atípicos; la raíz del error cuadrático medio  $RMSE$ , para penalizar principalmente los errores de magnitud mayor; el coeficiente de eficiencia de Nash-Sutcliffe  $NSE$ , sensible a valores extremos, y el coeficiente de correlación  $CC$  como medida de la asociación lineal entre estimaciones y observaciones.

En la evaluación de la calidad de los resultados de cada esquema de asimilación se optó por un procedimiento de validación cruzada, que consistió en remover hasta un 10% de los datos pluviométricos disponibles (preferentemente aquellos localizados cerca de los centroides de las celdas con mayor disponibilidad de información), evitando una modificación drástica de la densidad y distribución original de las observaciones, por lo que al no incluirlos en el proceso de asimilación, fue posible considerarlos como referencias en la evaluación de errores (Rozante *et al.*, 2010).

## 4.2. Aplicación del modelo

### 4.2.1. Análisis de una región con disponibilidad de pluviómetros alta

- **Recopilación de datos**

Con la intención de discutir sobre las bondades de la aplicación del modelo de Asimilación de datos construido para la estimación del estado de diagnóstico de un campo de lluvia histórico, es decir, determinar la distribución bidimensional de la lluvia sobre las celdas de una malla regular que representa una superficie de estudio, considerando para su integración datos de una red de pluviómetros, además de estimaciones obtenidas como productos de radar y satélite, se realizó una identificación de eventos diarios registrados sobre el territorio nacional, los cuales sean de interés debido a la magnitud de su acumulación, o por el comportamiento espacial de los valores.

De acuerdo con las necesidades de la modelación, se requiere que los valores de lluvia recopilados de las tres bases de datos correspondan a eventos simultáneos. En estos términos, la limitante principal para llevar a cabo el procedimiento de combinación propuesto está dada por la disponibilidad de datos de imágenes de reflectividad de los radares meteorológicos del país, la cual está indicada de manera general en la Tabla 1.7.

Dadas sus características fisiográficas y los cambios que han tenido lugar para la modificación de su cuenca original, la superficie de la actual Ciudad de México (CDMX) ha sido monitoreada ampliamente en atención a su fuerte densidad de población, y por lo tanto existen registros pluviométricos y pluviográficos disponibles en cantidad y longitud convenientes, como son las bases de datos que administran respectivamente la CONAGUA y el SACMEX. Por lo anterior, para la aplicación inicial de la metodología propuesta se juzgó apropiado considerar como datos los registros de eventos de lluvia históricos ocurridos sobre dicha entidad.

De las lluvias estimadas con información del radar meteorológico más cercano a la región de interés, durante la elaboración de este documento el SMN disponía de imágenes de reflectividad capturada por el radar Cerro Catedral banda C, ubicado en el Estado de México a una distancia menor que su radio de cobertura de 300 km, respecto a eventos posteriores a diciembre de 2007, entonces para la CDMX sólo se consideraron eventos ocurridos desde diciembre del año 2007.

Adicionalmente, con la descarga electrónica de archivos ráster desde el sitio web del Centro de datos GES DISC de la NASA, también fue posible disponer de valores de lluvia estimados con datos de la misión satelital GPM, correspondientes a láminas acumuladas sobre la superficie nacional en celdas de 0.1° de resolución espacial, para fechas posteriores al 2007.

Como una manera de identificar eventos que puedan considerarse representativos de la temporada de lluvias sobre la CDMX, ocurridos en los años más recientes, se consideró de utilidad la revisión del comportamiento de la lluvia media mensual registrada a través de estaciones convencionales y automáticas. Con información de Datos Abiertos de la Comisión Nacional del Agua respecto de lluvias medias mensuales a nivel estatal desde el año 1985, en la Tabla 4.1 se muestran los valores correspondientes a la CDMX, a partir del año 2007, [91].

Tabla 4.1: Lluvia media mensual sobre la Ciudad de México, desde el año 2007, en mm, [91].

Año	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Annual
2007	5.2	21.9	18.9	22.8	58.5	98.8	191.0	196.9	193.7	50.9	6.6	1.3	866.5
2008	0.1	2.9	3.1	39.4	42.6	152.8	<b>164.0</b>	<b>184.7</b>	113.6	37.7	0.0	0.0	740.9
2009	13.2	6.8	8.6	5.9	41.7	86.3	103.6	143.0	227.8	72.1	0.7	3.4	713.1
2010	25.8	57.3	1.6	14.9	27.3	94.3	203.3	174.5	93.7	3.7	1.4	0.1	697.9
2011	0.1	0.7	8.1	39.8	28.0	129.8	230.5	158.9	101.9	62.0	21.9	0.9	782.6
2012	6.9	28.2	26.5	15.1	16.3	96.0	160.5	108.4	83.4	14.9	12.7	0.0	568.9
2013	2.2	1.4	2.3	15.6	52.8	107.1	113.8	130.3	169.7	64.3	36.1	2.5	698.0
2014	0.4	1.4	10.9	21.8	60.0	96.4	99.0	108.2	98.3	92.6	7.1	19.6	615.6
2015	0.3	2.9	19.5	8.3	69.6	74.1	94.1	79.2	114.3	20.2	5.4	4.1	491.8
2016	6.7	0.8	27.8	19.0	51.8	113.0	131.9	128.4	113.7	14.6	45.5	2.7	655.9
2017	0.0	0.0	18.1	12.9	48.9	86.2	101.7	101.3	86.8	31.4	0.6	0.0	488.0
2018	7.9	7.7	12.2	44.5	59.1	127.4	85.9	198.2	120.7	86.8	31.1	4.8	786.2
2019	5.3	5.7	7.2	4.3	21.6	108.1	145.5	87.2	76.1	86.8	17.6	0.1	565.3
2020	11.3	5.3	8.1	21.9	32.8	59.7	103.8	107.8					

De la Tabla 4.1 se aprecia que, en general, durante los meses de julio y agosto se han presentado las lluvias medias de mayor magnitud, con promedios respectivos de 137.8 y 136.2 mm mensuales, valores que resultan de utilidad porque es de esperarse que los eventos de mayor acumulación de lluvia espacialmente distribuida sobre toda la Ciudad (no sólo eventos concentrados espacialmente, de menor interés en un Análisis objetivo de diagnóstico), ocurrieron en fechas en que estos valores medios fueron excedidos, como es el caso del año 2008, para el que se observa que las lluvias medias en julio y agosto fueron de 164.0 y 184.7 mm, respectivamente.

Debido a que los registros de pluviómetros y pluviógrafos son mediciones que permiten conocer la acumulación de lluvia sobre la superficie del terreno en milímetros, para propósitos de la asimilación se consideró indistinto el instrumento empleado para medir la precipitación diaria. Se procedió entonces a la revisión de datos de lluvia registrada mediante los pluviómetros [94] y pluviógrafos [93] operados respectivamente por la CONAGUA y el SACMEX, durante los meses de julio y agosto del 2008, para identificar eventos de mayor acumulación diaria y distribución espacial más amplia.

Con la identificación de cinco eventos de interés del año 2008 (precipitación del 17 de julio, y de los días 2, 5, 7 y 25 de agosto), se procedió a practicar una revisión de los datos para la identificación de valores atípicos que pudieran influir en el análisis del comportamiento de la lluvia. Por ejemplo, la Figura 4-1 representa el contorno de la superficie de la CDMX dentro de una malla uniforme constituida por 30 celdas de 0.1° de resolución espacial, donde también se aprecia la disposición que guardan los puntos de registro de lluvia observada el 17 de julio de 2008.

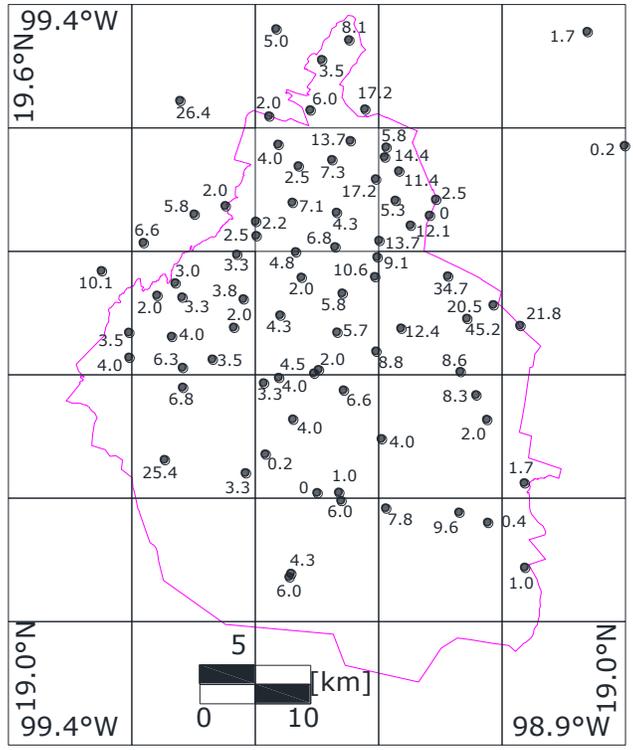


Figura 4-1: Distribución de lluvia registrada sobre la CDMX, julio 17 de 2008, en mm.

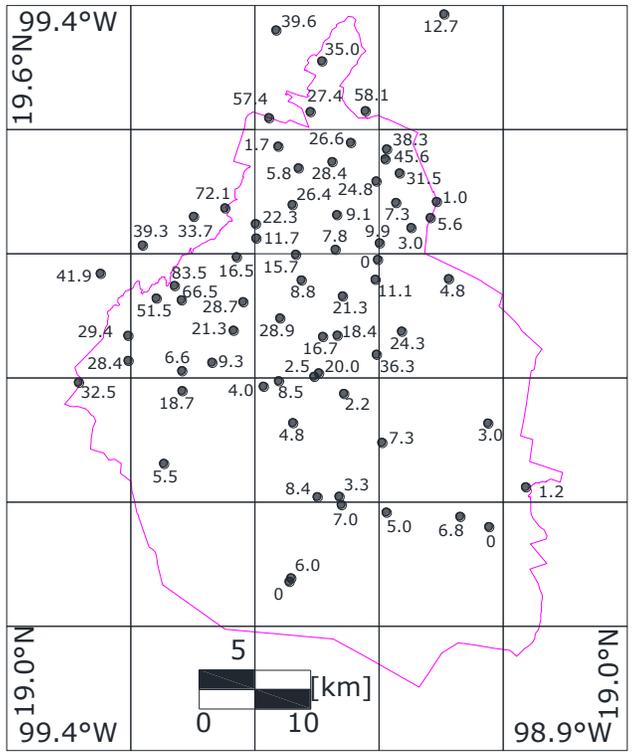


Figura 4-2: Distribución de lluvia registrada sobre la CDMX, agosto 25 de 2008, en mm.

Puede juzgarse que la distribución espacial de la lluvia fue amplia y de magnitud relativamente uniforme (próxima a los 7 mm), con acumulaciones máximas al oriente de la Ciudad (máxima de 45.2 mm). Otros registros superiores a la media se localizan en la zona poniente, pero son valores aislados sobre regiones donde no abundan mediciones para realizar la comparación correspondiente, y por ello se consideraron válidos.

Con respecto al mes de agosto del mismo año, durante el día 25 también se presentaron láminas de lluvia distribuidas sobre toda la CDMX, pero según se aprecia en la Figura 4-2, existe una mayor variación espacial de las acumulaciones (aspecto conveniente en la valoración de las ventajas de aplicar un análisis de asimilación), con registros máximos en la zona norte (58.1 mm) y poniente (83.5 mm) de la Ciudad, magnitudes que en general superan los valores medios de las lluvias máximas anuales (del orden de 47 mm) registradas por la red de pluviómetros desde antes de la mitad del siglo pasado, de acuerdo con Domínguez *et al.* 2017.

En la Tabla A.4.1 del Anexo 4 se concentra la información de lluvia correspondiente a las estaciones de registro pluviométrico consideradas en los cinco eventos históricos de interés para practicar la metodología de asimilación propuesta sobre la superficie de la Ciudad de México.

En relación con las cinco fechas identificadas, se continuó con la recopilación de datos necesarios. Para cada día de lluvia se recibió del SMN un conjunto de 96 imágenes de reflectividad, obtenidas mediante el uso del radar meteorológico Cerro Catedral, con resolución temporal de 15 minutos y resolución espacial menor a un kilómetro. En disposición de las imágenes se realizaron los pasos del procedimiento propuesto por Vilchis *et al.* (2011), utilizando la plataforma SIG-Idrisi para la obtención de lluvia distribuida espacialmente.

A manera de ejemplos, las figuras 4-3 y 4-4 son los mapas de lluvia acumulada diaria sobre la superficie de la malla de estudio (respectivamente los eventos de los días 17 de julio y 25 de agosto del 2008), donde la distribución espacial se aprecia en colores, a una resolución de 833 m. Cabe señalar que, para fines del procedimiento de asimilación, las láminas de lluvia correspondientes a la ubicación de los centroides de las 30 celdas de 0.1° de resolución fueron calculadas como el promedio de las láminas de las nueve micro celdas de 833 m más cercanas a cada uno de los centroides de interés [57], considerando que la lluvia precipitada sobre la región más próxima a cada centroide corresponde con la reflectividad capturada por el radar sobre una superficie de alrededor de 6 km<sup>2</sup>.

Adicionalmente, con la descarga electrónica de archivos tipo ráster desde el sitio electrónico del Centro de datos GES DISC de la NASA, se dispuso de valores de lluvia distribuida estimada a partir de trabajos de la misión satelital GPM, correspondientes a las láminas acumuladas sobre la región de interés y durante las cinco fechas identificadas. Respecto a la magnitud de las láminas obtenidas como productos de satélite, son valores derivados del algoritmo IMERG de estimación multi-satelital, y por corresponder a eventos con una anterioridad superior a 3.5 meses respecto de la fecha de realización de este documento, sí fue posible la descarga electrónica de archivos del producto de aproximación definitiva denominado IMERG Final, calibrados con toda la información disponible para tal propósito.

Como ejemplos, las figuras 4-5 y 4-6 son los mapas de lluvia acumulada diaria sobre la superficie de la malla de estudio, correspondientes al 17 de julio y 25 de agosto de 2008, respectivamente, con resolución espacial de 0.1°.

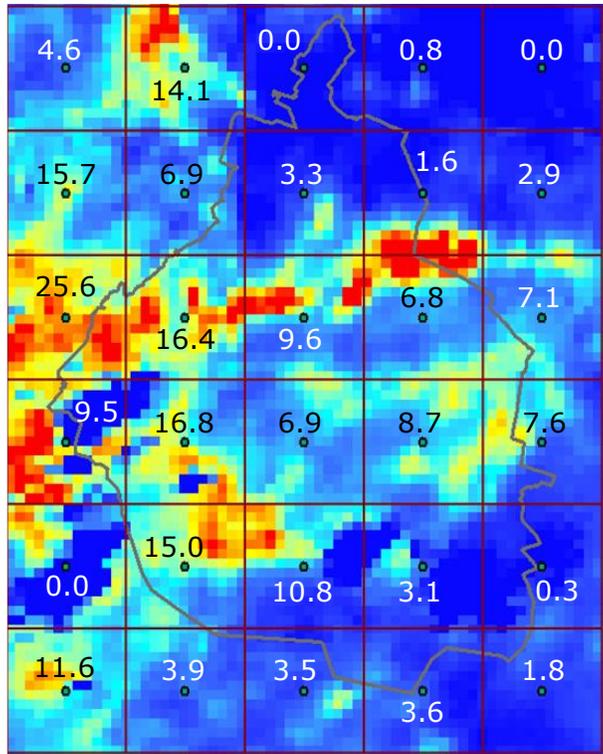


Figura 4-3: Distribución de lluvia estimada con el radar Cerro Catedral, julio 17 de 2008, en mm.

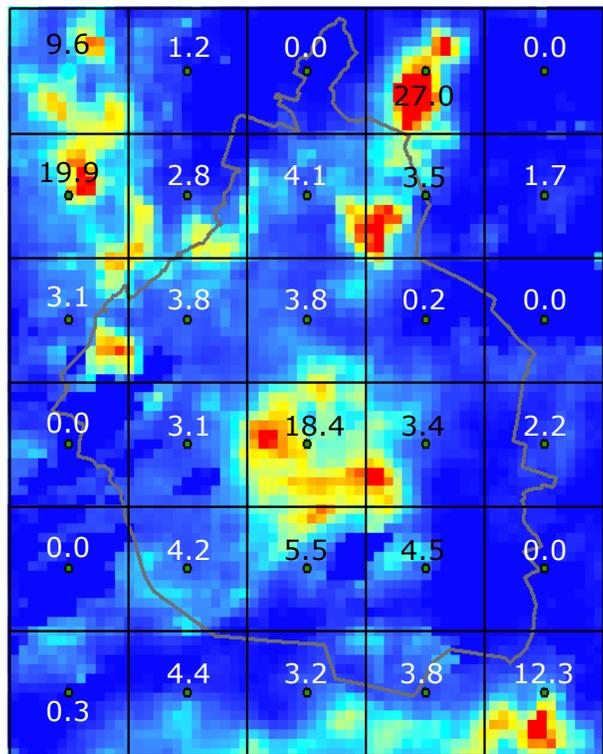


Figura 4-4: Distribución de lluvia estimada con el radar Cerro Catedral, agosto 25 de 2008, en mm.

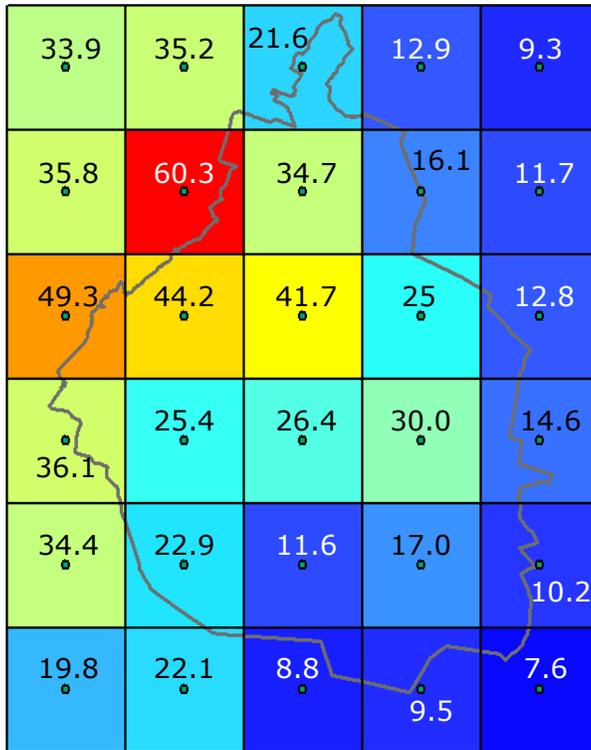


Figura 4-5: Distribución de lluvia estimada con satélite, julio 17 de 2008, en mm.

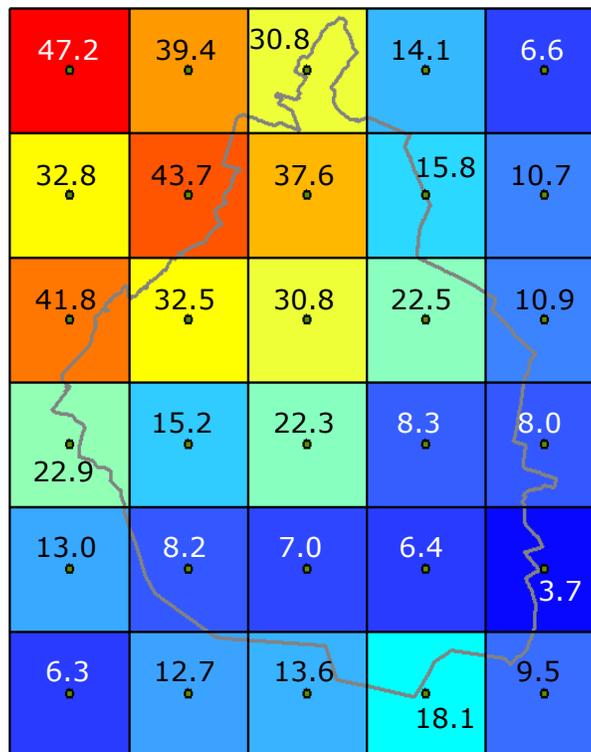


Figura 4-6: Distribución de lluvia estimada con satélite, agosto 25 de 2008, en mm.

Con la construcción de estos esquemas puede observarse cierta semejanza cualitativa entre las distribuciones de acumulación máxima definidas mediante satélite y radar, por lo que bajo esta consideración, el procedimiento de Asimilación para la combinación de lluvias (registros pluviométricos con estimaciones derivadas de radar meteorológico y satélite) se realizó utilizando la resolución de los datos satelitales, es decir, la distribución de la variable se analizó sobre un conjunto de celdas de  $0.1^\circ$  de resolución.

Finalmente, como elementos útiles para la revisión del comportamiento espacial de los datos de lluvia diaria obtenida a partir de productos de percepción remota, respecto a los eventos de interés del año 2008 se construyeron mapas de la distribución espacial histórica de la variable, obtenidos mediante el procedimiento de interpolación simple por Ponderación Inversa de la Distancia (IDW, por sus siglas en inglés), al considerarse que es apropiado aplicar este método porque la cantidad de observaciones puntuales disponibles es grande con distribución espacial amplia (red de hasta 84 registros pluviométricos), y por ello resulta factible una estimación relativamente confiable de la distribución real de las láminas de lluvia sobre la superficie de la región de estudio. Como ejemplos, las figuras 4-7 y 4-8 son los mapas de lluvia por interpolación de registros, correspondientes al 17 de julio y 25 de agosto de 2008, respectivamente.

De los resultados obtenidos, se concluyó que aun cuando existe diferencia en las magnitudes de las lluvias interpoladas respecto de las estimadas por percepción remota, la distribución general de las acumulaciones máximas sí está reflejada en ambas representaciones, aspecto que es de importancia para justificar el empleo de las estimaciones mediante radar y satélite en los procesos de Asimilación de datos para el diagnóstico de sistemas de lluvia.

- **Calibración del modelo**

Para procurar que la aplicación del procedimiento de Asimilación de tres bases de datos propuesto genere campos de precipitación más aproximados a la realidad de los eventos característicos de la Ciudad de México, se buscó la puesta a punto del modelo mediante la asignación de valores a sus parámetros conceptuales.

La calibración del modelo se realizó a partir de un conjunto de entradas conocidas, donde se pretendió la obtención de resultados similares para las correspondientes salidas, también conocidas. Para valorar la influencia de los parámetros del modelo, se evaluó su respuesta a través de la modificación ordenada de cada uno, dentro de un rango de valores determinado.

Respecto a la ecuación (2.16) del modelo, se consideró la posibilidad de asignar valores al parámetro de convergencia numérico  $\gamma$ , útil en el suavizado de la función de peso modificada  $w'_m$ . Con la estimación de lluvias para un mismo escenario,  $\gamma$  se hizo variar hasta obtener un mejor ajuste entre los valores observados y los calculados, considerando  $0 < \gamma < 1$ .

Por su parte, en la etapa del modelo que corrige el estado antecedente debido a las discrepancias entre observaciones del estado verdadero y observaciones “estimadas”, para las celdas periféricas donde el operador de observaciones es una interpolación IDW, se requiere de un parámetro de potencia  $\beta$  que controle la caída del peso respecto a la distancia. Por esta razón, en el proceso de calibración también se consideró la influencia de  $\beta$  en la calidad del estado asimilado.

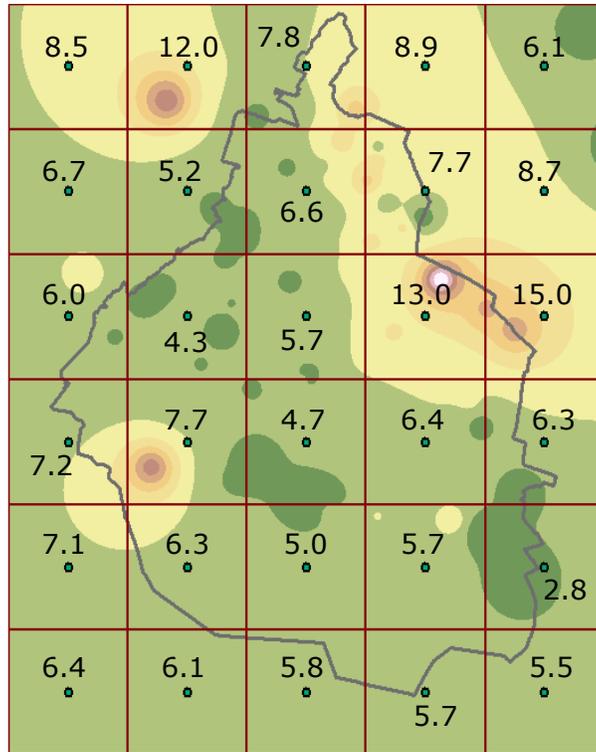


Figura 4-7: Distribución de lluvia estimada con interpolación IDW, julio 17 de 2008, en mm.

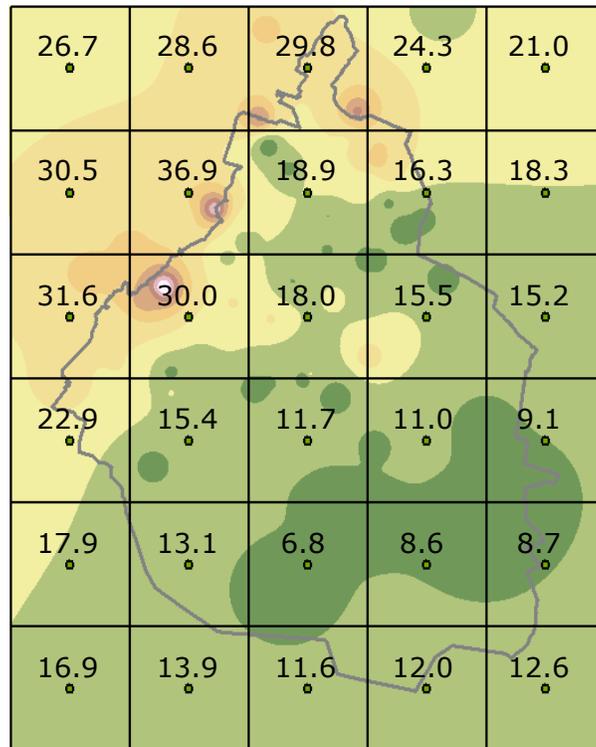


Figura 4-8: Distribución de lluvia estimada con interpolación IDW, agosto 25 de 2008, en mm.

En conocimiento de la distribución espacial de lluvias históricas sobre la CDMX, figuras 4-1, 4-3 y 4-5, correspondientes al evento del 17 de julio de 2008, se procedió a la calibración del modelo de asimilación. Debido a que el parámetro de convergencia  $\gamma$  tiene influencia en el cálculo de la lluvia estimada para cada celda de la malla de estudio, y el parámetro de potencia  $\beta$  influye únicamente en las estimaciones correspondientes a las celdas periféricas, la calibración se inició suponiendo un valor fijo  $\beta = 2$ , de aplicación común en la literatura [87], para observar el comportamiento de los estadísticos según la variación de  $\gamma$  entre 0.1 y 0.9.

En la Tabla A.4.2 del Anexo 4 se muestran los resultados de los criterios estadísticos considerados en la primera etapa de calibración, donde se aprecia que las mejores estimaciones se obtienen cuando  $\gamma$  se encuentra entre 0.3 y 0.5, con menores errores *ME* y *RMSE*.

Con el resultado anterior, se procedió a revisar el comportamiento de los estadísticos para una variación del parámetro  $\beta$  entre 1 y 10 unidades, considerando fijo el valor de  $\gamma$ , e igual a 0.3, 0.4 y 0.5. En la misma Tabla A.4.2 del Anexo 4 se muestran los cálculos de esta segunda etapa de calibración, donde se observa que las estimaciones de menor calidad se obtuvieron para  $\beta$  superior a 3 unidades. En estas condiciones se observó que la mejor estimación, con menores errores *MAE* y *RMSE*, se obtuvo con  $\gamma$  igual a 0.3 para valores de  $\beta$  iguales a 2 ó 3 unidades, pues las diferencias entre los demás estadísticos resultaron menos representativas.

Con los datos de lluvia del 25 de agosto de 2008 también se realizó un procedimiento de calibración similar al descrito, y los resultados se muestran en la Tabla A.4.3 del Anexo 4. Como ocurrió en el primer caso, se observó que las mejores estimaciones se presentan para  $\gamma$  entre 0.3 y 0.5, con menores errores *RMSE* y mejor correlación *CC*. Respecto al parámetro  $\beta$ , entre 1 y 10 unidades, las estimaciones de menor calidad también se obtuvieron para  $\beta$  superior a 3 unidades, siendo la correlación el estadístico más afectado. En este caso las mejores estimaciones se obtuvieron considerando valores de  $\beta$  en el intervalo de 1 a 3 unidades.

Como conclusión a los resultados obtenidos, se juzgó conveniente que en la aplicación del modelo de asimilación propuesto para el análisis de los campos de lluvia sobre la CDMX, los parámetros de convergencia  $\gamma$  y de potencia  $\beta$  sean iguales a 0.3 y 2 unidades, respectivamente.

- **Asimilación de datos de lluvia**

Debido a que la red de registros pluviométricos en la CDMX es relativamente densa para las cinco fechas identificadas, en la evaluación del procedimiento de asimilación se practicaron esquemas distintos:

- Combinación de las tres bases de datos, en consideración de la totalidad de los pluviómetros disponibles (densidad completa, que representa el 100% de los datos).
- Asimilación de tres bases de datos con redes de pluviómetros de densidad reducida (alta al 75%, media al 50% y baja al 25%), procurando conservar la configuración original de la red. De manera complementaria, se examinó la influencia de la distribución espacial de los pluviómetros con el análisis de asimilación de tres configuraciones de registros diferentes para cada densidad de red de pluviómetros.
- Asimilación que considera solo una base de datos de percepción remota a la vez, para las tres densidades y tres configuraciones de pluviómetros definidas.
- Asimilación de los datos de percepción remota de satélite y radar.

En disposición de la lluvia distribuida espacialmente sobre las celdas de la malla que representa la zona de estudio (figuras 4-1, 4-3 y 4-5), para el evento de tormenta del día 17 de julio de 2008 se aplicó el modelo de Asimilación propuesto para la combinación de los datos.

Se procedió a la aplicación de la herramienta programada en lenguaje MATLAB R2017a. Los datos adicionales a las láminas de lluvia que el programa requiere se indican en la Tabla 4.2.

Tabla 4.2: Datos requeridos para la ejecución de la herramienta de asimilación.

Datos ingresados al programa de MATLAB	
Nombre de la cuenca	CDMX
Longitud horizontal de malla en km	55.65
Longitud vertical de malla en km	66.78
Longitud de celda en km	11.13
Número de registros pluviométricos	72
Valor del parámetro beta	2
Valor del parámetro gama	0.3

Con la ejecución del programa se obtuvo el campo integrado de lluvia, correspondiente a la fecha de interés, cuya distribución espacial se presenta en la Figura 4-9.

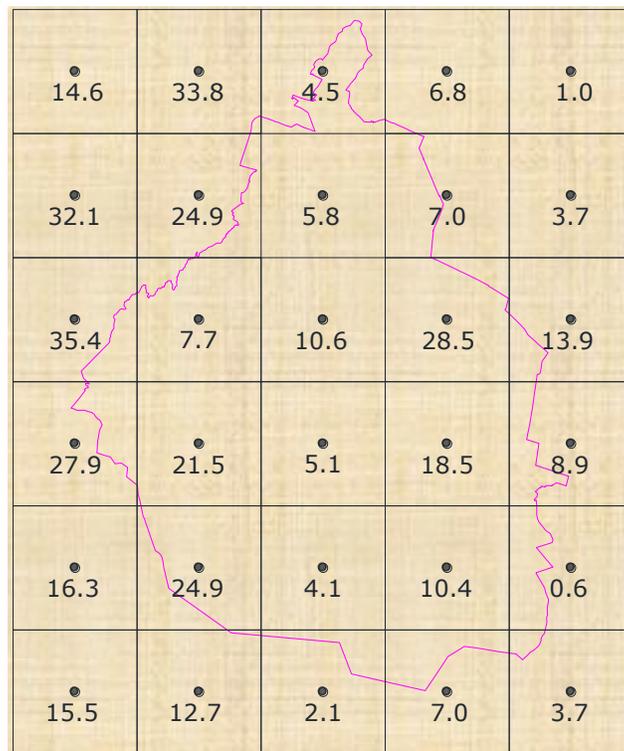


Figura 4-9: Distribución regular de lluvia integrada con asimilación de datos, julio 17, 2008, en mm.

En la Tabla 4.3 se muestran los valores de las siete láminas de lluvia seleccionadas para validación ( $P_o$ ), así como las correspondientes láminas estimadas ( $P_e$ ) por combinación de las tres bases de datos, para las celdas respectivas. Para comparación, también se indican las lluvias estimadas individualmente con satélite y con radar, para las celdas de validación. Además, en la última columna se incluyen las lluvias obtenidas con el procedimiento de interpolación IDW, calculadas con los registros pluviométricos considerados en la asimilación.

Tabla 4.3: Láminas de lluvia para comparación, julio 17 de 2008, en mm.

$P_o$ , Validación	$P_e$ , Combinación	Satélite	Radar	Pluviómetro
3.6	4.5	21.6	0.04	7.8
7.1	5.8	34.7	3.3	6.6
2.5	7.0	16.1	1.7	7.8
3.3	7.7	44.2	16.4	4.4
5.7	10.6	41.7	9.7	5.8
45.2	28.5	25.0	6.9	13.0
4.1	5.1	26.4	6.9	4.8

En la Tabla 4.4 se presentan los cálculos correspondientes a los distintos criterios estadísticos considerados para la evaluación de resultados.

Tabla 4.4: Criterios estadísticos para validación, julio 17 de 2008.

	$ME$ (mm)	$MAE$ (mm)	$RMSE$ (mm)	$NSE$	$CC$
Combinación	-0.3	4.8	7.0	0.8	0.97
Satélite	19.7	25.5	27.1	-2.6	-0.16
Radar	-3.8	9.5	15.5	-0.2	0.04
Pluviómetro	-3.0	6.3	12.4	0.2	0.88

Según se aprecia, la calidad de las lluvias combinadas con la asimilación de las tres bases de datos es superior a los valores estimados con las bases de datos individuales. Con relación al error medio  $ME$  como medición del sesgo, se observa que la lluvia asimilada presenta una ligera subestimación. Respecto al error medio absoluto  $MAE$  como indicador menos sensible en la penalización de errores de mayor magnitud, y la raíz del error cuadrático medio  $RMSE$  como medición de la exactitud que penaliza especialmente los errores de mayor magnitud, la lluvia asimilada presenta un mejor comportamiento. En cuanto al coeficiente  $NSE$ , el resultado obtenido puede interpretarse como un ajuste de buena calidad. Por último, del coeficiente de correlación  $CC$  se observa que sí existe una asociación lineal importante entre la lluvia asimilada y los datos observados de validación.

El procedimiento descrito se practicó para el análisis de la distribución espacial de lluvia diaria para los demás eventos de interés ocurridos sobre la Ciudad de México, utilizando la totalidad de la información disponible de las tres bases de datos. Con el propósito de representar de manera condensada los resultados de la lluvia asimilada respecto de las precipitaciones obtenidas de forma independiente con los datos de satélite, radar y las interpolaciones IDW calculadas con los registros pluviométricos considerados en cada asimilación, en la Figura 4-10 se muestran los criterios  $ME$ ,  $MAE$ ,  $RMSE$ ,  $NSE$  y  $CC$ , correspondientes a los estadísticos promedio del análisis de los cinco días de lluvia analizados.

En base a estos resultados puede concluirse que, en disposición de una red de pluviómetros de densidad y distribución de registro amplias, la combinación de las tres bases de datos mediante la aplicación de la metodología de asimilación propuesta genera estimaciones satisfactorias de un campo de lluvia histórico, por lo que se recomienda su aplicación. También se observa que efectivamente la magnitud de la lluvia asimilada presenta una ligera subestimación, incluso mayor que la calculada con la interpolación IDW, aunque en promedio la penalización de los errores extremos resultó menor, lo que también se ve reflejado en los coeficientes adimensionales que dan muestra de la bondad de los ajustes obtenidos. Asimismo, se aprecia una sobreestimación de las lluvias derivadas de datos satelitales, así como una subestimación de las lluvias calculadas con imágenes de radar, con la particularidad de que existe un orden de magnitud similar en los errores cometidos al utilizar ambas tecnologías de percepción remota en la estimación de lluvia.

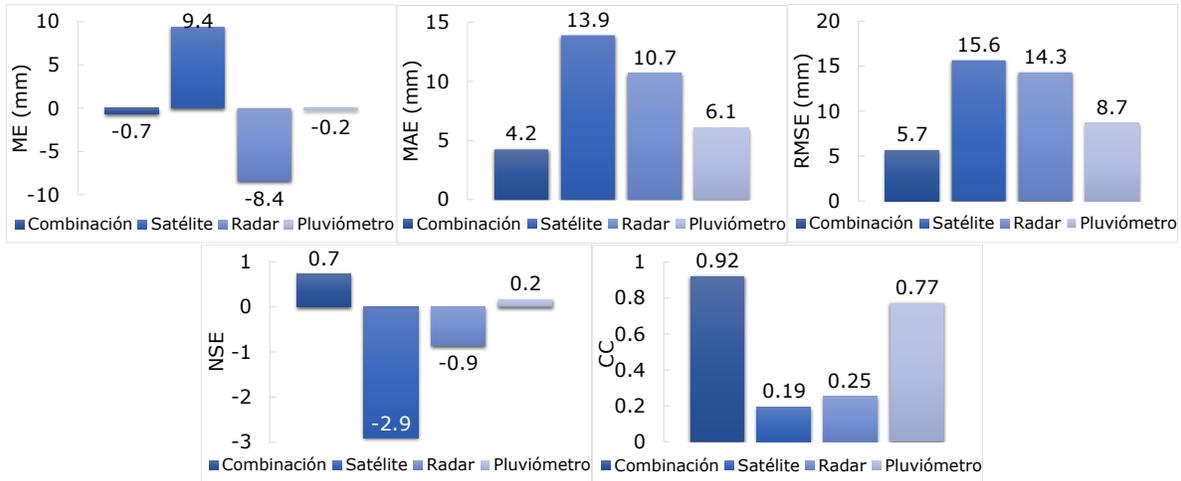


Figura 4-10: Criterios estadísticos promedio en la evaluación de cinco eventos de lluvia, CDMX.

En vista de los resultados satisfactorios que se obtuvieron de los ejemplos de estimación de campos de lluvia históricos sobre la CDMX, considerando en la asimilación toda la información disponible, se consideró importante revisar sobre la calidad de las estimaciones si el análisis tuviera que practicarse para eventos en los cuales la disponibilidad de registros pluviométricos estuviera más limitada. Con propósitos de comparación se analizaron los mismos eventos del año 2008, pero de cada fecha se estudiaron tres esquemas distintos respecto a la cantidad de pluviómetros considerados en la combinación de los datos, utilizando densidades de redes con el 75, 50 y 25% del total de las observaciones existentes. Al mismo tiempo, los eventos analizados del año 2008 se consideraron para evaluar la capacidad del modelo en el análisis de asimilación de los registros de lluvia respecto a tres versiones de configuración distintas en cuanto a la distribución espacial de los pluviómetros involucrados en la combinación de los datos. La Figura 4-11 muestra la variación de los parámetros  $ME$ ,  $MAE$  y  $RMSE$ , normalizados respectivamente con los errores máximos de cada evento, además de los coeficientes  $NSE$  y  $CC$  correspondientes.

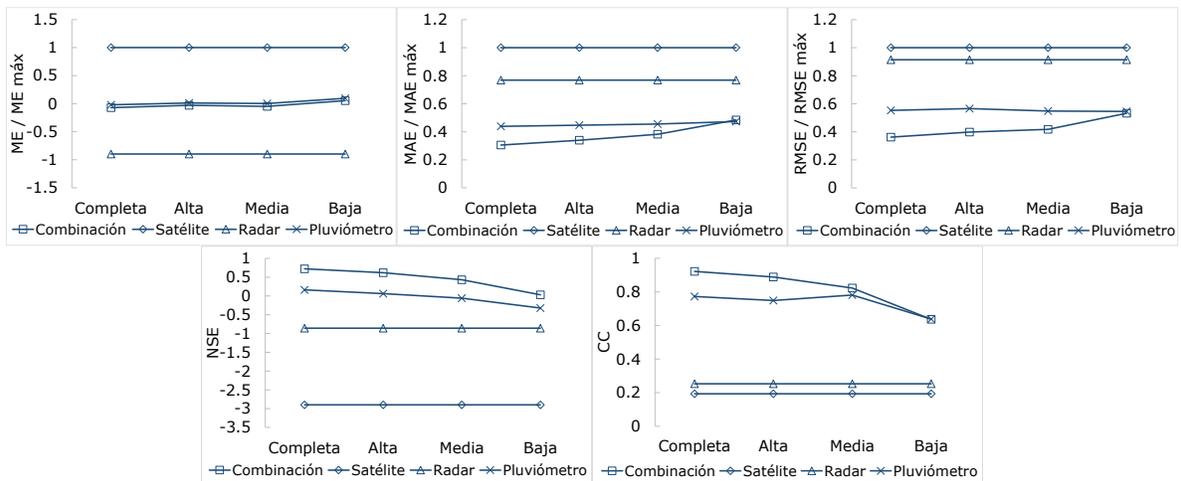


Figura 4-11: Variación de los parámetros normalizados  $ME$ ,  $MAE$  y  $RMSE$ , y los coeficientes  $NSE$  y  $CC$ , en la evaluación de cinco eventos de lluvia.

Estos resultados son una representación resumida de los estadísticos promedio calculados del análisis de las tres diferentes configuraciones de redes de pluviómetros consideradas, tomando en cuenta los cinco días de lluvia de interés.

Del análisis de redes de pluviómetros de diferentes densidades con distribución espacial de observaciones relativamente uniforme, los resultados reflejan consistencia en cuanto a la capacidad del modelo de asimilación para la estimación de los campos de lluvia históricos, pues se verifica que la combinación de las tres bases de datos permite determinar aproximaciones satisfactorias de los eventos registrados, independientemente de la configuración de una red, especialmente cuando la densidad de pluviómetros es alta o media, pero la calidad de las estimaciones se reduce si la densidad es baja, con errores comparables a los obtenidos con la interpolación IDW de los registros pluviométricos, aunque los estadísticos también muestran que aun con una red pluviométrica muy limitada, el resultado de la asimilación de los datos es mejor que las estimaciones individuales del satélite o radar. Cabe observar que, con el análisis de los eventos considerados, en los resultados promedio persiste la condición de subestimación de lluvias con los datos de radar, así como la sobreestimación de lluvias con los datos de satélite.

De manera complementaria, para verificar la conveniencia de practicar el análisis de asimilación propuesto en un caso desfavorable donde solamente se disponga de una base de datos adicional a la de registros pluviométricos, ya sean estimaciones de satélite o de radar, se optó por revisar la calidad de asimilaciones generadas con una sola base de datos de percepción remota a la vez. Se realizaron los ejercicios de combinación correspondientes, y la Figura 4-12 muestra un resumen del comportamiento de los criterios estadísticos respecto a los valores promedio obtenidos con el análisis de las redes de pluviómetros de diferentes densidades y configuraciones consideradas, tomando en cuenta los cinco días de lluvia de interés. Como elementos de comparación, en la misma figura se presentan los estadísticos correspondientes a la asimilación de los datos de satélite con los datos de radar.

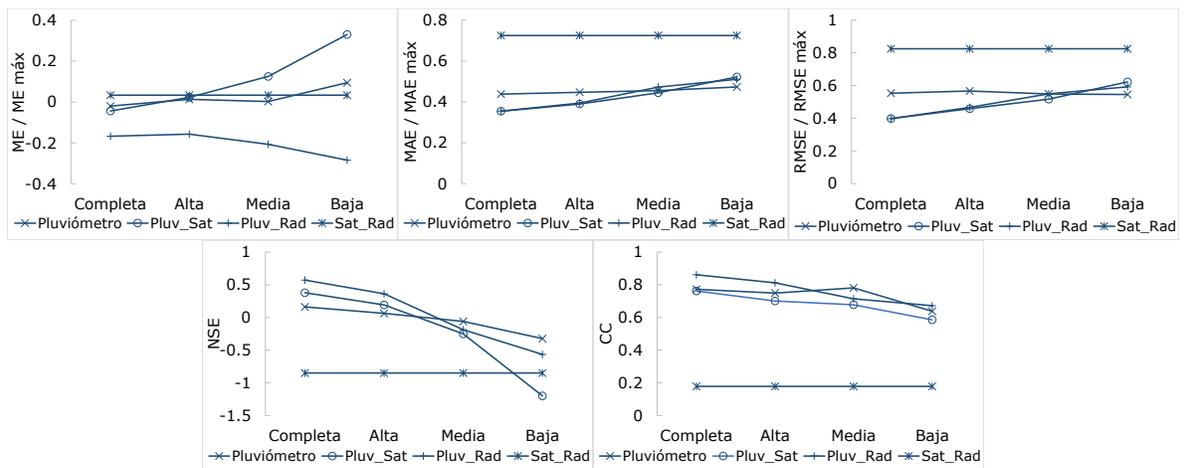


Figura 4-12: Variación de los parámetros normalizados *ME*, *MAE* y *RMSE*, y los coeficientes *NSE* y *CC*, en la evaluación de cinco eventos de lluvia, al combinar sólo dos bases de datos.

Según se aprecia, mediante la asimilación de registros pluviométricos con datos de satélite o datos de radar se obtuvieron mejores aproximaciones para las combinaciones que consideran mayor cantidad de registros, como son las redes de pluviómetros de densidad completa y alta, mientras que para una densidad media los resultados son de calidad semejante a las interpolaciones IDW calculadas con los pluviómetros empleados en la asimilación. Así, estos ejemplos de asimilación

con dos bases de datos permiten juzgar que la aplicación del procedimiento en dichas condiciones es de utilidad, porque ofrece de manera indistinta resultados de calidad similar al considerarse datos de satélite o de radar, observándose respectivamente errores de sobreestimación y subestimación. Además, los resultados de asimilación obtenidos al combinar solamente los datos estimados con las dos técnicas de percepción remota consideradas presentan los errores de mayor magnitud, por lo que los campos de lluvia así definidos son de menor interés.

#### 4.2.2. Análisis de regiones con disponibilidad de información escasa

En secciones anteriores del documento se discutió sobre la distribución espacial de las tormentas y la importancia de las redes de pluviómetros de densidad mínima para realizar el registro correspondiente de eventos históricos. Sin embargo, una revisión de la disponibilidad de estos registros sobre las cuencas que conforman la superficie del territorio nacional permite confirmar que por lo general la cantidad de información disponible resulta insuficiente para conocer la distribución espacial de los eventos de tormenta.

Debido a que mediante la asimilación de las tres bases de datos recomendadas se obtuvieron las estimaciones más satisfactorias de la distribución espacial de lluvia diaria sobre la Ciudad de México, pero la combinación de los registros pluviométricos con alguno de los productos de percepción remota representó una buena aproximación a dicho mejor resultado, se consideró conveniente aplicar este tipo de combinación sobre otras regiones del país donde la distribución espacial no uniforme de registros pluviométricos por unidad de superficie fuera sensiblemente menor. Entonces, para el análisis de una condición más desfavorable se practicó el procedimiento de asimilación utilizando solamente dos bases de datos, considerando registros pluviométricos y estimaciones de productos satelitales, aplicación de interés dado que la disponibilidad de registros de radar meteorológico en México es particularmente limitada.

De una revisión de documentos de carácter de investigación realizados recientemente en la Facultad de Ingeniería de la UNAM, con el propósito de una eventual comparación de resultados, se identificaron tres regiones en atención a su tamaño, disponibilidad de información pluviométrica e hidrométrica, localización geográfica que influye en la magnitud de la lluvia prevalente durante varios días consecutivos, así como su cercanía con centros urbanos de mayor dimensión. Los sitios se localizan dentro de los estados de Chiapas-Tabasco, Coahuila-Nuevo León y Oaxaca-Veracruz.

- **Aplicación del modelo, subcuenca de la E.H. Teapa (Chiapas-Tabasco)**

Como parte de un análisis de escurrimientos superficiales hasta el sitio de la estación hidrométrica E.H. Teapa, estado de Tabasco, en 2019 de Luna estudió sobre el comportamiento espacial de la precipitación ocurrida sobre la subcuenca de captación de dicha estación, particularmente el tren de lluvias diarias desarrollado entre el 19 y el 26 de noviembre de 2015, de manera que pudo comparar el resultado de su modelación hidráulica de escurrimientos con una estimación de caudales obtenida a partir de un registro de niveles de la superficie libre del agua en el sitio de la estación.

En consideración del sitio y evento señalado, para fines de la aplicación del modelo de asimilación de lluvia diaria, se procedió al análisis de las condiciones propias de la región. Al igual que el 95.5% de la superficie del estado de Tabasco, el municipio de Teapa presenta clima cálido húmedo y abundantes lluvias en verano. En el caso de la estación climatológica E.C. Teapa, clave 27044, el

valor medio de la precipitación total anual es de 3 130 mm, con temperatura media mensual de 26.2°C. Asimismo, la lluvia diaria máxima sucedió en septiembre de 1992 (301.2 mm), y el valor medio de láminas diarias máximas mensuales es de 76.5 mm.

Tomando en cuenta la información topográfica digital generada por el Instituto Nacional de Estadística y Geografía (INEGI), en particular el continuo de elevaciones mexicano 3.0 (celdas de 15 m de resolución) de los datos vectoriales de los estados de Tabasco y Chiapas, además de las redes hidrográficas digitales, escala 1:50 000 (edición 2.0, INEGI), en especial la que corresponde a la unidad hidrográfica RH30Di, R. de la Sierra, así como el empleo de un sistema de información geográfica, se identificó la unidad de captación de agua superficial para el sitio de la E.H. Teapa, en la corriente del río Teapa (cuenca del río Grijalva), definiéndose la subcuenca de interés mostrada en la Figura 4-13, cuya superficie tiene una magnitud de 418.6 km<sup>2</sup>, y su cauce principal una longitud de 60.0 km.

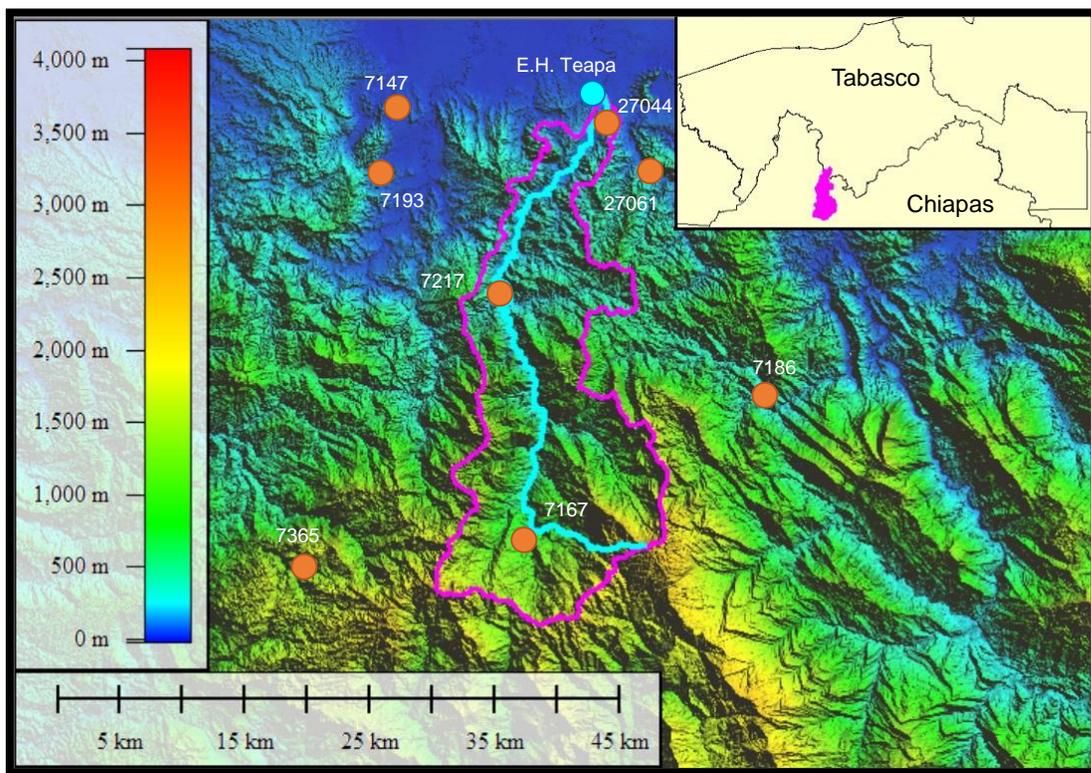


Figura 4-13: Delimitación de la subcuenca con aportación de escurrimiento superficial al sitio de la E.H. Teapa.

Con la intención de caracterizar el estado atmosférico del sitio, se realizó la identificación de las estaciones climatológicas convencionales comprendidas en la región más próxima a la subcuenca de interés. Se consideró la información disponible en la base de datos de las estaciones que conforman la Red Nacional de la CONAGUA [94], y las estaciones con información del evento del 2015 se indican en la Tabla 4.5, mientras su distribución espacial se aprecia en la Figura 4-13.

En la Tabla 4.6 se concentra la información registrada por la red de pluviómetros, que representan la altura de lámina de lluvia en mm, precipitada durante los días que conforman al tren de lluvias

ocurrido del 19 al 26 de noviembre de 2015, de interés porque las láminas distribuidas en los días de mayor acumulación se aproximan al valor medio de los registros máximos de la zona.

Tabla 4.5: Estaciones climatológicas para analizar la subcuenca de la E.H. Teapa.

Clave	Nombre	Latitud	Longitud	Clave	Nombre	Latitud	Longitud
7147	San Joaquín	17°33'29"N	93°6'29"O	7217	Solosuchiapa	17°25'41"N	93°1'44"O
7167	Tapilula	17°14'56"N	93°0'47"O	7365	Ocotepec	17°13'23"N	93°10'26"O
7186	Finca El Escalón	17°21'0"N	92°49'59"O	27044	Teapa	17°32'56"N	92°57'11"O
7193	Pichucalco	17°30'32"N	93°7'12"O	27061	Puyacatengo	17°30'47"N	92°55'12"O

Tabla 4.6: Lluvia en estaciones climatológicas, del 19 al 26 de noviembre de 2015, en mm.

Día	Clave de estación							
	7147	7167	7186	7193	7217	7365	27044	27061
19	19.1	0	26.5	19	13	3.9	3.5	5.8
20	0.8	21.3	0.1	0	0.1	0.4	32.5	24.1
21	4.4	6.7	8.2	0	100	1.5	0	0
22	63.9	11.6	116	0	157	248.8	0	0
23	51.1	76.4	62	68.5	89	48.3	50	63.4
24	0	8.2	0.1	0	3	0	83.8	99
25	9.5	0	0.1	8	0.1	0.3	0.5	0.5
26	0	0	0.1	0	0	0.1	14	6.3

La información hidrométrica de gastos históricos en la estación E.H. Teapa, clave 30032, está disponible en el Banco Nacional de Datos de Aguas Superficiales (BANDAS) de la CONAGUA [89]. Respecto al comportamiento de los escurrimientos relacionados con los eventos de lluvia de interés, se dispone de mayor detalle porque la variación horaria de los gastos fue estimada por de Luna (2019), a partir de registros de niveles de la superficie del agua y su relación con una curva elevaciones-gastos [28].

Como un elemento adicional para visualizar el comportamiento espacial y temporal de la precipitación respecto de la variación de los escurrimientos superficiales registrados en la E.H. Teapa en el periodo señalado, se construyó una representación de los trenes de lluvia diaria correspondientes a las estaciones climatológicas identificadas respecto de la avenida registrada simultáneamente a la salida de la subcuenca, esquema de utilidad para establecer apropiadamente el periodo que conviene analizar en una simulación de escurrimientos derivados de un tren de lluvias, por ejemplo. La Figura 4-14 es un esquema construido con apoyo del SIG Google Earth, donde se aprecian algunos detalles de los eventos de lluvia y escurrimiento.

En disposición de los registros pluviométricos, se procedió a la recopilación de datos de lluvia de origen satelital, obtenidos del sistema de datos en línea GIOVANNI de la NASA. De cada día del tren de lluvias se obtuvo una malla de celdas con láminas diarias, a una resolución de 0.1°x0.1°. En la Figura 4-15 se aprecia la malla uniforme de 20 celdas considerada para asimilación, empleada para representar la superficie de estudio donde se encuentra la subcuenca de la E.H. Teapa, y como ejemplo se muestra la distribución de la lluvia en mm que corresponde al día 23 de noviembre de 2015, cuando en simultáneo ocurrió el pico de la avenida registrada.

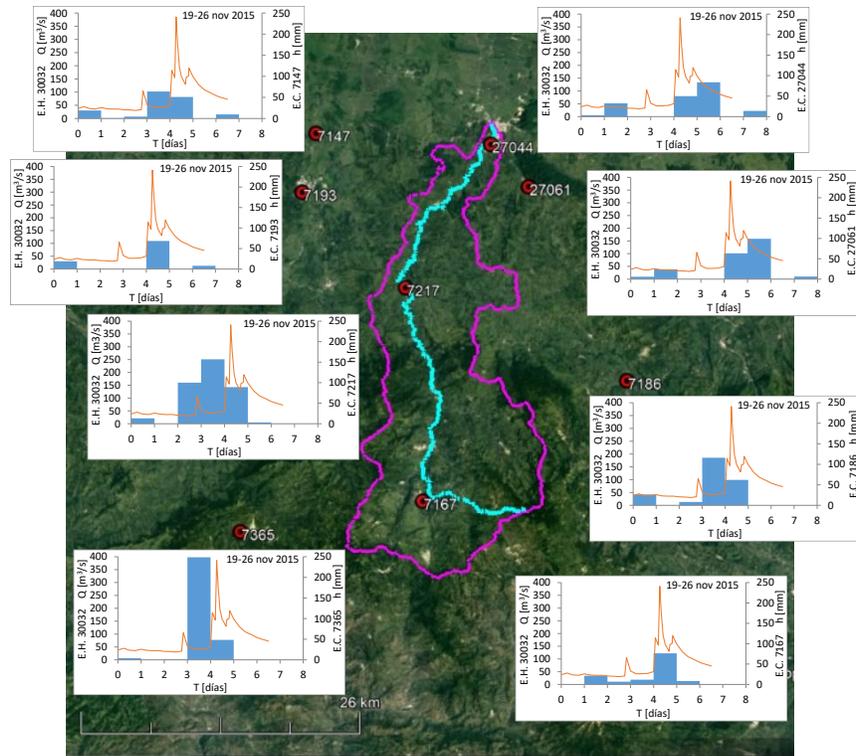


Figura 4-14: Variaciones de lluvia y escurrimiento en noviembre de 2015, subcuena E.H. Teapa.

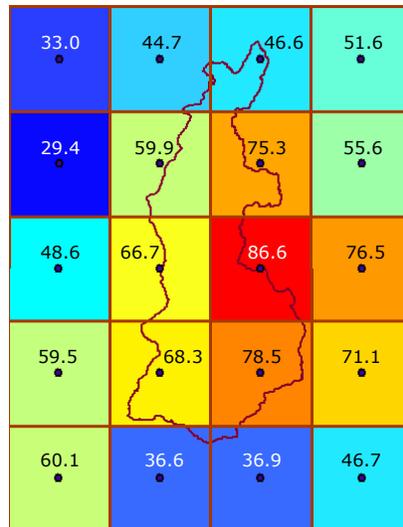


Figura 4-15: Distribución de lluvia estimada con satélite, noviembre 23 de 2015, en mm.

Con información de lluvia distribuida sobre la malla de estudio se realizó el procedimiento de calibración del modelo con la asignación ordenada de valores a los parámetros de convergencia y potencia  $\beta$ . Para la medición matemática del error de los valores estimados por asimilación  $P_e$  respecto de los observados  $P_o$  en los campos de lluvia, también se emplearon los criterios estadísticos  $ME$ ,  $MAE$ ,  $RMSE$ ,  $NSE$  y  $CC$ . Además, la evaluación del esquema se llevó a cabo con un procedimiento de validación cruzada, en el que se consideraron como observaciones de

validación las lluvias en las estaciones 27044, 7167 y 7147, procurando la asimilación de un conjunto de observaciones cuya densidad y distribución de registros guardara semejanza con la red de pluviómetros disponible originalmente.

La calibración se practicó con la asimilación de las dos bases de datos recopiladas respecto al día 23 de noviembre, por considerarse que la lluvia ocurrida en esa fecha tuvo un impacto generalizado sobre el área de interés, de manera que las mejores estimaciones del campo de lluvia se obtuvieron cuando los parámetros  $\gamma$  y  $\beta$  fueron iguales a 0.4 y 2 unidades, respectivamente.

Se llevó a cabo el procedimiento de asimilación de cada uno de los ocho días que componen al tren de lluvias del 19 al 26 de noviembre de 2015, estimándose así la distribución espacial de las láminas diarias sobre las celdas de la malla de estudio. Como ejemplo, la Figura 4-16 presenta la distribución de lluvia del 23 de noviembre, donde se aprecian por separado las láminas registradas por la red de pluviómetros, las lluvias del producto satelital, las obtenidas mediante la interpolación IDW de los registros pluviométricos considerados en el proceso de asimilación, así como las estimadas por asimilación.

Con el propósito de representar de manera condensada los resultados de la asimilación de las dos bases de datos respecto de las lluvias obtenidas al considerar cada base de manera individual, en la Figura 4-17 se muestra el cálculo de los criterios *ME*, *MAE*, *RMSE*, *NSE* y *CC*, que corresponden a los estadísticos promedio del análisis de los ocho días del tren de lluvias de interés. De manera adicional, en la Figura 4-18 se presenta el cálculo de los mismos criterios, pero en relación a los errores obtenidos al considerar la estimación de las láminas acumuladas durante los ocho días del tren de lluvias analizado.

De acuerdo con las figuras 4-17 y 4-18, respecto a los resultados obtenidos por interpolación IDW de datos de un tren de lluvias que incluye láminas de magnitud reducida o nula (Tabla 4.6), tanto los errores promedio como los obtenidos de láminas acumuladas pueden ser menores que los calculados al asimilar los datos pluviométricos con las lluvias de productos satelitales. Como en general esta circunstancia obedece a la falta de exactitud de las estimaciones con satélite, la cual se manifiesta principalmente respecto a los días de menor precipitación, se revisó el cálculo de los mismos criterios estadísticos, pero considerando solo los días con acumulación de lluvia más significativa, es decir, las lluvias ocurridas entre el 21 y el 23 de noviembre.

En estas condiciones, la Figura 4-19 representa los estadísticos promedio del tren de lluvias de tres días de duración, mientras que en la Figura 4-20 se muestran los estadísticos de las lluvias acumuladas durante los mismos tres días.

Puede concluirse que para los eventos en que las láminas de lluvia son más significativas, las estimaciones de los campos obtenidos de precipitación diaria o acumulada por asimilación de las dos bases de datos son una mejor aproximación que los cálculos obtenidos con las bases de datos analizadas individualmente.

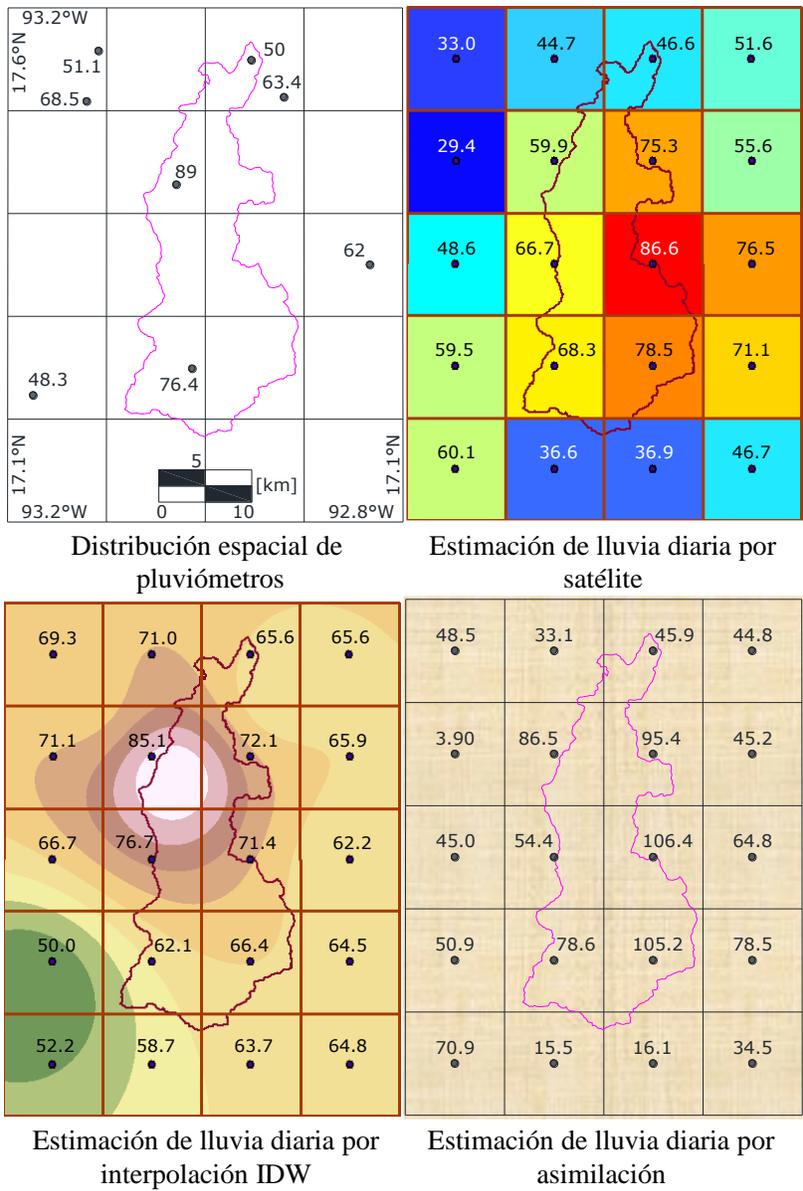


Figura 4-16: Distribución espacial de lluvia, 23 de noviembre de 2015, en mm.

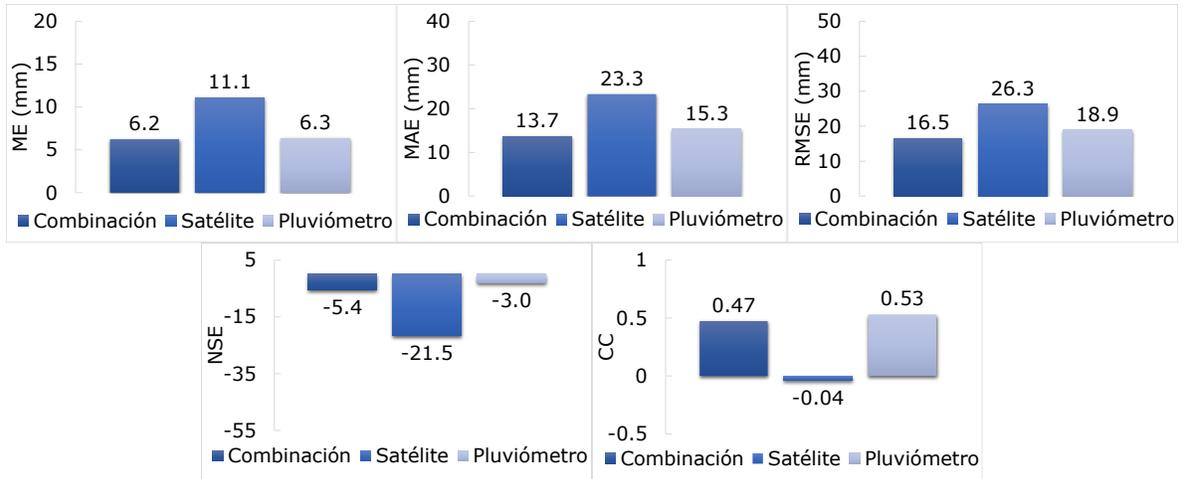


Figura 4-17: Criterios estadísticos promedio de los ocho días del tren de lluvias, noviembre de 2015.

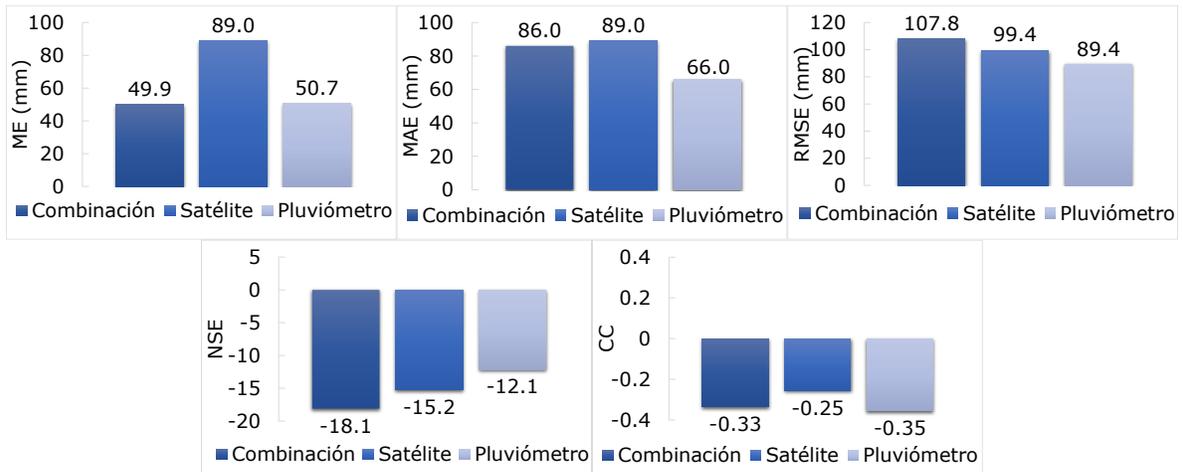


Figura 4-18: Criterios estadísticos del tren de lluvias acumulado, noviembre de 2015.

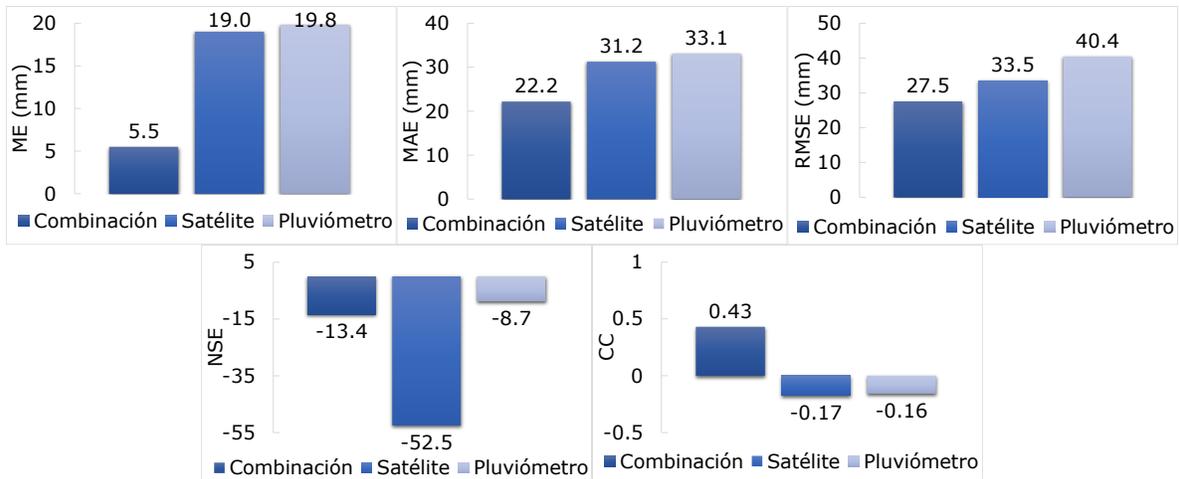


Figura 4-19: Criterios estadísticos promedio de tres días del tren de lluvias, noviembre de 2015.

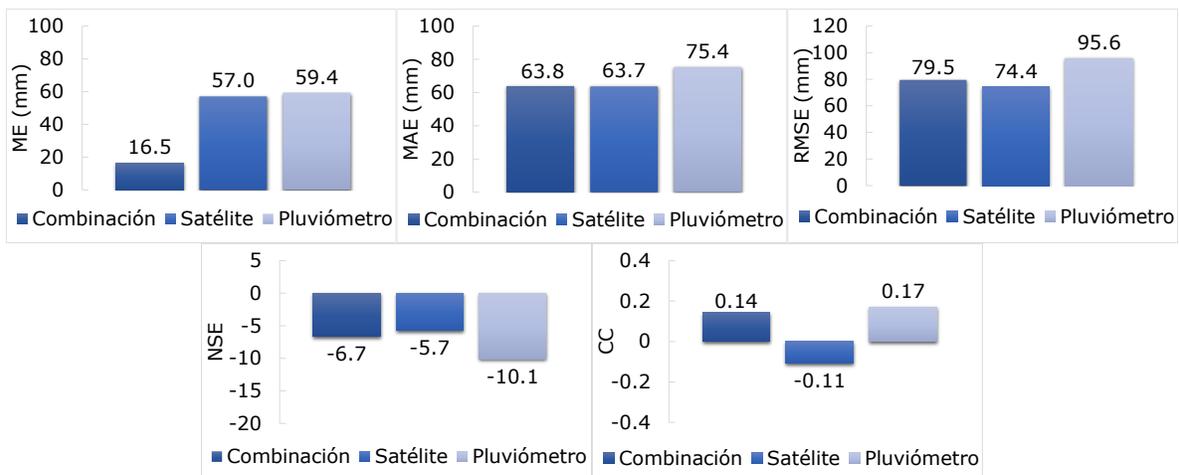


Figura 4-20: Criterios estadísticos de las lluvias acumuladas en tres días, noviembre de 2015.

- **Aplicación del modelo, subcuenca de la E.H. Cadereyta (Coahuila-Nuevo León)**

Con la intención de llevar a cabo un análisis de la variabilidad en la distribución espacial de la precipitación sobre la cuenca del río Santa Catarina, en 2016 Alcaraz realizó una estimación de láminas de lluvia máxima anual asociadas con periodos de retorno de magnitud superior al tamaño de los registros pluviométricos de la región, observando el efecto que sobre las estimaciones tiene la consideración u omisión de datos de origen ciclónico.

En relación con el sitio y eventos de interés, se procedió al análisis de las condiciones propias de la región. Al igual que el 68% de la superficie del estado de Nuevo León, el municipio de Cadereyta Jiménez presenta clima seco y semiseco, con lluvias en los meses de agosto y septiembre. En el caso de la estación climatológica E.C. Cadereyta, clave 19008, el valor medio de precipitación total anual es de 631 mm, con temperatura media mensual de 22.0°C. Asimismo, la lluvia diaria máxima sucedió en septiembre de 2017 (209.4 mm), y el valor medio de láminas diarias máximas mensuales es de 26.2 mm.

A partir de la información topográfica digital generada por el INEGI, en particular el continuo de elevaciones mexicano 3.0 de los datos vectoriales de los estados de Coahuila y Nuevo León, así como las redes hidrográficas digitales, edición 2.0, especialmente la que corresponde a la unidad hidrográfica RH24Bf, R. Monterrey, mediante el empleo de un sistema de información geográfica se identificó la unidad de captación de agua superficial para el sitio de la E.H. Cadereyta, en la corriente del río Santa Catarina (cuena del río San Juan), definiéndose la subcuenca mostrada en la Figura 4-21, cuya superficie tiene una magnitud de 1 812.3 km<sup>2</sup>, y su cauce principal una longitud de 165.6 km.

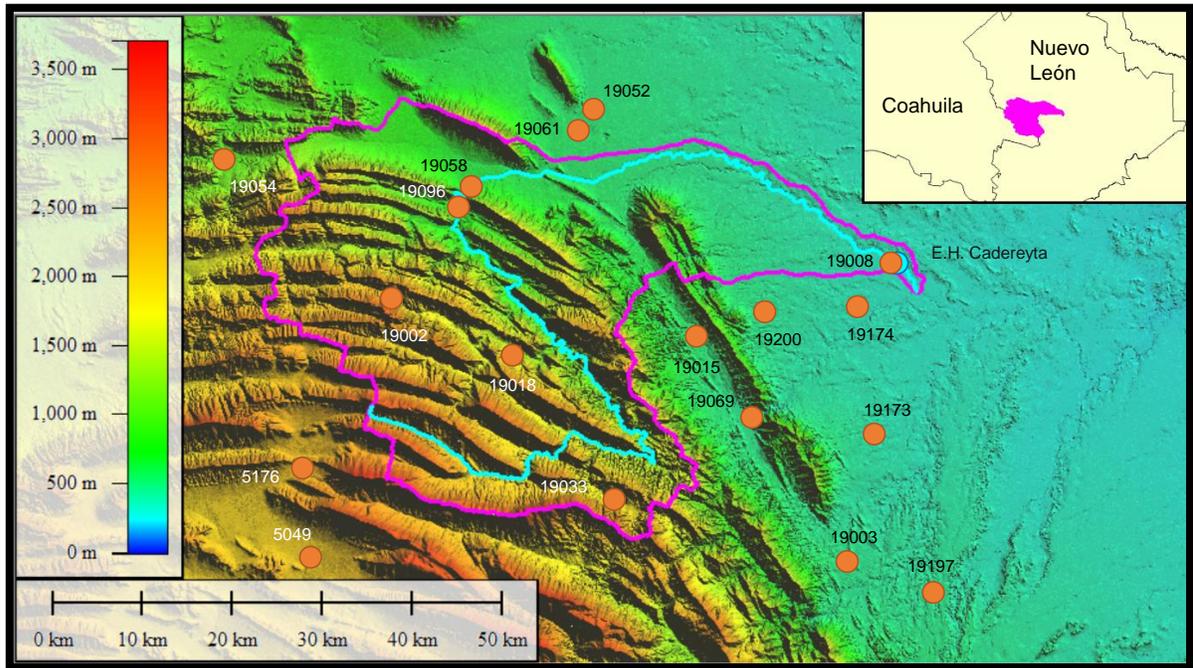


Figura 4-21: Delimitación de la subcuenca con aportación de escurrimiento superficial al sitio de la E.H. Cadereyta.

Para caracterizar el estado atmosférico del sitio se realizó la identificación de las estaciones climatológicas comprendidas en la región más próxima a la subcuenca de la estación Cadereyta, por lo que se consideró la información disponible en la base de datos de la CONAGUA, y su distribución espacial se aprecia en la Figura 4-21. Asimismo, la información hidrométrica de los gastos en la E.H. Cadereyta, clave 24327, se obtuvo de la base de datos BANDAS.

Con objeto de conocer el comportamiento de la precipitación y su efecto sobre la magnitud de los escurrimientos superficiales en la cuena del río Santa Catarina, se tomó en cuenta que sobre esta región del norte del país se han presentado inundaciones relacionadas principalmente con la ocurrencia de eventos ciclónicos, como sucedió con el paso del huracán Gilberto en 1988, el huracán Emily en 2005 y el huracán Alex en 2010. Para disponer de elementos de comparación, se revisaron en simultáneo la variación de los gastos medios diarios de la estación Cadereyta con los registros de lluvia máxima anual de la red de pluviómetros del estado de Nuevo León, con la intención de identificar eventos ciclónicos y no ciclónicos de la región, los que resultan de interés por la magnitud de su precipitación y su consecuente escurrimiento.

En función de lo anterior, en la Tabla 4.7 se indican las estaciones climatológicas identificadas con información de lluvia diaria registrada entre el 8 y el 22 de septiembre de 2002, por tratarse de un tren de lluvias que no muestra relación con evento ciclónico alguno. Por su parte, en la misma Tabla 4.7 se indican las estaciones climatológicas que tienen información correspondiente a la lluvia ocurrida entre el 16 y el 30 de julio de 2005, periodo durante el cual en la subcuenca se observó influencia del paso del huracán Emily.

Tabla 4.7: Estaciones climatológicas para analizar la subcuenca de la E.H. Cadereyta.

Clave	Nombre	Latitud	Longitud	Año	Clave	Nombre	Latitud	Longitud	Año
5049	San Antonio	25°16'59"N	100°37'1"O	2002/05	19054	Rinconada	25°40'52"N	100°43'5"O	2002/05
5176	Jame	25°22'1"N	100°37'16"O	2002	19058	Santa Catarina	25°39'25"N	100°26'46"O	2002
19002	Agua Blanca	25°32'38"N	100°31'23"O	2002/05	19061	Topo Chico	25°43'5"N	100°19'23"O	2002
19003	Allende	25°17'2"N	100°1'12"O	2002/05	19069	La Boca	25°25'44"N	100°7'44"O	2002/05
19008	Cadereyta	25°35'24"N	99°58'30"O	2005	19096	La Huastequita	25°38'20"N	100°27'18"O	2005
19015	El Cerrito	25°30'36"N	100°11'35"O	2002/05	19173	Palmitos (Ge)	25°25'1"N	99°59'49"O	2002/05
19018	El Pajonal	25°29'24"N	100°23'20"O	2002/05	19174	San Bartolo	25°32'31"N	100°0'32"O	2002/05
19033	Laguna de Sánchez	25°20'46"N	100°16'48"O	2002/05	19197	La Morita	25°15'18"N	99°55'52"O	2002
19052	Monterrey (Obs)	25°44'2"N	100°18'18"O	2002/05	19200	La Ciénega	25°32'10"N	100°7'16"O	2002/05

En la Tabla 4.8 se muestran alturas de lámina de lluvia acumuladas en mm, registradas durante el tren de lluvias ocurrido del 8 al 22 de septiembre de 2002, y en la Tabla 4.9 las láminas correspondientes al tren de lluvias del 16 al 30 de julio de 2005.

Para visualizar el comportamiento de la precipitación respecto de la variación de los escurrimientos superficiales en la E.H. Cadereyta, de cada periodo señalado se construyó una representación de los trenes de lluvia correspondientes a las estaciones climatológicas identificadas, respecto de las avenidas registradas simultáneamente a la salida de la subcuenca. En las figuras 4-22 y 4-23 se aprecian algunos detalles de los eventos de 2002 y 2005, respectivamente.

En disposición de registros pluviométricos para el análisis de los dos periodos de interés, se procedió a recopilar lluvias estimadas como productos satelitales del sistema en línea GIOVANNI. De cada día de ambos trenes de lluvia se obtuvo una malla de celdas con láminas diarias de resolución de 0.1°x0.1°. Como ejemplos, en la Figura 4-24 se aprecian las mallas uniformes de 60 celdas consideradas para representar la superficie donde está localizada la subcuenca de la E.H. Cadereyta, en particular los valores correspondientes al 10 de septiembre de 2002 y el 20 julio de 2005, cuando se presentaron las láminas máximas de los trenes de lluvia respectivos.

Considerando los datos de lluvia de las fechas significativas mencionadas, se realizó la calibración del modelo con la asignación de valores a los parámetros de convergencia  $\gamma$  y potencia  $\beta$ . También se emplearon los criterios estadísticos *ME*, *MAE*, *RMSE*, *NSE* y *CC* para la medición matemática del error de los valores estimados por asimilación  $P_e$  respecto de los observados  $P_o$ . Finalmente, la evaluación de los esquemas se realizó mediante validación cruzada, de modo que para el evento del año 2002 las observaciones de validación consideradas fueron las lluvias de las estaciones 19061 y 19200, y para el año 2005 se consideraron las estaciones 5049 y 19003, procurando la asimilación de grupos de observaciones con densidad y distribución de registros semejantes a las redes disponibles originalmente.

Tabla 4.8: Lluvia en estaciones climatológicas, del 8 al 22 de septiembre de 2002, en mm.

Día	Clave de estación							
	5049	5176	19002	19003	19015	19018	19033	19052
8	0	0	0	0	0	0	3.3	24.2
9	0	1	3	60	10	7	3.1	23
10	24	40	135	120	114	102	101	86
11	4	0	90	0	51	45	98.6	58.2
12	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0.5	0
14	16	0	12	38.2	34	4	9.7	1.8
15	23	6	42	115	22	43	48	48.9
16	8	0	8	0	6.5	6	145	41.7
17	4	1	11	7.3	0	2	0.5	2.5
18	7	0	12	0	0	3	3.5	1
19	0	0	20	15.6	32	13	22.7	0
20	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0

Día	Clave de estación							
	19054	19058	19061	19069	19173	19174	19197	19200
8	0	0	0	11.9	3.5	65.3	43	0.3
9	67	18.9	10	53.8	14.5	15.3	38	0.8
10	61	98.6	74.5	103	76.3	86.4	76	125
11	12	28.5	14.7	88.7	33.4	10	15	33
12	0	0	0	0	0	0	25	0
13	0	0	0	0	0	0	33	0
14	0	0	20.5	21	21.2	0	31	31
15	34	45	42.6	233.9	89.5	67.5	0	178
16	1	2	0.3	12.2	18.5	0	0	0.9
17	3	0	0	0	0	0	0	0
18	15	0	0	0.3	5	0	0	0
19	0	0	0	6.1	5.5	0	0	17
20	0	0	0	0	0	0	0	0
21	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0

Tabla 4.9: Lluvia en estaciones climatológicas, del 16 al 30 de julio de 2005, en mm.

Día	Clave de estación						
	5049	19002	19003	19008	19015	19018	19033
16	15	5	22.6	0.9	7	0	2.1
17	0	20	0	0	0	0	0
18	0	0	0	0	0	0	0
19	0	8	0	2	5.3	14.5	25
20	70	250	210.6	164.5	330	293	286.5
21	0	5	23.8	0	18.5	13.6	23.5
22	0	0	0	0	12.5	0	1
23	0	0	0	0	0	0	0
24	14	3	12.8	0	8	7	24.5
25	0	29	33.9	45.6	62	28	33.5
26	0	5	8.7	2	0.5	0	2
27	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0

Día	Clave de estación						
	19052	19054	19069	19096	19173	19174	19200
16	25.3	0	31.2	0	2	6	11
17	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0
19	9.9	5	3	4	4	0	0
20	190.9	173	328	236	217.2	218	234
21	47	0	12	19	0	20	0.2
22	0	0	6	0	0	0	0
23	0	0	0	0	0	12	0
24	0.3	0	13	3	2.5	22	0.2
25	40	7	54.5	17	52	0	44
26	1.2	0	0	0	0	0	0
27	0.8	4	0	0	0	0	0
28	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0

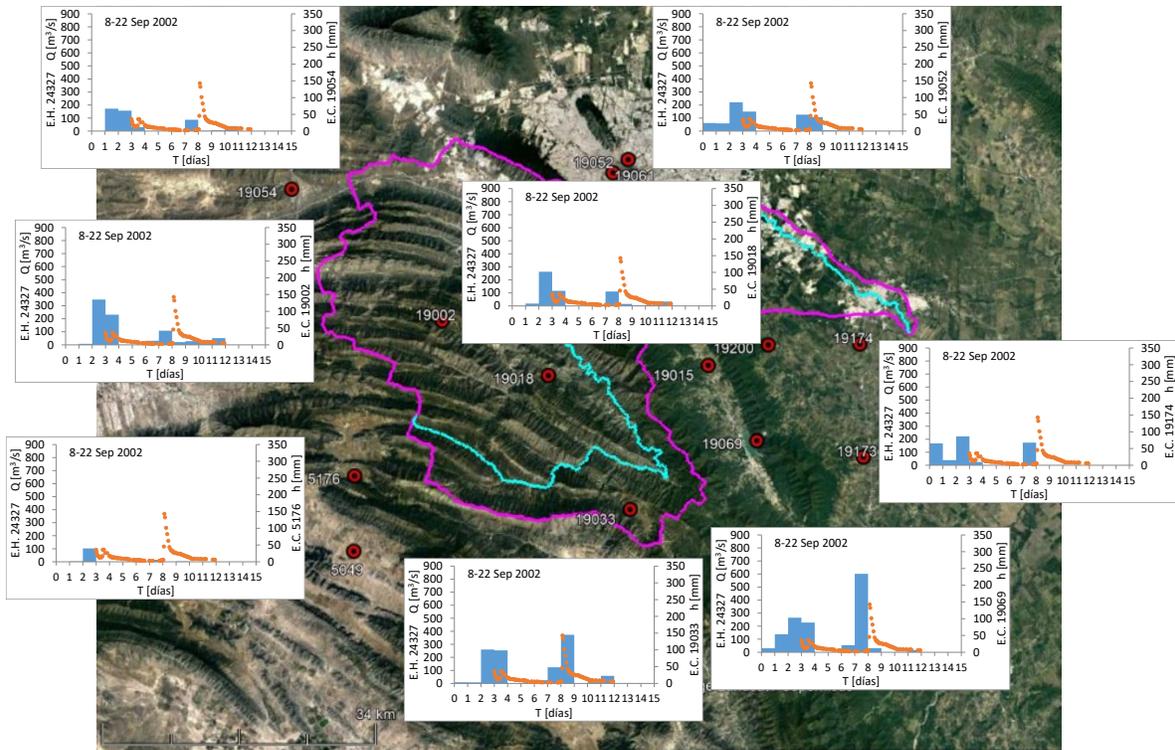


Figura 4-22: Variaciones de lluvia y escurrimiento en septiembre de 2002, subcuena E.H Cadereyta.

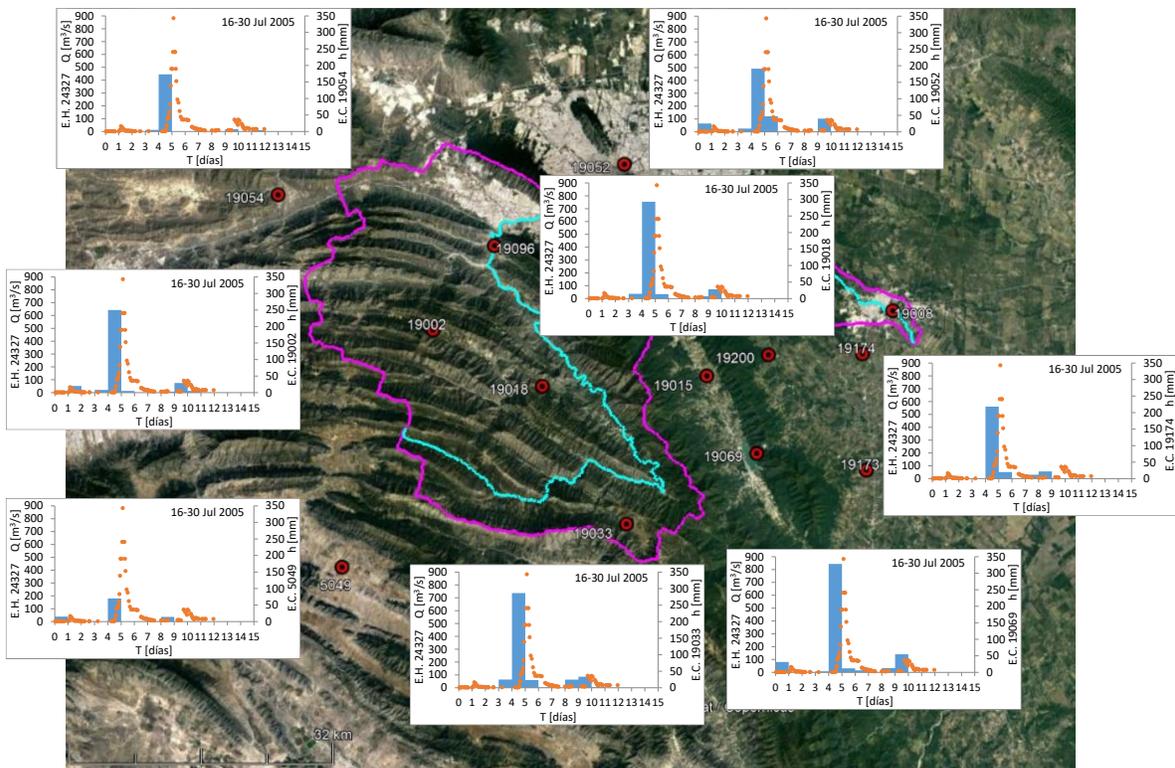


Figura 4-23: Variaciones de lluvia y escurrimiento en julio de 2005, subcuena E.H Cadereyta.

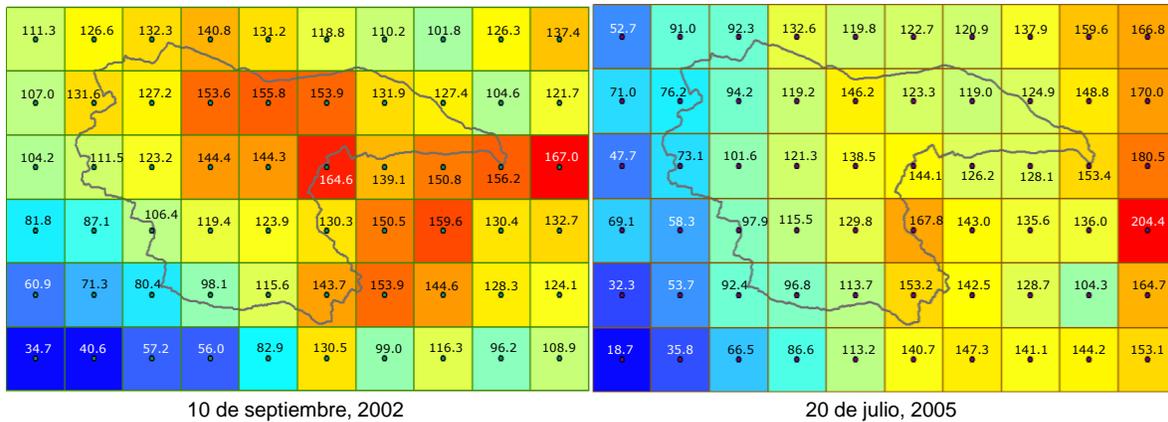


Figura 4-24: Distribución de lluvia estimada con satélite, eventos de 2002 y 2005, en mm.

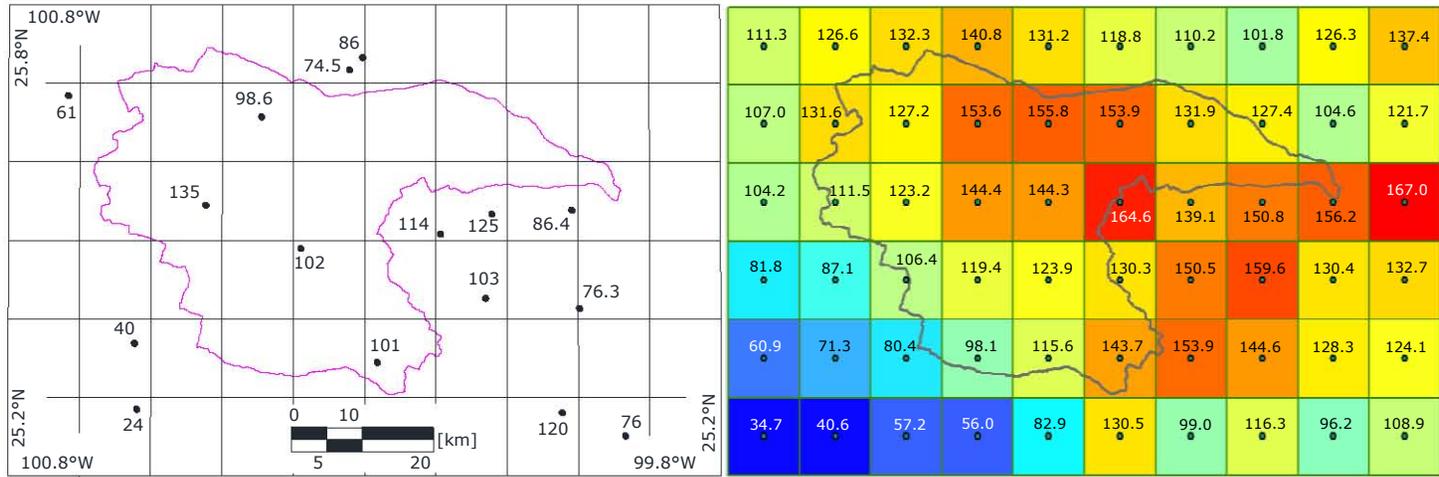
La calibración se practicó considerando los datos de las fechas señaladas, procurando definir una combinación de valores apropiados de  $\gamma$  y  $\beta$  para los dos eventos analizados, y se observó una semejanza en la magnitud de los errores calculados para una misma combinación, independientemente de la fecha considerada. En estas condiciones, se concluyó que la mejor estimación de un campo de lluvia sobre la subcuenca de estudio se obtiene cuando los parámetros  $\gamma$  y  $\beta$  se igualan a 0.4 y 1 unidades, respectivamente.

Se efectuó la asimilación de cada uno de los quince días que componen a los dos trenes de lluvia analizados, estimándose las láminas diarias sobre la malla de interés. Como ejemplos, las figuras 4-25 y 4-26 presentan las distribuciones correspondientes a los días 10 de septiembre de 2002 y 20 de julio de 2005, donde se muestran por separado las lluvias de la red de pluviómetros, las de origen satelital, las interpolaciones IDW de los registros empleados en la asimilación, así como las estimadas por asimilación.

Para representar de manera condensada los resultados de asimilar las dos bases de datos, así como las lluvias obtenidas con cada base individualmente, en las figuras 4-27 y 4-29 se muestra el cálculo de los criterios *ME*, *MAE*, *RMSE*, *NSE* y *CC*, como estadísticos promedio del análisis de los quince días de los trenes de lluvia analizados de los años 2002 y 2005, respectivamente. Adicionalmente, en las figuras 4-28 y 4-30 se presenta el cálculo de los mismos criterios, pero respecto a los errores cometidos al considerar la estimación de las láminas acumuladas durante los quince días de los trenes analizados.

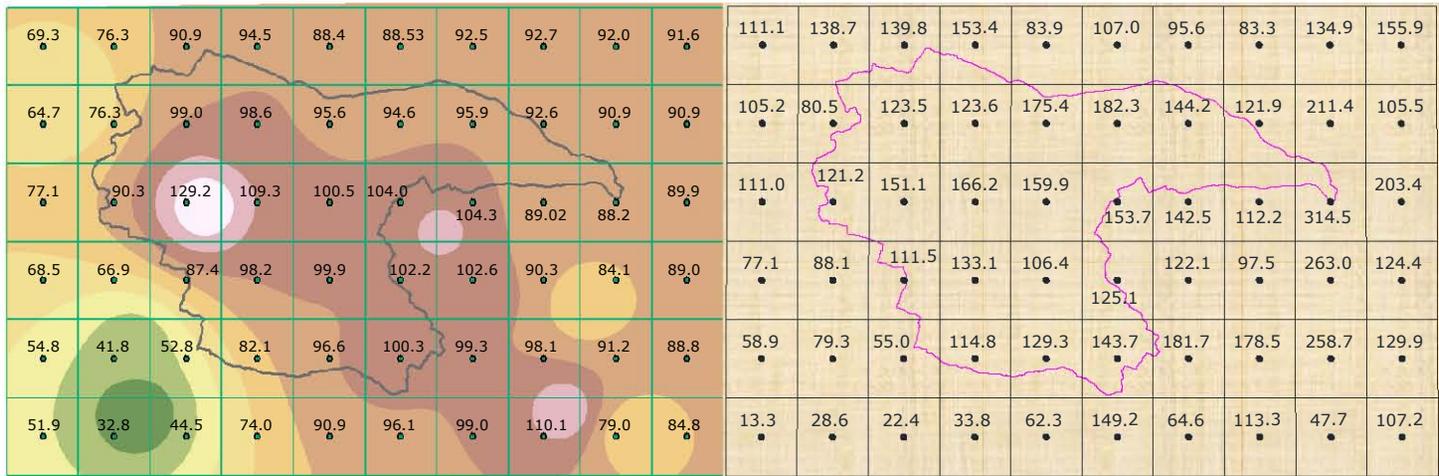
De acuerdo con las figuras 4-27 a 4-30, tanto los errores promedio como los obtenidos de láminas acumuladas de las estimaciones diarias con interpolación IDW de datos pluviométricos de los trenes de lluvia, pueden ser menores que los calculados con la asimilación de los datos pluviométricos y de origen satelital, como consecuencia principalmente de la falta de exactitud en las estimaciones satelitales de los días de menor precipitación.

Para verificar lo anterior, se practicó el cálculo de los mismos criterios estadísticos, pero considerando únicamente los días con acumulación de lluvia más significativa, es decir, los días 10, 11, 14 y 15 de septiembre de 2002, así como los días 16, 20 y 25 de julio de 2005.



Distribución espacial de pluviómetros

Estimación de lluvia diaria por satélite



Estimación de lluvia diaria por interpolación IDW

Estimación de lluvia diaria por asimilación

Figura 4-25: Distribución espacial de lluvia, 10 de septiembre de 2002, en mm.



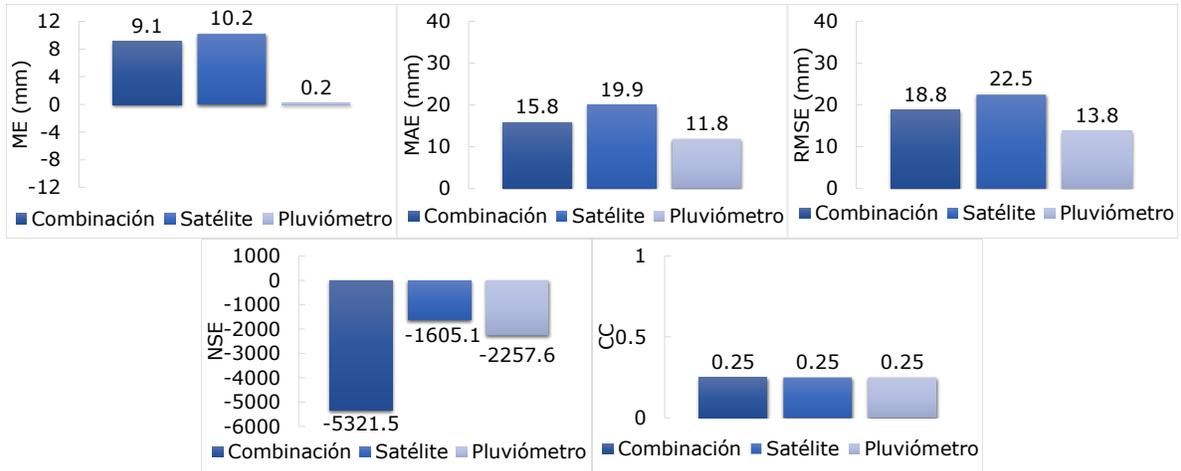


Figura 4-27: Criterios estadísticos promedio de los quince días del tren de lluvias, septiembre de 2002.

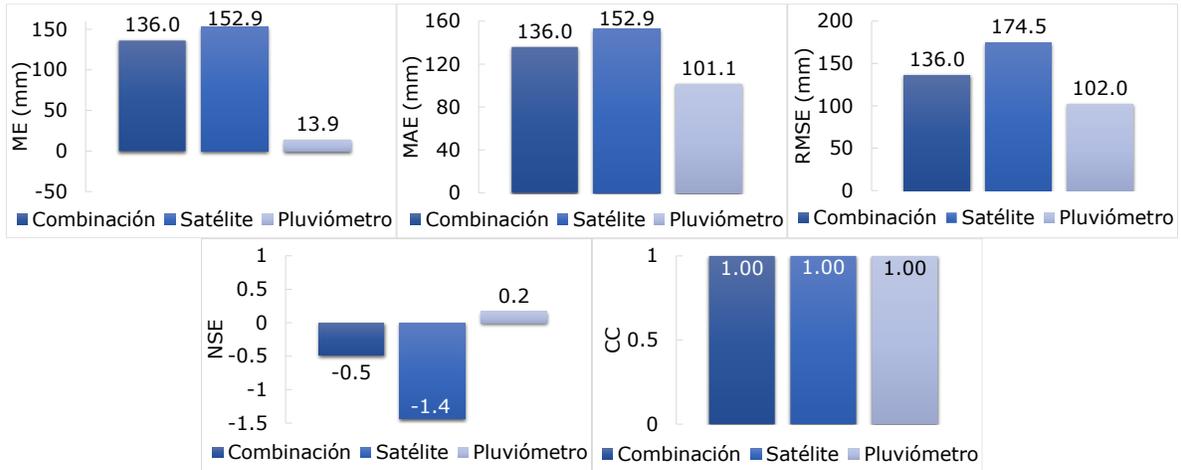


Figura 4-28: Criterios estadísticos del tren de lluvias acumulado, septiembre de 2002.

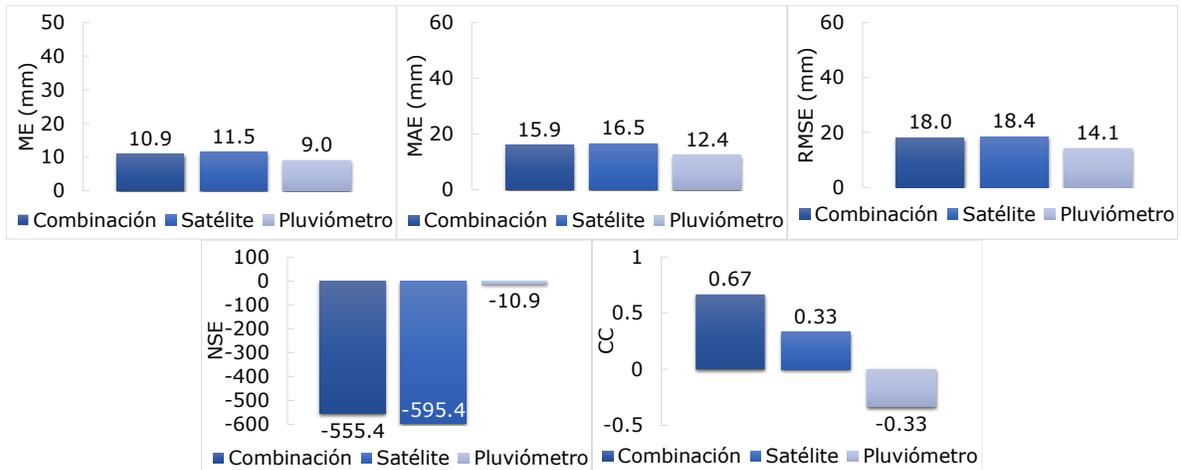


Figura 4-29: Criterios estadísticos promedio de los quince días del tren de lluvias, julio de 2005.

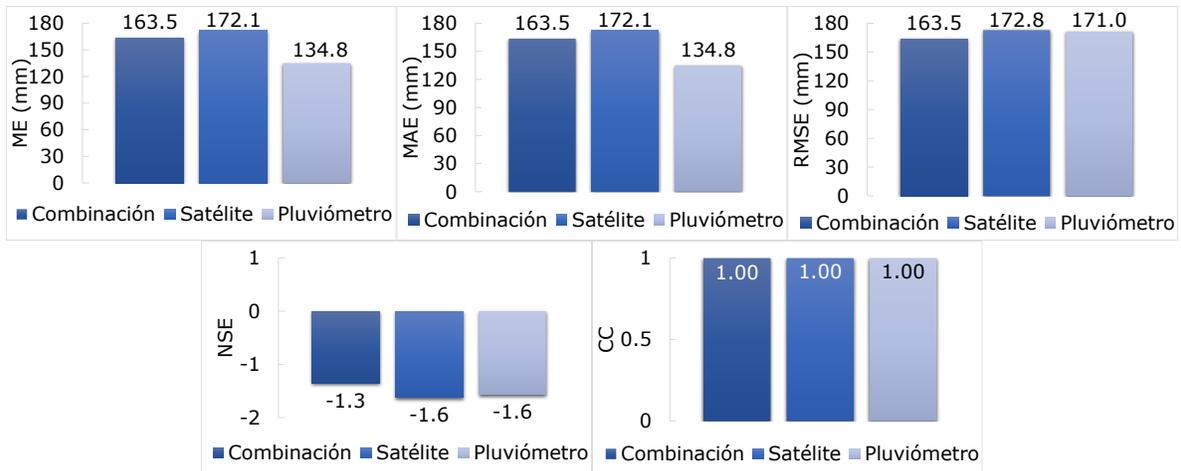


Figura 4-30: Criterios estadísticos del tren de lluvias acumulado, julio de 2005.

Las figuras 4-31 y 4-33 representan los estadísticos promedio de los días de lluvia más significativa de los eventos de 2002 y 2005, respectivamente, mientras que las figuras 4-32 y 4-34 muestran los estadísticos de las precipitaciones acumuladas en los mismos días, donde se observa que en general las estimaciones por asimilación de las dos bases de datos son una mejor aproximación que las obtenidas con las bases individuales.

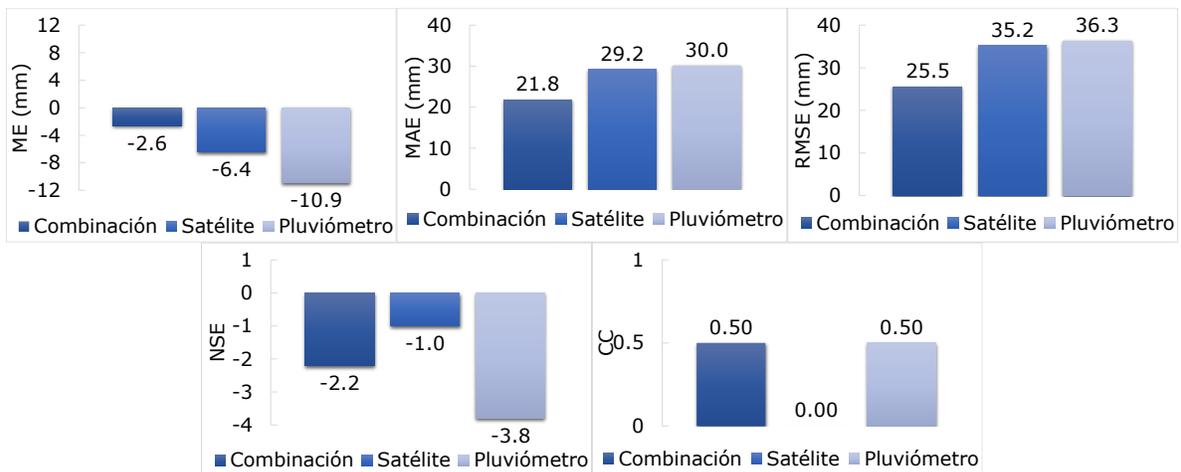


Figura 4-31: Criterios estadísticos promedio de cuatro días del tren de lluvias, septiembre de 2002.

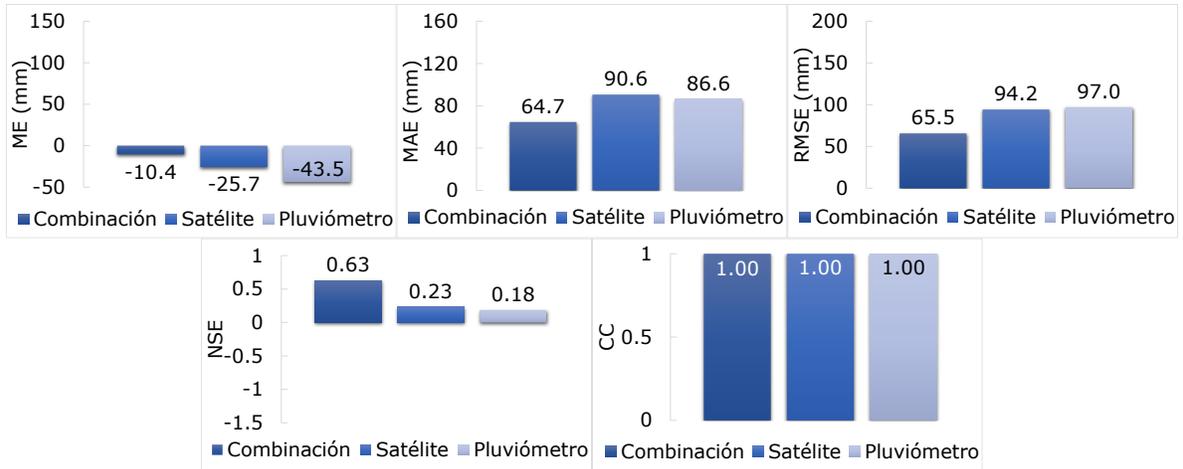


Figura 4-32: Criterios estadísticos de lluvias acumuladas de cuatro días, septiembre de 2002.

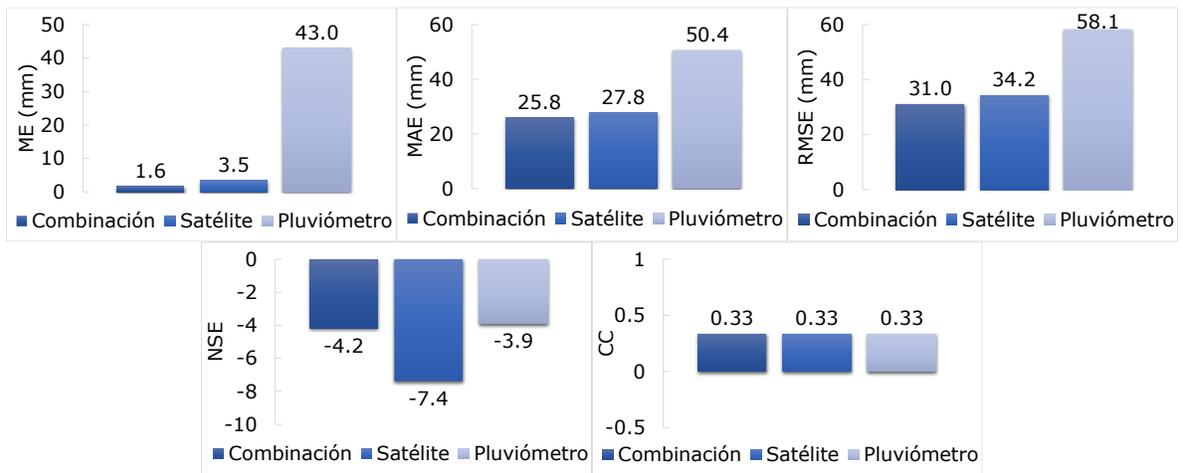


Figura 4-33: Criterios estadísticos promedio de tres días del tren de lluvias, julio de 2005.

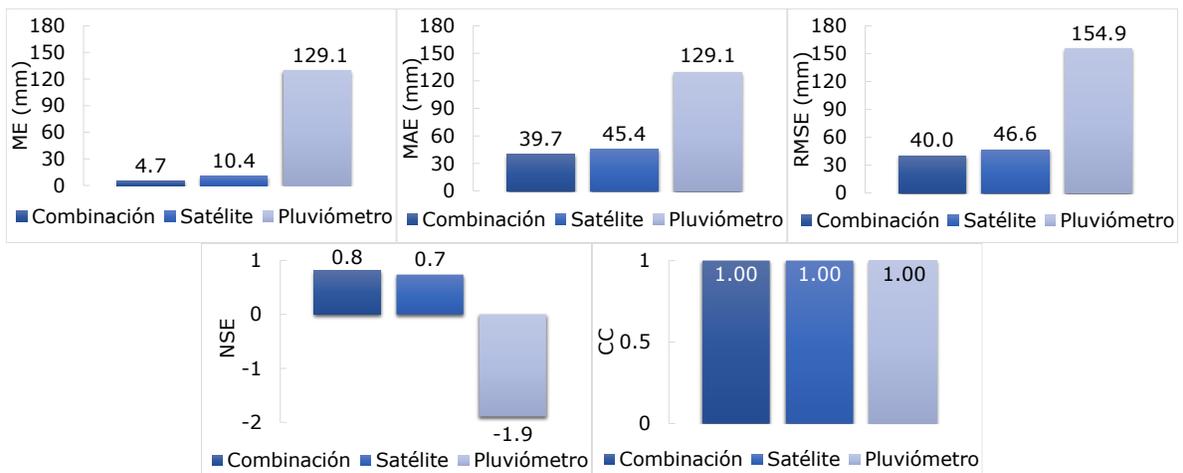


Figura 4-34: Criterios estadísticos de lluvias acumuladas de tres días, julio de 2005.

- **Aplicación del modelo, subcuenca de la E.H. Las Perlas (Oaxaca-Veracruz)**

Con el objetivo principal de aplicar la metodología de asimilación de lluvia sobre una región de dimensiones considerables, especialmente por la disponibilidad limitada de registros pluviométricos que pudieran describir satisfactoriamente la distribución espacial de la precipitación asociada con la trayectoria de un evento ciclónico, se identificaron las lluvias ocurridas entre el 15 de septiembre y el 3 de octubre de 2010, relacionadas con el paso de la tormenta tropical Matthew, información que resultó de interés dada la disponibilidad de registros simultáneos de carácter hidrométrico reportados en la E.H. Las Perlas, estado de Veracruz.

En atención del sitio y evento señalado, se procedió al análisis de las condiciones propias de dicha región. Al igual que el 41% de la superficie del estado de Veracruz, el municipio de Jesús Carranza presenta clima cálido húmedo, con abundantes lluvias en verano. En el caso de la estación climatológica E.C. Las Perlas, clave 30090, el valor medio de precipitación total anual es de 2 441 mm, con temperatura media mensual de 25.8°C. Asimismo, la lluvia diaria máxima sucedió en julio de 2001 (213.6 mm), y el valor medio de láminas diarias máximas mensuales es de 52.6 mm.

A partir de la información topográfica digital del continuo de elevaciones mexicano 3.0, generado por el INEGI, en particular los datos vectoriales de los estados de Oaxaca y Veracruz, además de las redes hidrográficas digitales, edición 2.0, especialmente las que corresponden a la cuenca RH29B, R. Coatzacoalcos, y con el empleo de un sistema de información geográfica, se identificó la unidad de captación de agua superficial para el sitio de la E.H. Las Perlas, en la corriente del río Coatzacoalcos, definiéndose la subcuenca mostrada en la Figura 4-35, cuya superficie tiene una magnitud de 8 993.1 km<sup>2</sup>, y su cauce principal una longitud de 257.4 km.

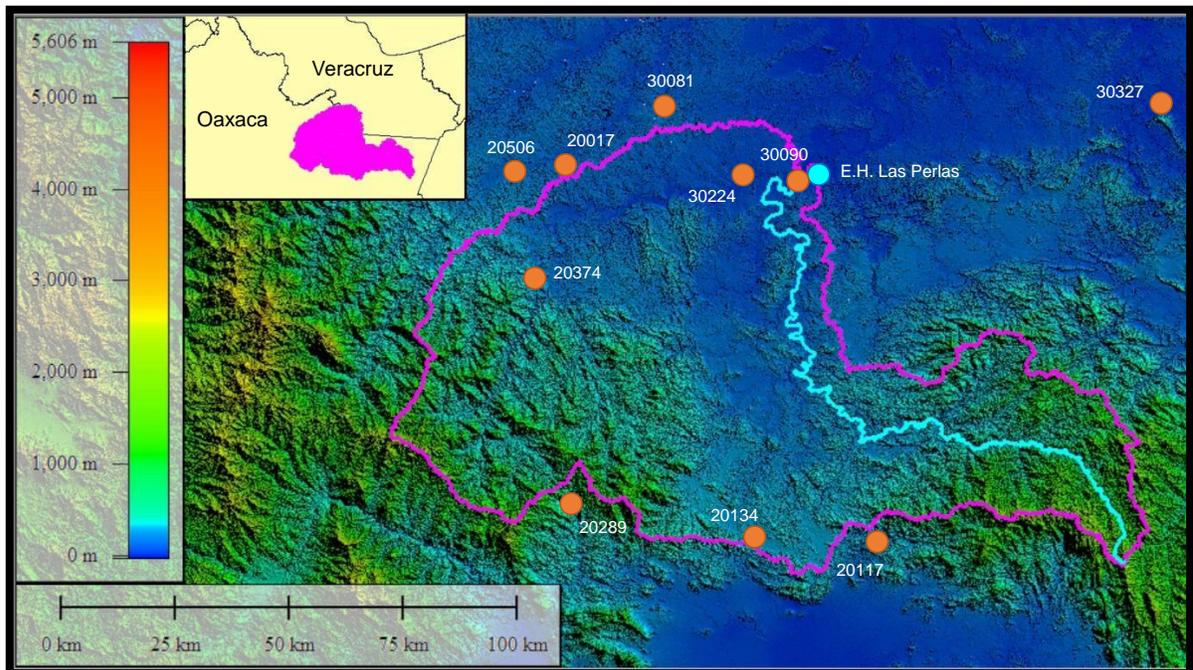


Figura 4-35: Delimitación de la subcuenca con aportación de escurrimiento superficial al sitio de la E.H. Las Perlas.

La caracterización del estado atmosférico del sitio se inició con la identificación de las estaciones climatológicas comprendidas en la región próxima a la subcuenca de interés, considerando la información disponible en la base de datos de la CONAGUA. Por su parte, la información hidrométrica de gastos en la E.H. Las Perlas, clave 29005, se obtuvo de la base de datos BANDAS.

Las estaciones con registro del evento del 2010 se indican en la Tabla 4.10, cuya distribución espacial se aprecia en la Figura 4-35.

Tabla 4.10: Estaciones climatológicas para analizar la subcuenca de la E.H. Las Perlas.

Clave	Nombre	Latitud	Longitud	Clave	Nombre	Latitud	Longitud
20017	Zihualtepec	17°27'0"N	95°22'59"O	20506	María Lombardo	17°26'2"N	95°29'38"O
20117	San Miguel Chimalapa	16°42'43"N	94°44'53"O	30081	La Florencia	17°34'30"N	95°10'52"O
20134	Santiago Chivela	16°43'1"N	95°0'0"O	30090	Las Perlas	17°25'1"N	94°54'50"O
20289	Guevea de Humboldt	16°47'20"N	95°22'19"O	30224	Jesús Carranza	17°25'44"N	95°1'8"O
20374	Tierra Negra	17°13'1"N	95°27'0"O	30327	Tierra Morada	17°34'19"N	94°9'47"O

En la Tabla 4.11 se concentran los datos registrados por la red de pluviómetros, que representan la altura de lámina de lluvia en mm, precipitada durante el tren de lluvias ocurrido entre el 21 de septiembre y el 5 de octubre de 2010.

Tabla 4.11: Lluvia en estaciones climatológicas, 21 de septiembre al 5 de octubre de 2010, en mm.

Día	Clave de estación									
	20017	20117	20134	20289	20374	20506	30081	30090	30224	30327
21	0	15.5	29	46.2	8.2	13	2	48	14	38
22	12.5	56	60.5	8.3	26.5	25	0	27	1.6	3
23	12.8	6	1	1.8	12.8	140	74.5	129	35.5	25
24	27.4	6	3.9	52.5	17.3	15.2	33.5	38	57	122
25	94.5	90	22.4	71.5	56.7	48	80	43	49	64
26	74	27.7	201	49.2	76	98.5	80	108	133	23
27	235	119	104.5	251.8	221.4	213.8	153.5	37	74	36
28	7.5	56	30.5	23	14.5	15.3	19	16	12.5	80
29	0	0	0	0	0	0	0	0	0	0
30	0	24	0	0	0	0	0	0	5	0
1	0	0	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	0	2.3	0	0
3	0	0	0	0	0	0	0	0	0	6
4	0	2	0	0	0	0.2	0	0	0	27.5
5	12.4	11	0	40	14	8.6	27	17	23.8	29

Como complemento, se construyó también una representación de los trenes de lluvia diaria correspondientes a las estaciones climatológicas identificadas respecto de la avenida registrada en simultáneo a la salida de la subcuenca, útil para visualizar el comportamiento espacial y temporal de la precipitación respecto a la variación de los escurrimientos en la E.H. Las Perlas, durante el periodo señalado. En la Figura 4-36 se aprecian algunos detalles del evento.

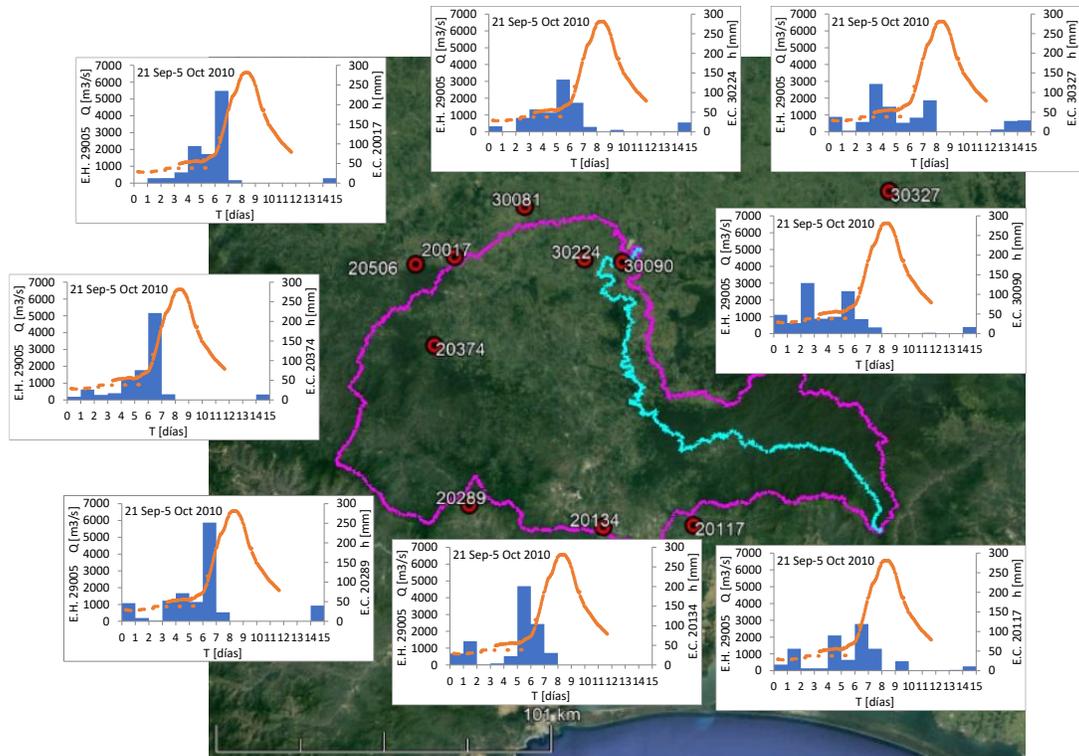


Figura 4-36: Variaciones de lluvia y escurrimiento, entre septiembre y octubre de 2010, subcuenca E.H Las Perlas.

En disposición de registros pluviométricos, se recopilaron las lluvias estimadas como productos satelitales del sistema en línea GIOVANNI, y de cada día del tren de lluvias se obtuvo una malla de celdas con láminas diarias, a una resolución de  $0.1^\circ \times 0.1^\circ$ . En la Figura 4-37 se aprecia la malla uniforme de 170 celdas considerada para asimilación, y empleada para representar la superficie de estudio en que está ubicada la subcuenca de la E.H. Las Perlas, en particular la distribución de la lluvia en mm correspondiente al 27 de septiembre de 2010, cuando se presentaron las láminas máximas del tren de lluvias de interés.

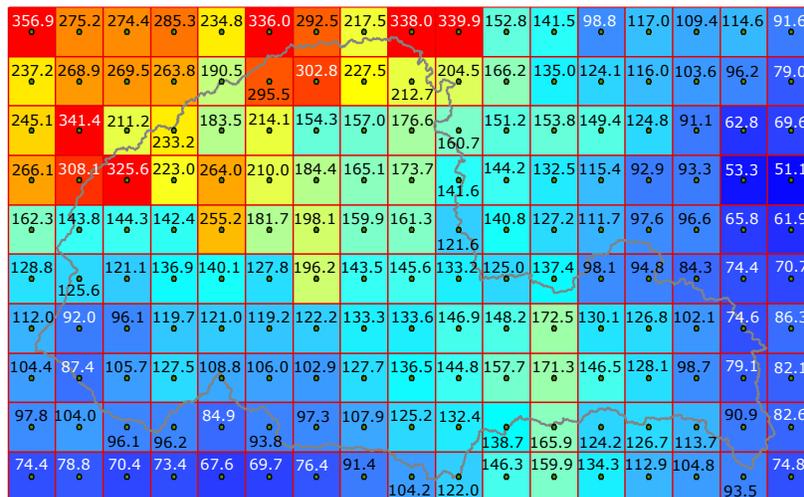


Figura 4-37: Distribución de lluvia estimada con satélite, 27 de septiembre de 2010, en mm.

Con los datos de lluvia distribuida sobre la malla de estudio, se llevó a cabo el procedimiento de calibración del modelo mediante la asignación ordenada de valores a los parámetros de convergencia  $\gamma$  y potencia  $\beta$ . La medición matemática del error de los valores estimados  $P_e$  respecto de los observados  $P_o$  en los campos de lluvia se practicó con el empleo de los criterios estadísticos *ME*, *MAE*, *RMSE*, *NSE* y *CC*, mientras que la evaluación del esquema se realizó con un procedimiento de validación cruzada, donde las observaciones de validación fueron los registros de las estaciones 30327 y 20117, de modo que la asimilación se realizó con un conjunto de observaciones cuya densidad y distribución guarda semejanza con la red original.

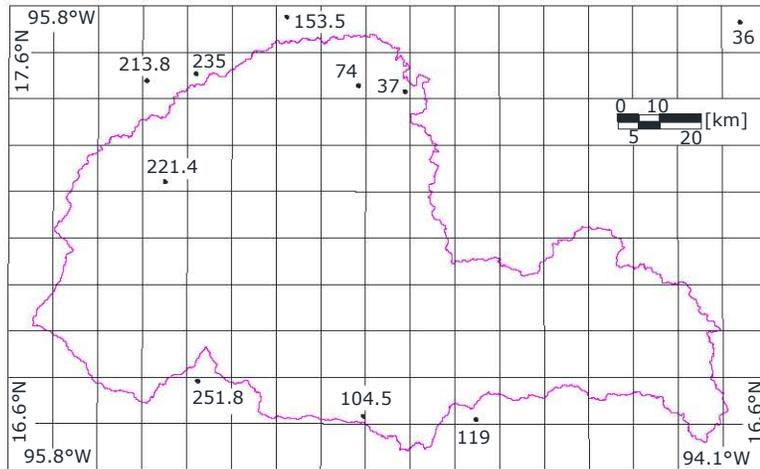
La calibración se ejecutó al asimilar la información de las bases de datos del 27 de septiembre, pues durante esa fecha se considera que el evento se desarrolló de manera generalizada sobre toda el área de interés, de modo que las mejores estimaciones del campo de lluvia se obtuvieron cuando los parámetros  $\gamma$  y  $\beta$  fueron iguales a 0.4 y 3 unidades, respectivamente.

Se procedió a la asimilación de cada uno de los quince días que componen al tren de lluvias del 21 de septiembre al 5 de octubre de 2010, obteniéndose la distribución espacial de las láminas diarias estimadas sobre las celdas de la malla de estudio. Como ejemplo, la Figura 4-38 muestra la distribución de lluvia del 27 de septiembre, donde pueden apreciarse las láminas correspondientes a la red de pluviómetros, las lluvias del producto satelital, las obtenidas por interpolación IDW de los registros pluviométricos considerados en la asimilación, así como las estimadas por asimilación.

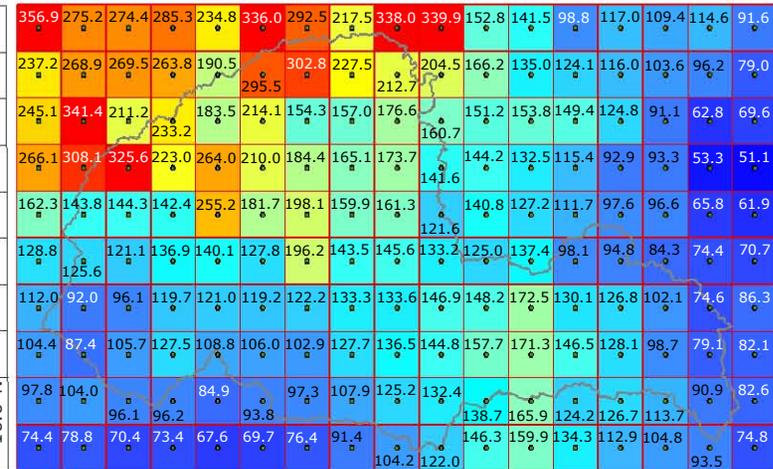
La Figura 4-39 es una representación condensada de los resultados de asimilar las dos bases de datos, así como los valores calculados al considerar cada base de manera individual, donde se muestran los criterios *ME*, *MAE*, *RMSE*, *NSE* y *CC*, correspondientes a los estadísticos promedio del análisis de los quince días del tren de lluvias estudiado. Asimismo, en la Figura 4-40 se presentan los mismos criterios estadísticos, pero respecto a los errores cometidos en la estimación de las láminas acumuladas durante los quince días del tren de lluvias.

De acuerdo con las figuras 4-39 y 4-40, tanto los errores promedio como los obtenidos de láminas acumuladas de las estimaciones diarias con interpolación IDW de los datos pluviométricos del tren de lluvias, pueden ser menores que los calculados con la asimilación de los datos pluviométricos y los de origen satelital. Dado que comúnmente esta circunstancia puede obedecer a la falta de exactitud de las estimaciones con satélite, se revisó la variación de los errores obtenidos de cada día asimilado, y se observó que en general las lluvias estimadas con satélite mostraron un orden de magnitud semejante a los registros pluviométricos, por lo que la falta de exactitud en las estimaciones satelitales de los días de menor precipitación fue menos significativa.

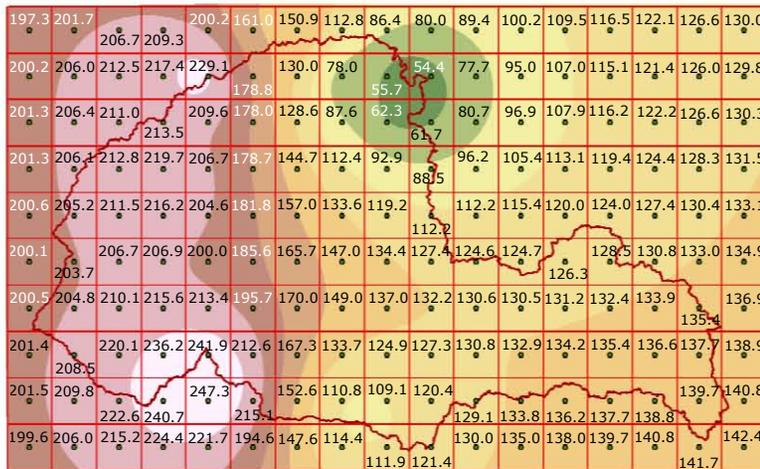
En este caso particular, la revisión permitió identificar que los errores (promedio y de acumulación) mayores se obtuvieron especialmente por los errores cometidos en las estimaciones del día 26 de septiembre, donde los registros pluviométricos considerados para validación tienen una magnitud mucho menor (de cuatro a seis veces menores) que los datos satelitales. Por esta razón, se realizó el cálculo de los criterios estadísticos del tren de lluvias, pero sin considerar las láminas del 26 de septiembre. La Figura 4-41 representa los estadísticos promedio respecto a los catorce días de lluvia considerados, mientras que la Figura 4-42 muestra los estadísticos de las lluvias acumuladas de los mismos catorce días, y se observa que en general las estimaciones por asimilación de las dos bases de datos son mejores aproximaciones que los valores obtenidos con las bases analizadas individualmente.



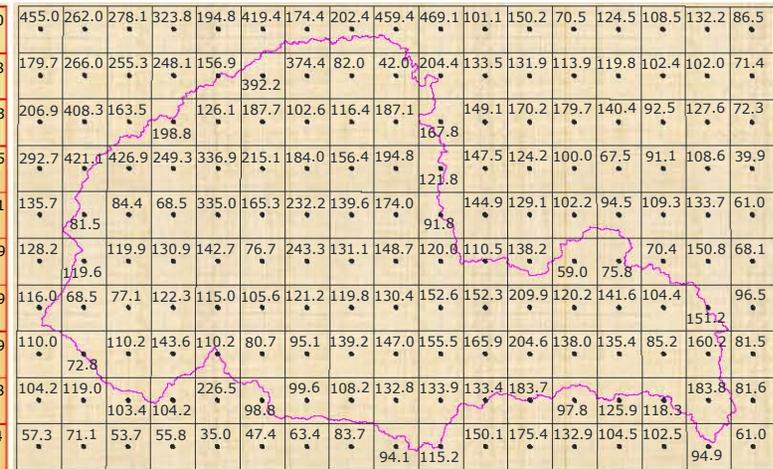
Distribución espacial de pluviómetros



Estimación de lluvia diaria por satélite



Estimación de lluvia diaria por interpolación IDW



Estimación de lluvia diaria por asimilación

Figura 4-38: Distribución espacial de lluvia, 27 de septiembre de 2010, en mm.

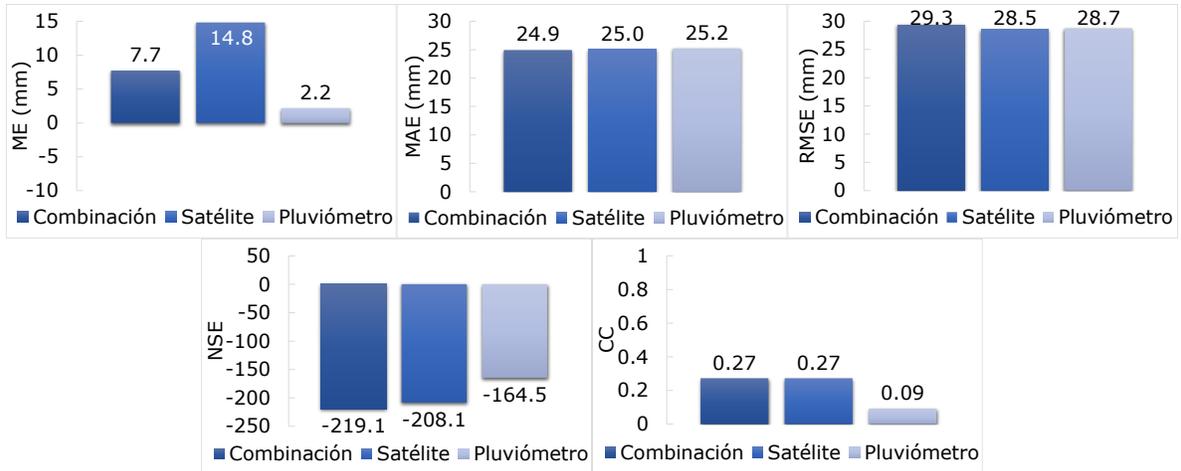


Figura 4-39: Criterios estadísticos promedio de los quince días del tren de lluvias, evento de 2010.

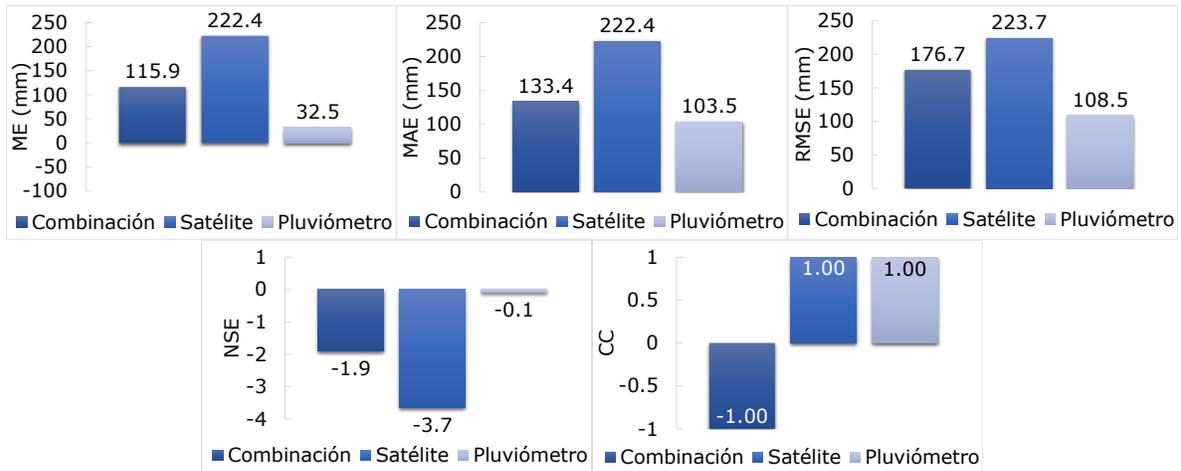


Figura 4-40: Criterios estadísticos del tren de lluvias acumulado, evento de 2010.

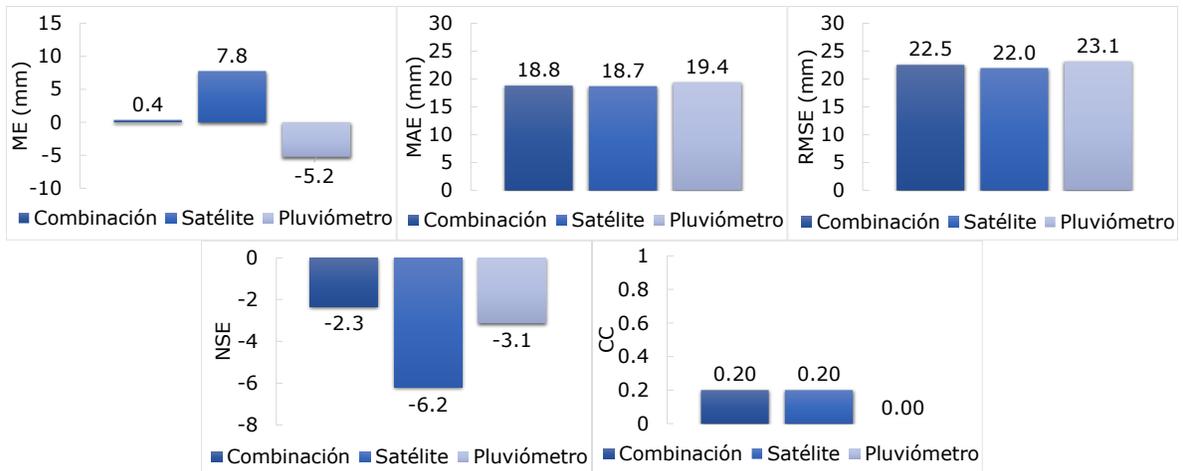


Figura 4-41: Criterios estadísticos promedio de catorce días del tren de lluvias, evento de 2010.

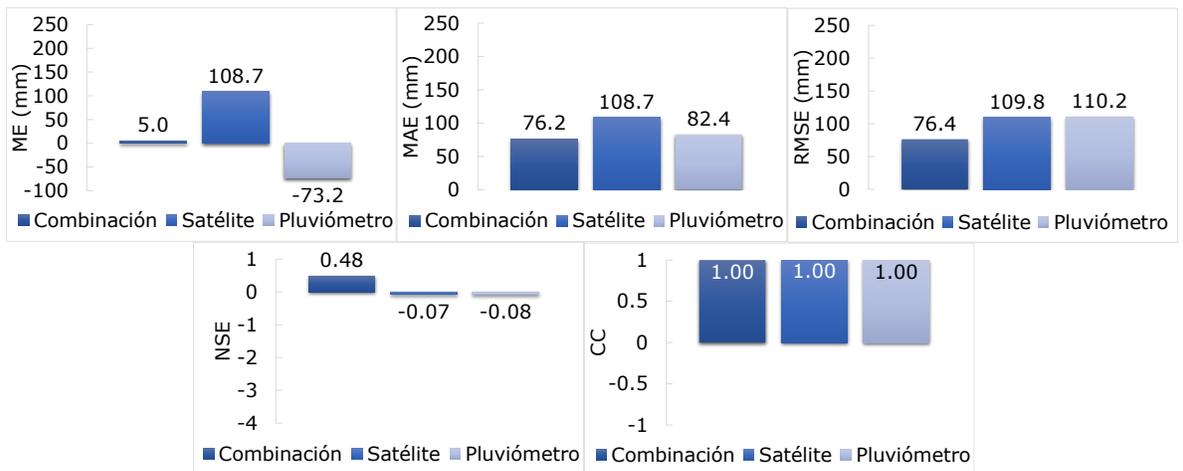


Figura 4-42: Criterios estadísticos de las lluvias acumuladas de catorce días, evento de 2010.

## **5. Consideración de lluvia distribuida espacialmente como insumo de un modelo para la simulación de escurrimiento superficial**

### **5.1. La modelación de escurrimientos en la revisión de la calidad de estimaciones de lluvia distribuida espacialmente**

Debido a que el objetivo principal de esta investigación es proponer un procedimiento que permita obtener mejores estimaciones de la distribución espacial de lluvia diaria, principalmente sobre regiones relativamente amplias del territorio mexicano, se ha optado por una metodología basada en el concepto de asimilación.

Tomando en cuenta que la disponibilidad de registros pluviométricos generalmente no es abundante ni está distribuida de manera regular, estos datos se combinaron con productos de percepción remota, alcanzándose resultados satisfactorios en comparación con las estimaciones obtenidas al considerar la información de cada una de las bases de datos de manera individual, tanto para lluvias diarias como para láminas acumuladas durante varios días consecutivos o trenes de lluvia. En estas condiciones, resultó de interés complementar la revisión de la calidad de la lluvia estimada por asimilación, al considerarla como el insumo principal de un análisis de escurrimiento superficial por lluvia en cuenca propia.

Esta revisión fue factible debido a que el análisis de precipitación realizado sobre cada una de las regiones estudiadas inició con la definición de la subcuenca de captación, a partir de la ubicación de una estación hidrométrica con registro de gastos en las fechas correspondientes a los trenes de lluvia considerados: estaciones 30032 Teapa, 24327 Cadereyta y 29005 Las Perlas.

### **5.2. Modelación de escurrimientos superficiales**

La precipitación puede considerarse como el estímulo capaz de producir una respuesta de escurrimiento dentro de una cuenca, que es una porción de superficie terrestre sobre la cual las gotas de lluvia son drenadas por gravedad hacia un punto (de salida en cuenca exorreica), a través de un conjunto de corrientes naturales o red de drenaje, conformado por un cauce principal y sus incorporaciones provenientes de zonas elevadas denominadas corrientes tributarias [34].

El inicio de la relación entre la precipitación y el escurrimiento empieza desde que las gotas de lluvia ingresan a la cuenca, y una parte del agua escurre sobre la superficie terrestre (escurrimiento superficial), otra se infiltra (flujo subterráneo), y una tercera atraviesa la superficie terrestre sin llegar al manto freático (escurrimiento subsuperficial). El escurrimiento superficial es el más rápido

de los flujos, e inicialmente consiste de una lámina de agua delgada que va incrementándose al aproximarse a la salida de la cuenca. Por esta razón, el gasto de salida es creciente en el tiempo, aunque posteriormente alcanza una condición de equilibrio y se vuelve constante, pero en el cese de la tormenta el gasto comienza a disminuir, hasta presentarse el final del escurrimiento directo [34].

Debido a que el área de una cuenca corresponde con la superficie de proyección horizontal de la región encerrada por su parteaguas, la cantidad de precipitación captada se relaciona con el tamaño de dicha superficie, y en consecuencia la magnitud de los escurrimientos se incrementa conforme el área también aumenta, aunque cabe señalar que la forma de la cuenca, que es la configuración geométrica del parteaguas, está más relacionada con el tiempo de recorrido de una gota de lluvia a través de la red de drenaje [34].

La pendiente de la cuenca es un factor que controla el tiempo del flujo de agua sobre el terreno, pues conforme la pendiente aumenta, la velocidad media del escurrimiento es mayor, provocando menores infiltraciones, así como un incremento en la magnitud de los gastos máximos de las avenidas. Respecto al cauce principal, su longitud se relaciona con la respuesta hidrológica, de manera que los cauces cortos se manifiestan más rápidamente, y los gastos aumentan conforme la longitud es más grande. Asimismo, la pendiente del cauce tiene un efecto en la duración de los escurrimientos, especialmente en el cálculo del tiempo de ocurrencia del gasto máximo [34].

En resumen, se tiene que el volumen del escurrimiento superficial está condicionado por factores climáticos como el balance entre la precipitación y la evapotranspiración, además de factores fisiográficos como el área de captación de lluvia, la pendiente, el tipo y uso de suelo, y la altitud de la cuenca. Por otra parte, respecto a la capacidad de respuesta de una cuenca que define la distribución de los escurrimientos en el tiempo, influyen factores meteorológicos como el tipo, intensidad, duración, distribución y trayectoria de la precipitación, además de factores fisiográficos como la topografía, geología, el suelo y la red de drenaje, y factores antropogénicos como estructuras hidráulicas, prácticas agrícolas y urbanización [34].

Desde un punto de vista práctico, es recurrente la necesidad de conocer la magnitud de escurrimientos para el diseño y operación de obras de protección y aprovechamientos hidráulicos, aunque es más común que el aforo de corrientes en los sitios de interés sea inexistente, o el registro de gastos sea insuficiente en cantidad o calidad. Es bajo estas condiciones que conviene tomar en cuenta que la precipitación es la variable ligada directamente al escurrimiento de cuyos registros se tiene menor incertidumbre respecto a la ocurrencia y magnitud de eventos históricos.

Anteriormente se ha discutido que la modelación es el método a partir del cual se crea un objeto modelado con los rasgos esenciales de un objeto real, para simular particularidades del comportamiento del sistema original de una manera más simple, dada la necesidad de omitir algunos rasgos del objeto real por su menor importancia o complejidad.

En el ámbito de las relaciones entre la precipitación y el escurrimiento, el análisis puede realizarse considerando solo aspectos hidrológicos, o mediante la modelación de la hidrodinámica de los cursos fluviales [45].

El flujo como función del tiempo se analiza en un modelo hidrológico a partir de la ecuación de la conservación de la masa, con un balance entre los volúmenes de entrada, salida y almacenamiento de agua en una cuenca. Entre los modelos hidrológicos clasificados por las características de los datos espaciales de entrada al sistema, los métodos de parámetros concentrados o agregados son aquellos en que la conversión de lluvia a escurrimiento se realiza

utilizando una distribución uniforme de los promedios espaciales de la precipitación y de las características fisiográficas de la cuenca, de modo que con una función de transferencia el volumen de precipitación efectiva se transforma en escurrimiento directo.

Con los modelos de parámetros distribuidos es factible considerar los atributos principales de las unidades areales en que convenga dividir una región de estudio, para describir las variaciones de lluvia y las características del terreno, de manera que utilizando modelos de producción se calcula la porción de lluvia que no se transforma en escurrimiento, y mediante modelos de transferencia se estima la atenuación por almacenamiento y la descarga de escurrimientos hacia la salida de la cuenca [34].

Por otro lado, los modelos que toman en cuenta la dinámica de los flujos dentro de una cuenca permiten la simulación de los escurrimientos superficiales a través de las secciones transversales de un cauce principal, y su conexión con los demás elementos de la red de drenaje, además de tomar en consideración la posible conectividad lateral entre un río y su planicie de inundación, de tal manera que puede establecerse la transformación de hietogramas de precipitación distribuida espacial y temporalmente en hidrogramas de escurrimiento, con la opción de analizar simultáneamente el ingreso de escurrimientos a través de las fronteras del sistema.

Como resultados pueden obtenerse la variación de los gastos, la velocidad de las corrientes y la magnitud de las profundidades hidráulicas a través del tiempo y en cualquier punto de la región de estudio, información de utilidad para conocer la atenuación de una avenida hacia la salida de la cuenca, ya sea por infiltración o por desbordamiento sobre las superficies adyacentes, y en consecuencia establecer zonas de inundación probables y los tiempos de permanencia de los anegamientos correspondientes a eventos asociados con determinado periodo de retorno. Por lo anterior, el nivel de certidumbre en los resultados de los modelos hidrodinámicos se relaciona especialmente con la calidad de los datos hidrológicos distribuidos espacial y temporalmente, la caracterización suficiente de los rasgos topográficos y de los tipos y usos del suelo, la consideración apropiada de las condiciones de frontera del sistema, y la capacidad de simular la intermitencia de los fenómenos de precipitación y escurrimiento durante un tren de tormentas, como consecuencia de una modelación adecuada de los procesos de conectividad entre los cauces de la red de drenaje, los cuerpos de agua existentes y las planicies de inundación [28].

Se sabe que la disipación de la energía que ocurre durante el movimiento de un fluido viscoso está definida por la solución de las ecuaciones de Navier-Stokes, pues éstas describen la relación que existe entre la velocidad, presión, temperatura y densidad de un fluido en movimiento, analizado en consideración de la 2ª Ley de Newton [75].

Dado que cualquier fenómeno físico se desarrolla dentro de un sistema de tres dimensiones espaciales y una dimensión temporal, el conjunto de ecuaciones en diferenciales parciales de Navier-Stokes consiste de una ecuación de continuidad dependiente del tiempo para la conservación de la masa, tres ecuaciones de conservación de la cantidad de movimiento dependientes del tiempo, y una ecuación de conservación de energía dependiente del tiempo. El problema está conformado por cuatro variables independientes, que incluyen las tres coordenadas espaciales del dominio, además del tiempo, así como seis variables dependientes: la presión, la densidad, la temperatura y las tres componentes del vector de velocidad. En estas condiciones, un flujo se obtiene con la solución simultánea del correspondiente sistema de cinco ecuaciones, y por ello se requiere la consideración adicional de una ecuación de estado que relacione la presión, temperatura y densidad del fluido [60].

Una simplificación del análisis tridimensional de los flujos se practica al considerar que el movimiento se desarrolla en un espacio donde la escala de longitud horizontal es mucho mayor que la vertical. Las ecuaciones de Saint-Venant en dos dimensiones, conocidas como ecuaciones en aguas poco profundas, tienen su origen en la integración en profundidad de las ecuaciones de Navier-Stokes, de modo que del análisis de la conservación de la masa se establece que la escala de velocidad horizontal del fluido es muy grande respecto a la vertical, y del análisis de la conservación de la cantidad de movimiento se observa que los gradientes de presión verticales son casi hidrostáticos, mientras que los gradientes de presión horizontales surgen por el desplazamiento de la superficie de presión, y en consecuencia, el campo de velocidades en la horizontal tiene una distribución relativamente uniforme en la profundidad [12].

Como una herramienta de uso libre para la simulación del flujo en ríos y estuarios, la plataforma Iber consiste en un modelo bidimensional basado en volúmenes finitos, desarrollado desde la administración pública española, en colaboración del Grupo de Ingeniería del Agua y del Medio Ambiente, del Grupo de Ingeniería Matemática, del Instituto Flumen, y promovido por el Centro de Estudios Hidrográficos del CEDEX [13]. El módulo hidrodinámico de Iber resuelve las ecuaciones de Saint-Venant bidimensionales, incorporando los efectos de la turbulencia y rozamiento superficial por viento [45]:

$$\frac{\partial h}{\partial t} + \frac{\partial h U_x}{\partial x} + \frac{\partial h U_y}{\partial y} = M_s \quad (5.1)$$

$$\frac{\partial h U_x}{\partial t} + \frac{\partial h U_x^2}{\partial x} + \frac{\partial h U_x U_y}{\partial y} = -gh \frac{\partial Z_s}{\partial x} + \frac{\tau_{s,x}}{\rho} - \frac{\tau_{b,x}}{\rho} - \frac{g}{\rho} \frac{h^2}{2} \frac{\partial \rho}{\partial x} + 2\Omega \text{sen} \lambda U_y + \frac{\partial h \tau_{xx}^e}{\partial x} + \frac{\partial h \tau_{xy}^e}{\partial y} + M_x \quad (5.2)$$

$$\frac{\partial h U_y}{\partial t} + \frac{\partial h U_x U_y}{\partial x} + \frac{\partial h U_y^2}{\partial y} = -gh \frac{\partial Z_s}{\partial y} + \frac{\tau_{s,y}}{\rho} - \frac{\tau_{b,y}}{\rho} - \frac{g}{\rho} \frac{h^2}{2} \frac{\partial \rho}{\partial y} + 2\Omega \text{sen} \lambda U_x + \frac{\partial h \tau_{xy}^e}{\partial x} + \frac{\partial h \tau_{yy}^e}{\partial y} + M_y \quad (5.3)$$

donde:

$h$	tirante
$U$	velocidades horizontales promediadas en la profundidad
$g$	aceleración de la gravedad
$Z_s$	elevación de la lámina libre
$\tau_s$	fricción en la superficie libre debida al rozamiento producido por el viento
$\tau_b$	fricción debida al rozamiento del fondo
$\rho$	densidad del agua
$\Omega$	velocidad angular de rotación de la Tierra
$\lambda$	latitud del punto considerado
$\tau^e$	tensiones tangenciales efectivas horizontales
$M$	términos de masa y de momento, para la modelación de precipitación, infiltración y sumideros.

Dada la intención de complementar la revisión de la calidad de la lluvia diaria estimada con el método de asimilación propuesto, de los eventos analizados en las subcuencas de las estaciones hidrométricas Teapa, Cadereyta y Las Perlas, se consideró la lluvia distribuida espacialmente como el insumo de un análisis de escurrimiento superficial por lluvia en cuenca propia.

Debido a que existen antecedentes satisfactorios del empleo del modelo Iber para el cálculo del hidrograma a la salida de una cuenca, con respecto a los resultados de un modelo de parámetros distribuidos [45], por ejemplo, a continuación se describen los procesos realizados para la obtención de hidrogramas a la salida de las subcuencas indicadas, correspondientes a las lluvias

estimadas por asimilación y las obtenidas por interpolación IDW de los registros pluviométricos considerados en la asimilación, para su comparación con los gastos registrados en las estaciones hidrométricas localizadas a la salida de cada región de estudio.

### 5.2.1. Escurrimientos superficiales en la subcuenca de la E.H. Teapa

El análisis de lluvia diaria sobre la subcuenca (418.6 km<sup>2</sup>) con aportación de escurrimiento superficial al sitio de la estación hidrométrica E.H. 30032 Teapa, en el estado de Tabasco, se llevó a cabo atendiendo la información del tren de lluvias desarrollado del 19 al 26 de noviembre de 2015, estimándose la distribución espacial de las láminas diarias sobre las veinte celdas que componen la malla de estudio. A manera de ejemplo, la Figura 5-1 presenta la distribución de lluvia correspondiente al 23 de noviembre, donde se aprecian por separado las láminas estimadas por asimilación y las calculadas mediante la interpolación IDW de los registros pluviométricos considerados en el proceso de asimilación.

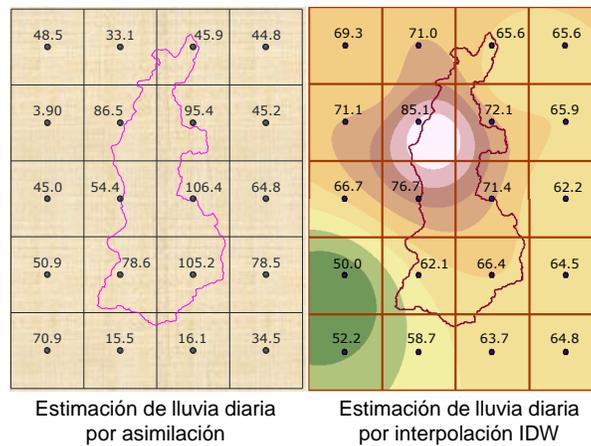


Figura 5-1: Distribución de lluvia en mm, noviembre 23 de 2015, subcuenca E.H. Teapa.

Como el volumen del escurrimiento superficial depende de la lámina de precipitación en exceso o lluvia efectiva, es necesario realizar su estimación correspondiente. Para el modelado de la infiltración en el terreno, Iber considera a la tasa de infiltración real como un término fuente negativo en la ecuación de conservación de la masa, y la define como el mínimo entre la tasa de infiltración potencial y la cantidad de agua superficial disponible para infiltrarse, donde la infiltración potencial se calcula con la implementación del modelo de Green-Ampt o el de Horton [45].

En vista de que la aplicación apropiada de los modelos de infiltración requiere de información que generalmente no está disponible en la cantidad o calidad necesarias para caracterizar las condiciones propias de cualquier región de México, como es el caso de la permeabilidad, porosidad y humedad inicial del suelo, la estimación de la lluvia efectiva en cada celda de la malla se obtuvo con el método de los números de escurrimiento.

El método de los Números de Escurrimiento, propuesto por el U.S. Soil Conservation Service, permite la estimación directa de la lámina de lluvia efectiva que produce una tormenta aislada, tomando en cuenta los tipos y usos de suelo y su vegetación, así como la pendiente del terreno y sus condiciones iniciales de humedad [3]. Dado un evento de tormenta que aporta la totalidad de agua susceptible de escurrir superficialmente dentro de una cuenca, la profundidad de la lámina de

lluvia efectiva  $P_e$  que produce escurrimiento directo siempre es menor o igual que la profundidad de la precipitación total  $P$ . Como parte del método se considera que existe una cierta cantidad de precipitación  $I_0$  para la cual no ocurrirá escurrimiento (lámina de infiltración inicial antes del encharcamiento), por lo que el escurrimiento potencial puede calcularse como la diferencia entre la precipitación  $P$  y la infiltración  $I_0$ . Asimismo, después de iniciado el escurrimiento, la lámina adicional de agua absorbida por el suelo,  $F$ , es menor o igual que alguna infiltración potencial máxima  $S$ . Del principio de continuidad se establece [20]:

$$P = P_e + I_0 + F \quad (5.4)$$

La hipótesis del método consiste en que las relaciones de las dos cantidades reales y las dos cantidades potenciales son iguales, es decir, se acepta una proporcionalidad entre el umbral de lámina de pérdida y el umbral de lámina de infiltración máxima (diferencia potencial máxima entre la lluvia y el escurrimiento), respecto de la lluvia efectiva y la lluvia efectiva máxima [28]:

$$\frac{F}{S} = \frac{P_e}{P - I_0} \quad (5.5)$$

Combinando las ecuaciones (5.4) y (5.5) se obtiene la expresión para el cálculo de la lluvia efectiva:

$$P_e = \frac{(P - I_0)^2}{P - I_0 + S} \quad (5.6)$$

donde se recomienda utilizar una relación empírica entre las láminas de infiltración inicial y potencial, siendo  $I_0$  el 20% de la lámina de infiltración máxima  $S$ , que en mm puede estimarse con:

$$S = \frac{25400}{N} - 254 \quad (5.7)$$

Respecto a las condiciones de humedad antecedente, como resultado de la precipitación consecutiva de varias láminas de lluvia diaria (tren de tormentas), se aplicó una metodología propuesta por Fuentes *et al.* 2014, que toma en cuenta el concepto del índice de precipitación antecedente:

$$I_{ant} = CI_{i-1} \quad (5.8)$$

donde  $I_{i-1}$  es la lámina de infiltración del día anterior;  $C$  es el coeficiente de recesión, de cuyo valor depende la pérdida progresiva de la magnitud de la lámina de infiltración respecto al tiempo, con un valor de 0.7 para lluvias diarias.

En el cálculo de la lluvia efectiva que considera la infiltración del día previo, se supone que el valor de la infiltración máxima  $S$  no es variable, y que la lámina de infiltración mínima  $I_0$  se mantiene como el 20% de la infiltración máxima. La infiltración  $I$  correspondiente a una lámina de precipitación total, cuya magnitud es superior a la lámina de infiltración mínima ( $P \geq 0.2S$ ), se calcula mediante:

$$I = (P - P_e) + I_{ant} \quad (5.9)$$

Al considerar la ecuación (5.9) en la ecuación de continuidad (5.4), la lámina de agua infiltrada  $F$  resulta:

$$F = I - I_0 = (P - P_e) + I_{ant} - I_0 \quad (5.10)$$

Sustituyendo la ecuación (5.10) en la ecuación (5.5):

$$P_e = (P - I_0) \left( \frac{(P - I_0) - P_e + I_{ant}}{S} \right) \quad (5.11)$$

Después de un proceso algebraico, la lámina de lluvia efectiva puede estimarse como:

$$P_e = \frac{(P - 0.2S)^2}{P + 0.8S} + \frac{(P - 0.2S)}{P + 0.8S} (I_{ant}) \quad (5.12)$$

donde el segundo término del lado derecho de la ecuación no forma parte del análisis tradicional del método de los números de escurrimiento, pero se considera que su aplicación mejora la estimación de la lluvia efectiva [28].

En estas condiciones, los valores de la lluvia efectiva  $P_e$  y la infiltración  $I$  correspondiente a la precipitación total, calculados respectivamente con las ecuaciones (5.12) y (5.9), se consideran correctos siempre que  $0.2S \leq I \leq S$ .

En caso de que  $I < 0.2S$ , considerar que  $I = 0.2S$ , y la lluvia efectiva resulta:

$$P_e = P - I \quad \text{si } P \geq I \quad (5.13)$$

$$P_e = 0 \quad \text{si } P < I \quad (5.14)$$

En el caso de que  $I > S$ , considerar que  $I = S$ , y la lluvia efectiva resulta:

$$P_e = P - S \quad (5.15)$$

Para el cálculo de la lluvia efectiva en cada celda de la malla que representa la región de estudio, se obtuvo el valor medio del parámetro  $N$  por celda, denominado Número de la Curva de Escurrimiento del Complejo Hidrológico Suelo-Cobertura [19]. Respecto a la ecuación (5.7), el número  $N$  depende del uso de suelo y vegetación de la superficie de escurrimiento, y varía desde cero hasta 100 para superficies impermeables o cuerpos de agua, tomándose en cuenta el grupo hidrológico al que pertenece el tipo de suelo (muy permeable a impermeable), además de la pendiente del terreno. Aunque existe una diversidad de tablas con información del número  $N$ , se consideró conveniente tomar en cuenta los valores que se indican en el Manual del Modelo para pronóstico de escurrimiento [29], porque es una recopilación de varias publicaciones, con la ventaja de que las definiciones en el uso de suelo son las que se utilizan en los mapas publicados por el INEGI.

Se consultó la información disponible en el Portal de Geoinformación de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, particularmente el conjunto de datos vectoriales de la edafología, escala 1:250 000 – 1:1 000 000 (1995), así como los correspondientes al uso de suelo y vegetación, escala 1:250 000 (Serie V, 2013). En la Figura 5-2 se muestra la distribución de los diferentes tipos y usos de suelo y su vegetación sobre la región de estudio, además del valor medio del número de escurrimiento  $N$  que le corresponde a cada celda de la malla.

En conocimiento de las láminas de lluvia efectiva distribuidas sobre las celdas de la malla de estudio, se procedió a establecer la distribución temporal de dichas láminas diarias, las que forman parte del tren de lluvias ocurrido entre el 19 y el 26 de noviembre de 2015. Al no disponer de registros de la variación temporal de la lluvia sobre la región y fechas de interés, se consideró pertinente tomar en cuenta la relación que existe entre las precipitaciones de máxima intensidad de una hora de duración con las acumulaciones correspondientes a 24 horas.

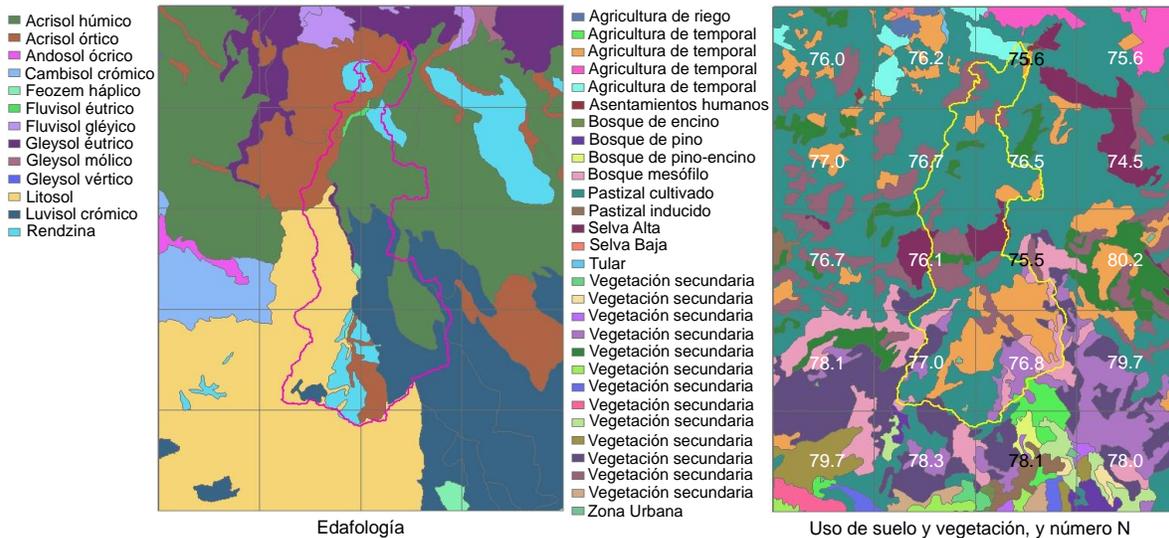


Figura 5-2: Distribución de tipos y usos de suelo y su vegetación, y números N correspondientes.

Como resultado de sus investigaciones, Chen encontró que el cociente R, que relaciona la altura de lluvia con la duración (1 hora y 24 horas), varía entre 10% y 60% con la ubicación geográfica, y que probablemente es independiente del periodo de retorno. Apoyado en estudios previos, Baeza se concentró en el análisis del factor R para la República Mexicana, al que denominó factor de conectividad, el cual es difícil de estimar con precisión debido a la escasez de información pluviográfica en México. Con un manejo regional de la información pluviográfica recopilada, utilizó el concepto de conectividad al considerar las características del relieve en el entorno de los sitios de medición, y su relación con los fenómenos meteorológicos predominantes por región. La Figura 5-3 muestra el mapa de los factores de conectividad propuestos por Baeza [54].

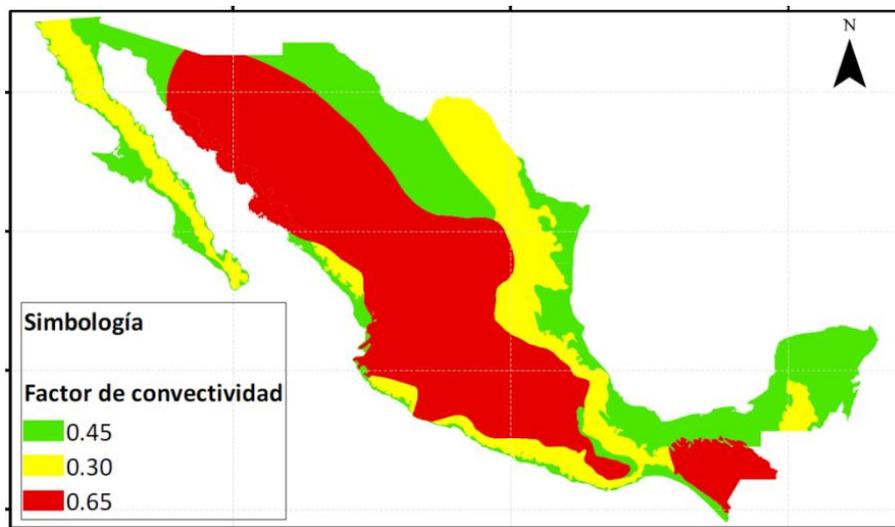


Figura 5-3: Mapa de Factores de Conectividad. Fuente: [54].

De acuerdo con la ubicación de la región de estudio, y considerando a las celdas de la malla de análisis como la distribución de la lluvia dentro de un arreglo matricial, a las celdas de los tres primeros renglones (de norte a sur) les corresponde un factor de convectividad R de 0.45, mientras que a las celdas de los dos renglones ubicados en la zona sur del arreglo les corresponde un factor R de 0.65. Adicionalmente, y de acuerdo con los resultados obtenidos por Chen, para estimar los valores asociados a otras duraciones menores que 24 h, se multiplicó la precipitación correspondiente a 1 h por los factores de reducción K ( $K = P_d^{Tr}/P_1^{Tr}$ ), indicados en la Tabla 5.1, que son valores modificados respecto a una tabla obtenida por Luna y Domínguez en 2013, [54].

Tabla 5.1: Relación  $K = P_d^{Tr}/P_1^{Tr}$ , en función de R y la duración d, [54].

d (min)	d (h)	K en función de d y del factor de convectividad $R = P_1^{Tr}/P_{24}^{Tr}$						
		R=0.10	R=0.20	R=0.30	R=0.40	R=0.50	R=0.60	R=0.65
10	0.17	0.293	0.390	0.432	0.454	0.469	0.481	0.487
15	0.25	0.380	0.485	0.536	0.565	0.584	0.600	0.608
30	0.50	0.612	0.699	0.745	0.773	0.793	0.809	0.816
60	1.00	1.000	1.000	1.000	1.000	1.000	1.000	1.000
120	2.00	1.646	1.424	1.317	1.250	1.203	1.166	1.151
150	2.50	1.934	1.595	1.435	1.337	1.268	1.215	1.193
180	3.00	2.207	1.750	1.538	1.410	1.322	1.254	1.226
210	3.50	2.468	1.892	1.631	1.475	1.367	1.286	1.253
240	4.00	2.719	2.024	1.715	1.532	1.407	1.314	1.275
270	4.50	2.961	2.148	1.793	1.584	1.443	1.337	1.294
300	5.00	3.196	2.266	1.865	1.631	1.475	1.358	1.311
330	5.50	3.425	2.378	1.933	1.675	1.504	1.377	1.326
360	6.00	3.649	2.485	1.997	1.716	1.531	1.395	1.339
390	6.50	3.867	2.587	2.057	1.755	1.556	1.410	1.351
420	7.00	4.081	2.686	2.115	1.791	1.579	1.425	1.362
450	7.50	4.291	2.781	2.170	1.825	1.601	1.438	1.373
480	8.00	4.497	2.874	2.223	1.858	1.621	1.451	1.382
510	8.50	4.700	2.963	2.273	1.889	1.641	1.463	1.391
540	9.00	4.899	3.050	2.322	1.919	1.659	1.474	1.399
570	9.50	5.096	3.134	2.369	1.947	1.677	1.484	1.407
600	10.00	5.289	3.216	2.414	1.975	1.694	1.494	1.415
630	10.50	5.480	3.297	2.458	2.001	1.710	1.504	1.422
660	11.00	5.669	3.375	2.501	2.026	1.725	1.513	1.429
690	11.50	5.855	3.452	2.542	2.051	1.740	1.522	1.435
720	12.00	6.039	3.527	2.582	2.074	1.754	1.530	1.441
750	12.50	6.221	3.600	2.622	2.097	1.768	1.538	1.447
780	13.00	6.401	3.672	2.660	2.120	1.781	1.545	1.452
810	13.50	6.580	3.743	2.697	2.141	1.794	1.553	1.458
840	14.00	6.756	3.812	2.734	2.162	1.807	1.560	1.463
870	14.50	6.931	3.880	2.769	2.183	1.819	1.567	1.468
900	15.00	7.104	3.947	2.804	2.203	1.831	1.573	1.473
930	15.50	7.275	4.013	2.838	2.222	1.842	1.580	1.477
960	16.00	7.445	4.078	2.872	2.241	1.853	1.586	1.482
990	16.50	7.614	4.142	2.904	2.260	1.864	1.592	1.486
1 020	17.00	7.781	4.205	2.936	2.278	1.875	1.598	1.490
1 050	17.50	7.947	4.266	2.968	2.296	1.885	1.604	1.495
1 080	18.00	8.112	4.328	2.999	2.313	1.895	1.609	1.499
1 110	18.50	8.275	4.388	3.029	2.330	1.905	1.615	1.502
1 140	19.00	8.437	4.447	3.059	2.347	1.914	1.620	1.506
1 170	19.50	8.598	4.506	3.088	2.363	1.924	1.625	1.510
1 200	20.00	8.758	4.564	3.117	2.379	1.933	1.630	1.513
1 230	20.50	8.917	4.621	3.146	2.395	1.942	1.635	1.517
1 260	21.00	9.075	4.678	3.174	2.411	1.951	1.640	1.520
1 290	21.50	9.232	4.734	3.201	2.426	1.959	1.645	1.524
1 320	22.00	9.388	4.789	3.228	2.441	1.968	1.649	1.527
1 350	22.50	9.542	4.843	3.255	2.455	1.976	1.654	1.530
1 380	23.00	9.696	4.897	3.282	2.470	1.984	1.658	1.533
1 410	23.50	9.849	4.951	3.308	2.484	1.992	1.662	1.536
1 440	24.00	10.001	5.004	3.333	2.498	2.000	1.667	1.539

Como ejemplo, la Figura 5-4 presenta la distribución temporal de la lluvia efectiva por asimilación en un acomodo en bloque alterno, correspondiente a la celda del cuarto renglón - tercera columna del arreglo matricial de la lluvia, espacio donde se localiza la sección del cauce principal más alejada respecto a la E.H. Teapa.

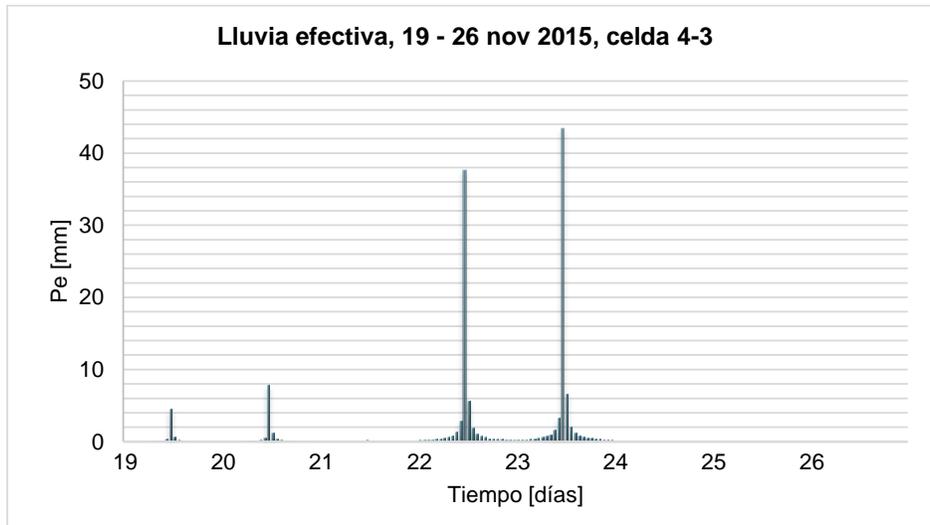


Figura 5-4: Hietograma de lluvia efectiva estimada por asimilación, celda 4-3, en intervalos de una hora y duración total de ocho días.

En disposición de lluvia efectiva estimada tanto por asimilación como por interpolación IDW, distribuida espacial y temporalmente durante el periodo del 19 al 26 de noviembre del 2015, se procedió a la simulación de la transformación de dicha lluvia en escurrimiento directo hasta el sitio de la E.H. Teapa, utilizando el módulo hidrodinámico de la herramienta Iber. En la simulación se utilizó la información topográfica digital generada por el INEGI, particularmente el continuo de elevaciones mexicano 3.0 (celdas de 15 m de resolución) de los datos vectoriales de los estados de Chiapas y Tabasco. Asimismo, se consideró un conjunto de valores del coeficiente de rugosidad de Manning para el análisis del flujo sobre la superficie general de la subcuenca y dentro del cauce, con las magnitudes siguientes [21]: agricultura 0.04, pastizal 0.03, vegetación secundaria 0.035, zona urbana 0.016 y cauce del río 0.025.

La Figura 5-5 muestra la variación de los escurrimientos superficiales obtenidos mediante análisis hidrodinámico, correspondientes a los esquemas de lluvia estimada por asimilación y por interpolación IDW, además del comportamiento de la avenida considerada real, definida por de Luna (2019) en consideración de registros de escalas, dado que en el registro de la E.H. Teapa, disponible en la base de datos BANDAS, no se tiene información de los gastos correspondientes al periodo de estudio.

Como puede apreciarse, existe una coincidencia notable en el tiempo de inicio y de finalización de la avenida, así como una semejanza en cuanto al orden de magnitud de los gastos máximos alcanzados a consecuencia de la variación de las precipitaciones durante el tren de lluvias analizado, en especial las lluvias estimadas con el método de asimilación propuesto, pues los escurrimientos asociados con las láminas obtenidas por interpolación IDW manifiestan una sobreestimación más grande.

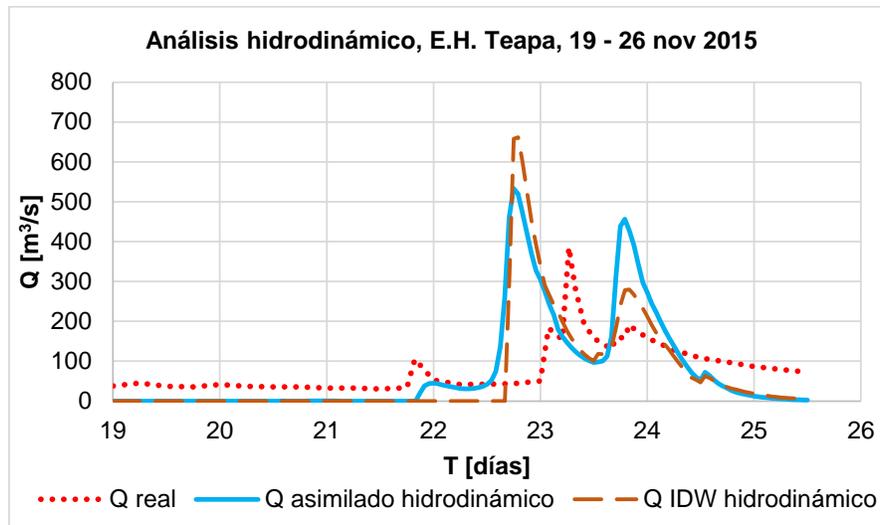


Figura 5-5: Escurrimiento superficial, real y simulado con análisis hidrodinámico, sitio de la E.H. Teapa.

Como complemento a la revisión de la calidad de las lluvias estimadas con el método de asimilación propuesto, se realizó también el cálculo de la avenida al sitio de la E.H. Teapa con un procedimiento hidrológico de parámetros concentrados, pues dada la magnitud de la superficie de análisis el resultado podría ser comparable al de una simulación hidrodinámica, pero con requerimientos de información y tiempo para ejecución sensiblemente menores.

De acuerdo con la información topográfica digital del INEGI considerada, la subcuenca de área A igual a 418.64 km<sup>2</sup>, tiene una longitud y pendiente media de cauce principal de 60 048.46 m y 0.0088, respectivamente, y con la expresión de Kirpich se calculó un tiempo de concentración ( $t_c$ ) de 9.59 horas. El tiempo y gasto de pico se obtuvieron bajo las consideraciones siguientes [34]:

$$\text{Tiempo de retraso: } t_r = 0.6t_c = 5.75 \text{ h}$$

$$\text{Duración en exceso: } d_e = \frac{t_r}{5.5} = 1.05 \text{ h}$$

$$\text{Tiempo de pico: } t_p = t_r + \frac{d_e}{2} = 6.28 \text{ h}$$

$$\text{Gasto de pico: } q_p = \frac{0.205A}{t_p} = 13.67 \text{ m}^3/\text{s}/\text{mm}$$

El hidrograma unitario correspondiente, asociado a una duración en exceso de una hora, se definió a partir de la forma del hidrograma unitario adimensional propuesto por el U.S. Soil Conservation Service, Figura 5-6.

A partir de las láminas de lluvia efectiva distribuidas espacialmente sobre las celdas donde se localiza el área de captación de la subcuenca, obtenidas por asimilación y por interpolación IDW, se definió un solo valor de lluvia efectiva media para cada día del tren de precipitaciones, valores que se concentran en la Tabla 5.2.

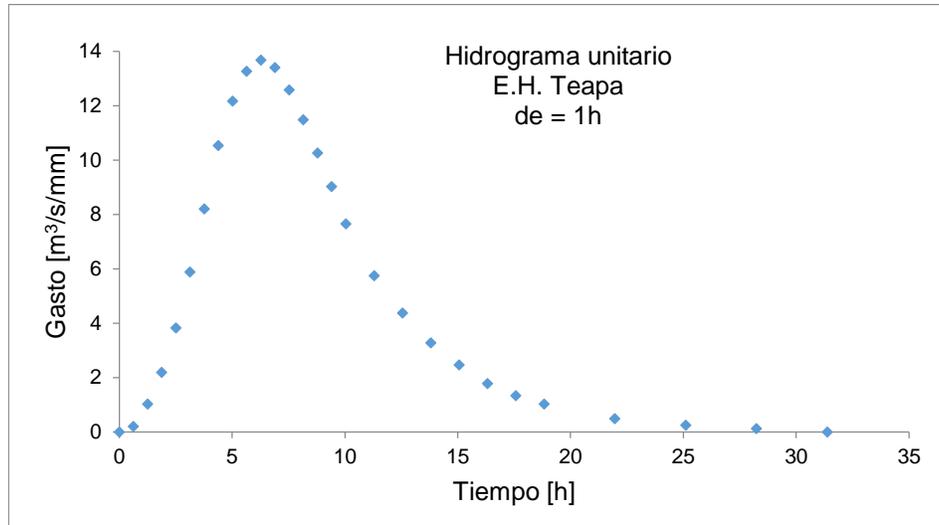


Figura 5-6: Hidrograma unitario asociado a una duración en exceso de una hora, correspondiente a la subcuenca de la E.H. Teapa.

Tabla 5.2: Láminas de lluvia efectiva media, subcuenca de la E.H. Teapa.

Día	19	20	21	22	23	24	25	26
Pe asimilada (mm)	7.94	15.50	13.01	44.06	32.44	5.40	0.00	0.00
Pe IDW (mm)	0.02	0.07	8.47	68.15	35.04	5.81	0.00	0.00

En consideración de la Figura 5-3 y la ubicación de la región de estudio, se tomaron en cuenta la cantidad de celdas de la malla a las que les corresponden factores de convectividad R de 0.45 y 0.65, por lo que en el cálculo de la intensidad de lluvia máxima de una hora se utilizó un factor de convectividad medio de 0.50. Asimismo, la distribución temporal de las láminas de los trenes de lluvia efectiva media establecidos se definió con los factores de reducción K indicados en la Tabla 5.1, para R igual a 0.50.

El cálculo de los escurrimientos en el sitio de la E.H. Teapa se realizó mediante la convolución del hidrograma unitario de la Figura 5-6 con los respectivos hietogramas de lluvia efectiva distribuida en intervalos de una hora. La Figura 5-7 muestra los escurrimientos correspondientes a las lluvias definidas por asimilación y por interpolación IDW, junto al comportamiento de la avenida considerada real.

De este ejemplo puede concluirse que existe una similitud importante en la magnitud y variación de los gastos obtenidos con un mismo esquema de lluvia (asimilado o IDW), analizado indistintamente con un modelo hidrodinámico o hidrológico, lo que aporta confiabilidad a los resultados. Adicionalmente, y como ocurrió en el análisis hidrodinámico, la menor sobreestimación de gastos se obtuvo al considerar las lluvias estimadas con el método de asimilación propuesto.

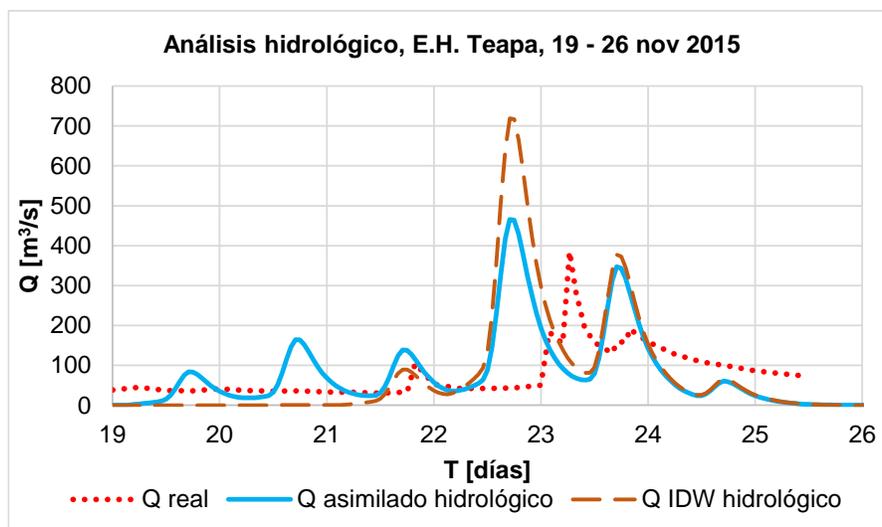


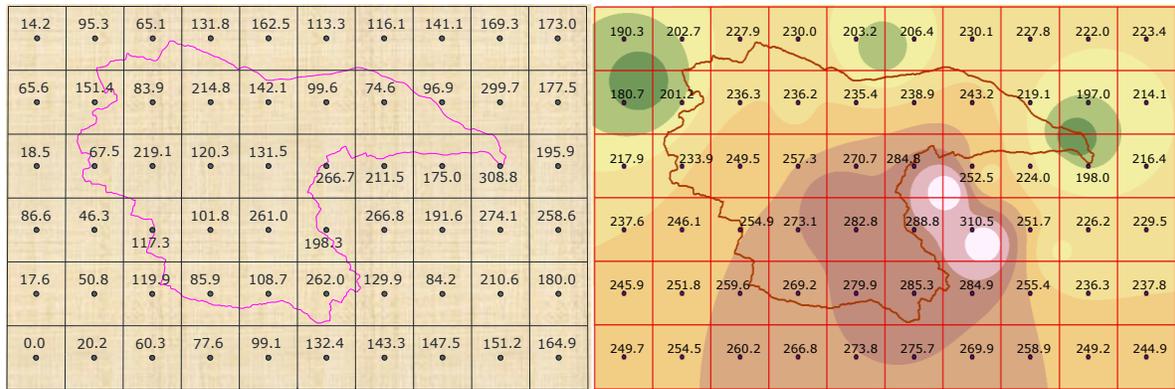
Figura 5-7: Escurrimiento superficial, real y simulado con análisis hidrológico, sitio de la E.H. Teapa.

### 5.2.2. Escurrimientos superficiales en la subcuenca de la E.H. Cadereyta

Anteriormente se realizó el análisis de lluvia diaria sobre la subcuenca de aportación al escurrimiento superficial en el sitio de la estación hidrométrica E.H. 24327 Cadereyta, estado de Nuevo León, en atención a las láminas históricas correspondientes a dos periodos distintos, el primero ocurrido entre el 8 y el 22 de septiembre de 2002 (tren de lluvias no relacionado con evento ciclónico), así como las lluvias precipitadas entre el 16 y el 30 de julio de 2005, asociadas al paso del huracán Emily. Como un ejemplo de la distribución espacial de las láminas diarias estimadas sobre las sesenta celdas que componen la malla definida para la caracterización de la subcuenca de estudio, en la Figura 5-8 se muestran por separado las láminas estimadas por asimilación y las calculadas mediante la interpolación IDW de los registros pluviométricos considerados en el proceso de asimilación, correspondientes al día 20 de julio de 2005.

Para realizar el análisis del escurrimiento superficial derivado de las precipitaciones sobre la subcuenca de la estación Cadereyta, primero se procedió al cálculo de la lámina de lluvia efectiva en cada celda de la malla, la cual se obtuvo por el método de los números de escurrimiento. Se consultó el conjunto de datos vectoriales de la edafología, escala 1:250 000 – 1:1 000 000 (1995), así como los correspondientes al uso de suelo y vegetación, escala 1:250 000 (Serie II, 2001), disponibles en el Portal de Geoinformación de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. La Figura 5-9 presenta la distribución de los diferentes tipos y usos de suelo y su vegetación sobre la zona de estudio, así como el número de escurrimiento N que le corresponde a cada celda de la malla.

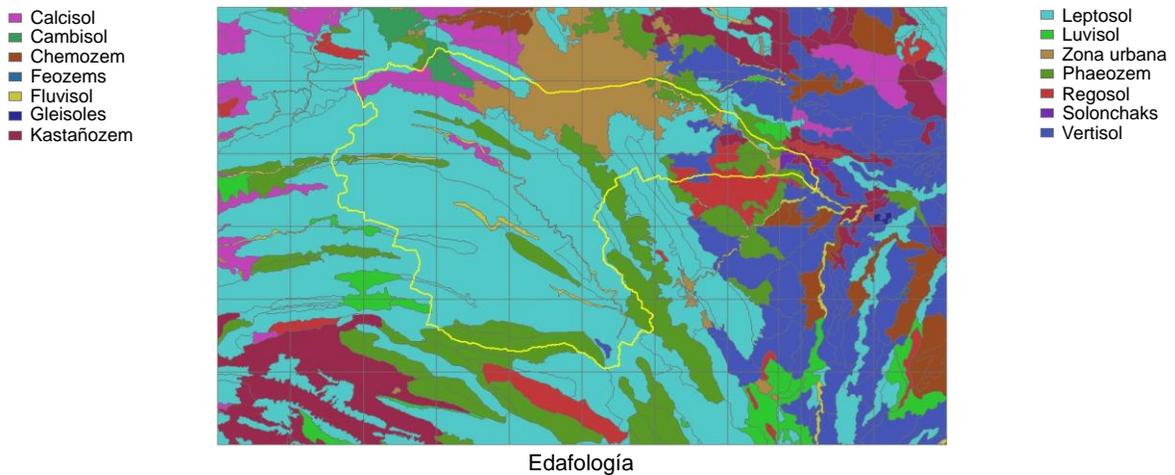
En conocimiento de la distribución espacial de la lluvia efectiva diaria, se procedió a establecer la distribución temporal de su precipitación, pero debido a que no se dispone de registros de la variación temporal de las tormentas sobre la región y fechas de interés, se tomó en cuenta la relación entre las precipitaciones de máxima intensidad de una hora de duración y las acumuladas en 24 horas.



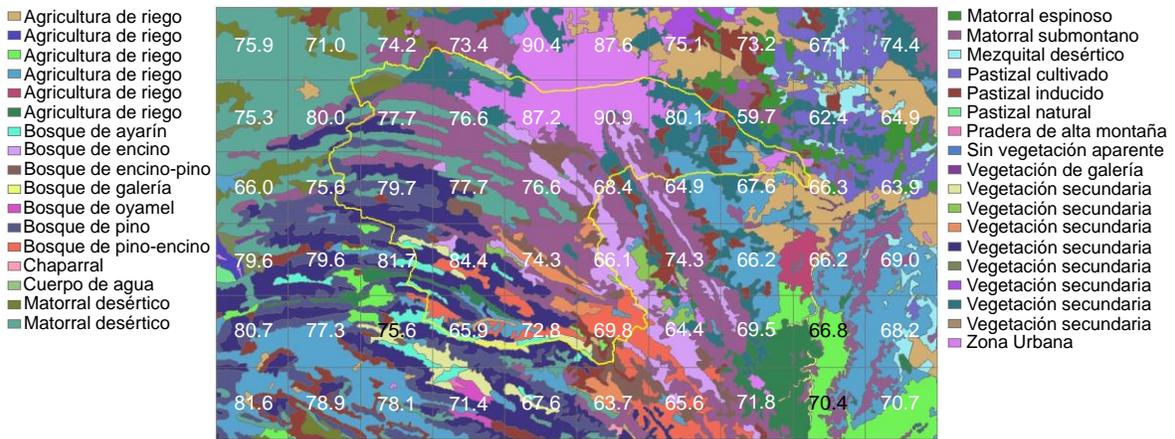
Estimación de lluvia diaria por asimilación

Estimación de lluvia diaria por interpolación IDW

Figura 5-8: Distribución de lluvia en mm, julio 20 de 2005, subcuena E.H. Cadereyta.



Edafología



Uso de suelo y vegetación, y número N

Figura 5-9: Distribución de tipos y usos de suelo y su vegetación, y números N correspondientes.

Dada la ubicación de la región de estudio y la distribución de factores de convectividad propuesta por Baeza (Figura 5-3), a todas las celdas de la malla les corresponde un factor R de 0.30. Asimismo, los valores de lluvia asociados a duraciones menores que 24 h se obtuvieron al multiplicar la precipitación correspondiente a 1 h por los factores de reducción K ( $K = P_d^{Tr} / P_1^{Tr}$ ) indicados en la Tabla 5.1.

A manera de ejemplo, la Figura 5-10 muestra la distribución temporal de la lluvia efectiva por asimilación en un acomodo en bloque alterno, que corresponde a la celda del cuarto renglón - tercera columna del arreglo matricial que representa la distribución espacial de la lluvia sobre la malla de estudio, lugar donde se localiza la sección del cauce principal más alejada de la estación Cadereyta, para los eventos de interés del año 2005.

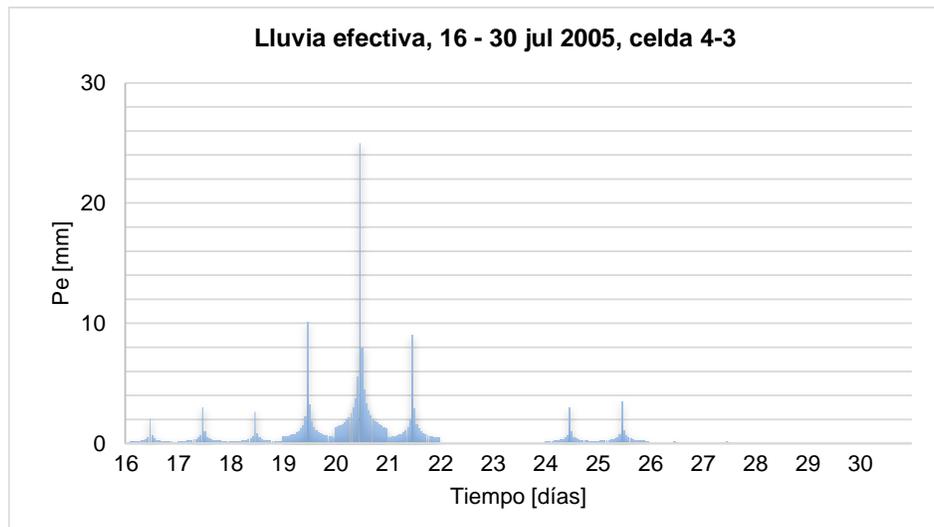


Figura 5-10: Hietograma de lluvia efectiva estimada por asimilación, celda 4-3, en intervalos de una hora y duración total de quince días.

En conocimiento de la lluvia efectiva estimada por asimilación y por interpolación IDW, distribuida espacial y temporalmente durante el periodo de precipitación analizado, se procedió a simular la transformación de la lluvia en escurrimiento directo hasta el sitio de la E.H. 24327 Cadereyta, utilizando el módulo hidrodinámico de la herramienta Iber. Para la simulación se utilizó la información topográfica digital generada por el INEGI, continuo de elevaciones mexicano 3.0 (celdas de 15 m de resolución), en particular los datos vectoriales de los estados de Coahuila y Nuevo León. Por su parte, para el análisis del flujo sobre la superficie general de la subcuenca y dentro del cauce, se tomó en cuenta un conjunto de valores del coeficiente de rugosidad de Manning [21]: agricultura 0.04, bosque 0.15, chaparral 0.03, cuerpo de agua 0.04, matorral 0.05, mezquital 0.03, pastizal 0.03, pradera 0.03, sin vegetación 0.025, vegetación 0.035, zona urbana 0.016 y cauce del río 0.025.

En la Figura 5-11 se aprecia la variación de los escurrimientos superficiales obtenidos mediante simulación hidrodinámica, para los esquemas de lluvia estimada por asimilación y por interpolación IDW de los eventos del 2005, así como el comportamiento simultáneo de los gastos registrados instantáneamente a lo largo del tiempo en el sitio de la estación Cadereyta, valores disponibles en el banco de datos BANDAS de la CONAGUA.

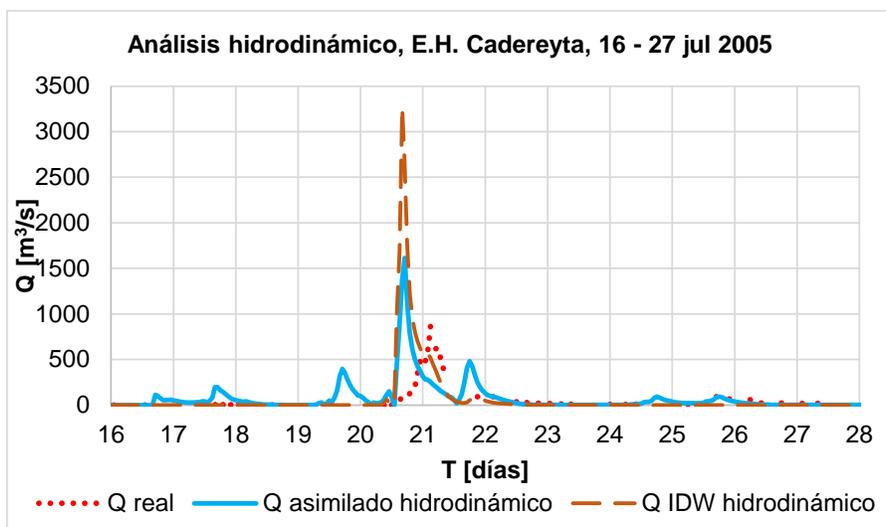


Figura 5-11: Escurrimiento superficial, real y simulado con análisis hidrodinámico, sitio de la E.H. Cadereyta.

Cabe señalar que en el comportamiento de los escurrimientos se tomó en cuenta la mitigación de la avenida debido al funcionamiento de la Presa Rompepicos construida para el año 2004, localizada en las coordenadas 25°33'23" N y 100°23'52" O. Los resultados mostrados representan la diferencia entre la avenida correspondiente a la captación de lluvia sobre la totalidad de la superficie de la subcuenca de estudio, que coincide con el sitio de la estación hidrométrica, y los escurrimientos correspondientes a la subcuenca de la Presa Rompepicos. Como se observa, al tomar en cuenta las precipitaciones estimadas con el método de asimilación propuesto, existe cierta coincidencia respecto al tiempo en que se presentó el gasto máximo de la avenida histórica, además de una mayor aproximación en la magnitud de dicho gasto, pues los escurrimientos derivados de láminas estimadas por interpolación IDW presentan una sobreestimación mayor.

A manera de complemento para la revisión de la calidad de las lluvias estimadas con el método de asimilación propuesto, se realizó el cálculo de la avenida al sitio de la E.H. Cadereyta mediante un procedimiento hidrológico de parámetros concentrados, con la intención de observar si bajo dicho nivel de análisis también se manifiesta un beneficio al comparar los escurrimientos derivados de la lluvia obtenida por asimilación respecto de la calculada con el método de interpolación IDW.

En consideración de la información topográfica digital del INEGI utilizada para este análisis, la subcuenca de interés tiene un área de 1 812.39 km<sup>2</sup>, así como una longitud y pendiente media de cauce principal de 165 614.12 m y 0.0073, respectivamente, por lo que el tiempo de concentración calculado con la expresión de Kirpich resultó de 22.50 horas. El tiempo y gasto de pico se obtuvieron considerando:

Tiempo de retraso:  $t_r = 13.50$  h

Duración en exceso:  $d_e = 2.45$  h

Tiempo de pico:  $t_p = 14.72$  h

Gasto de pico:  $q_p = 25.23$  m<sup>3</sup>/s/mm

El hidrograma unitario asociado a la duración en exceso se definió a partir de la forma del hidrograma unitario adimensional propuesto por el U.S. Soil Conservation Service (Figura 5-12).

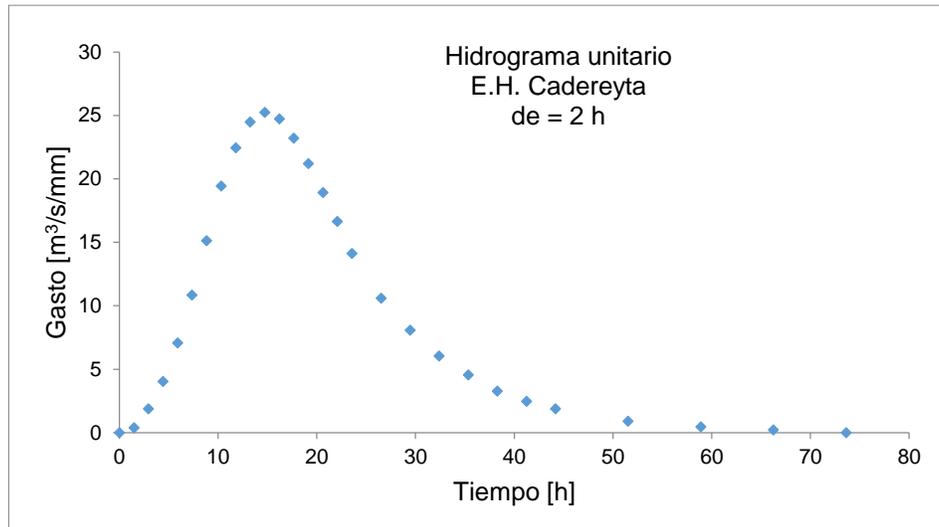


Figura 5-12: Hidrograma unitario asociado a una duración en exceso de dos horas, correspondiente a la subcuenca de la E.H. Cadereyta.

Para realizar el análisis hidrológico, de cada día del tren de precipitaciones se estableció un solo valor de lluvia efectiva media a considerarse sobre la totalidad de las celdas de captación, valores que se concentran en la Tabla 5.3.

Tabla 5.3: Láminas de lluvia efectiva media, subcuenca de la E.H. Cadereyta.

Día	16	17	18	19	20	21	22	23	24	25	26	27
Pe asimilada (mm)	7.86	4.88	0.37	29.32	95.68	15.63	0	0	5.44	7.52	0.21	0.20
Pe IDW (mm)	0.15	0.05	0	0.04	184.59	1.38	0	0	0	5.35	0	0

De acuerdo con la Figura 5-3 y la ubicación de la región de estudio, en el cálculo de la intensidad de lluvia máxima de una hora se utilizó un factor de convectividad de 0.30, mientras que la distribución temporal de las láminas de precipitación de los trenes de lluvia efectiva media se definió con los factores de reducción K indicados en la Tabla 5.1, para R igual a 0.30.

La obtención de los escurrimientos en el sitio de la E.H. Cadereyta se realizó mediante la convolución del hidrograma unitario de la Figura 5-12 con los hietogramas de lluvia efectiva distribuida en intervalos de dos horas. Cabe mencionar que en este análisis de los escurrimientos también se tomó en cuenta la mitigación de la avenida debido al funcionamiento de la Presa Rompepicos, y la Figura 5-13 muestra la variación de los gastos correspondientes a las lluvias definidas por asimilación y por interpolación IDW, junto al comportamiento de la avenida real registrada en la estación hidrométrica.

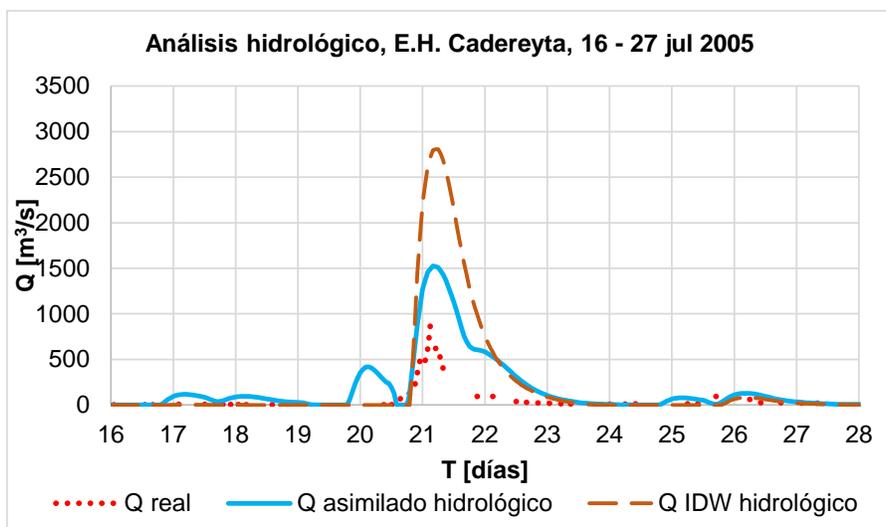


Figura 5-13: Escurrimiento superficial, real y simulado con análisis hidrológico, sitio de la E.H. Cadereyta.

A partir de estos resultados, donde se aprecia que existe similitud entre magnitudes y variaciones del gasto obtenidos para un mismo esquema de lluvia, analizado mediante un modelo hidrodinámico e hidrológico (figuras 5-11 y 5-13), se confirma que las diferencias respecto al registro histórico de gastos en la estación hidrométrica son consecuencia de la sobreestimación de las precipitaciones, aunque el resultado que representa una mejor aproximación al evento real se obtuvo al considerar las lluvias estimadas con el método de asimilación propuesto.

### 5.2.3. Escurrimientos superficiales en la subcuenca de la E.H. Las Perlas

Como un ejemplo adicional de la revisión de la calidad de las estimaciones de lluvia espacialmente distribuida y estimada con el método de asimilación propuesto, se tomaron en cuenta los resultados del análisis de lluvia diaria sobre la subcuenca de aportación al escurrimiento superficial en el sitio de la estación hidrométrica E.H. 29005 Las Perlas, estado de Veracruz, para lo cual se consideraron las láminas correspondientes a las precipitaciones ocurridas entre el 21 de septiembre y el 5 de octubre de 2010, relacionadas con el paso de la tormenta tropical Matthew. A manera de muestra, en la Figura 5-14 se presentan por separado las láminas estimadas por asimilación y las calculadas mediante la interpolación IDW de los registros pluviométricos considerados en el proceso de asimilación, correspondientes a las ciento setenta celdas que componen la malla definida para caracterizar la subcuenca de estudio, estimadas para el día 27 de septiembre de 2010.

Con la intención de realizar el análisis del escurrimiento superficial derivado de las precipitaciones sobre la subcuenca de la estación Las Perlas, también se procedió al cálculo de la lluvia efectiva en cada celda de la malla, por el método de los números de escurrimiento. Se consultó el conjunto de datos vectoriales de la edafología, escala 1:250 000 – 1:1 000 000 (1995), así como los correspondientes al uso de suelo y vegetación, escala 1:250 000 (Serie IV, 2009) disponibles en el Portal de Geoinformación de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad.

En la Figura 5-15 se muestra la distribución de los diferentes tipos y usos de suelo y su vegetación sobre la zona de estudio, así como el número de escurrimiento N que le corresponde a cada celda de la malla.

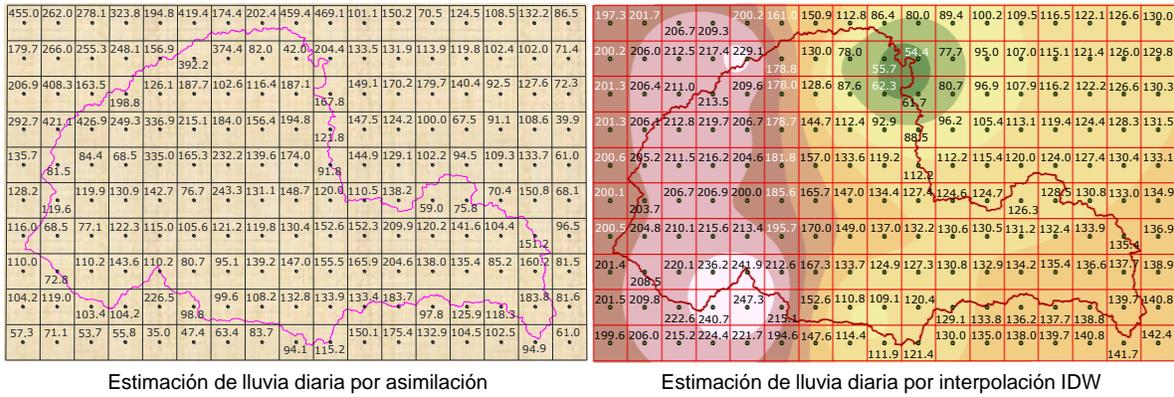


Figura 5-14: Distribución de lluvia en mm, septiembre 27 de 2010, subcuenca E.H. Las Perlas.

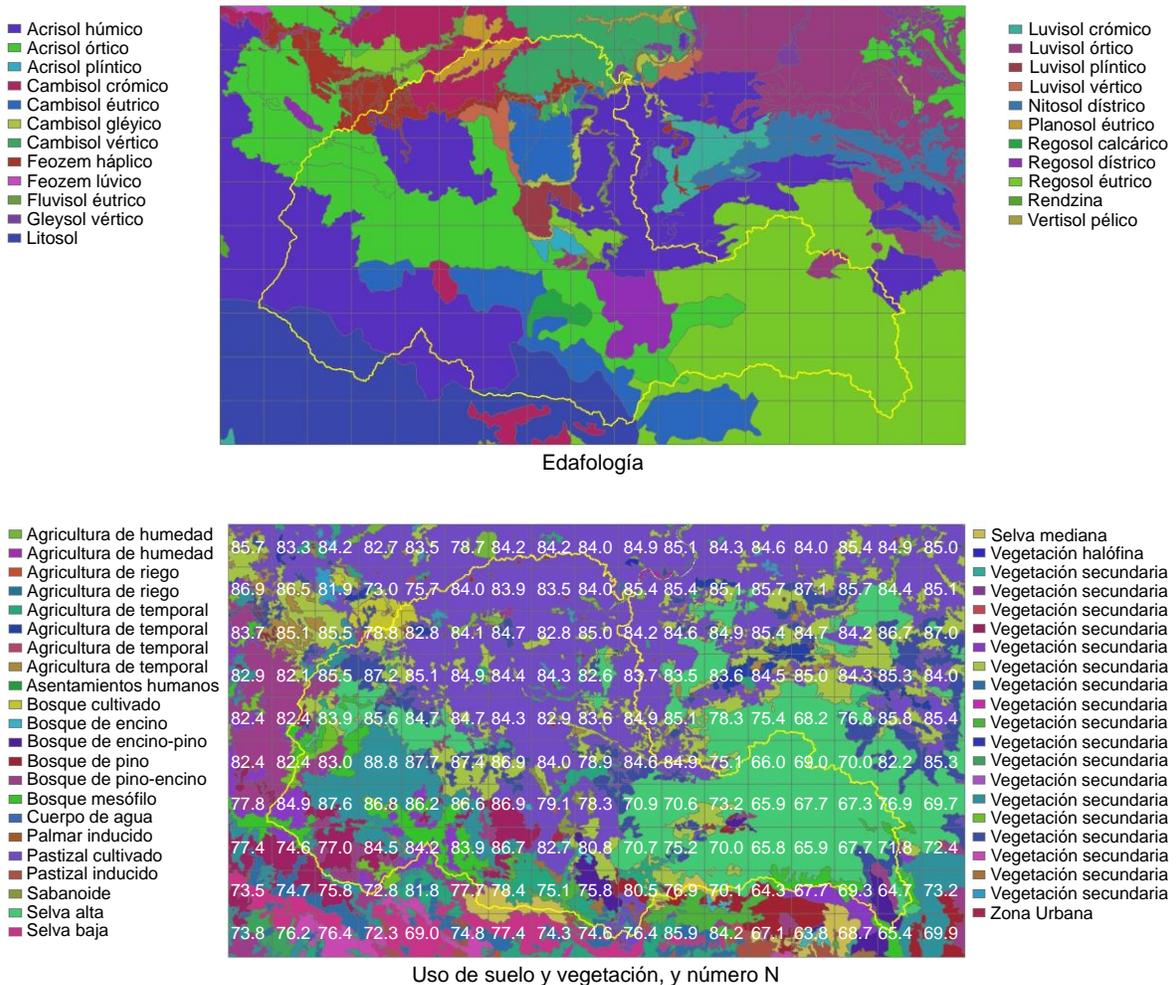


Figura 5-15: Distribución de tipos y usos de suelo y su vegetación, y números N correspondientes.

Una vez conocida la distribución espacial de la lluvia efectiva diaria, se procedió a definir la distribución temporal de su precipitación. Se consideró apropiado tomar en cuenta la relación existente entre las precipitaciones de máxima intensidad de una hora y las acumuladas en 24 horas.

De acuerdo con la distribución de factores de convectividad propuesta por Baeza (Figura 5-3), se identificó que a un total de 57 celdas de captación de la subcuenca les corresponde un coeficiente R de 0.30, a 29 celdas un coeficiente de 0.45, y a 3 celdas un coeficiente de 0.65. Posteriormente, las lluvias asociadas a duraciones menores que 24 h se obtuvieron al multiplicar las precipitaciones de 1 h de duración por los correspondientes factores de reducción K ( $K = P_d^{Tr} / P_1^{Tr}$ ) de la Tabla 5.1.

En la Figura 5-16 se presenta un ejemplo de la distribución temporal de la lluvia efectiva por asimilación en un acomodo en bloque alterno, que corresponde a la celda del décimo renglón – decimosexta columna del arreglo matricial que representa la distribución espacial de la lluvia en la malla de estudio, porque allí se localiza la sección del cauce principal más alejada de la estación Las Perlas.

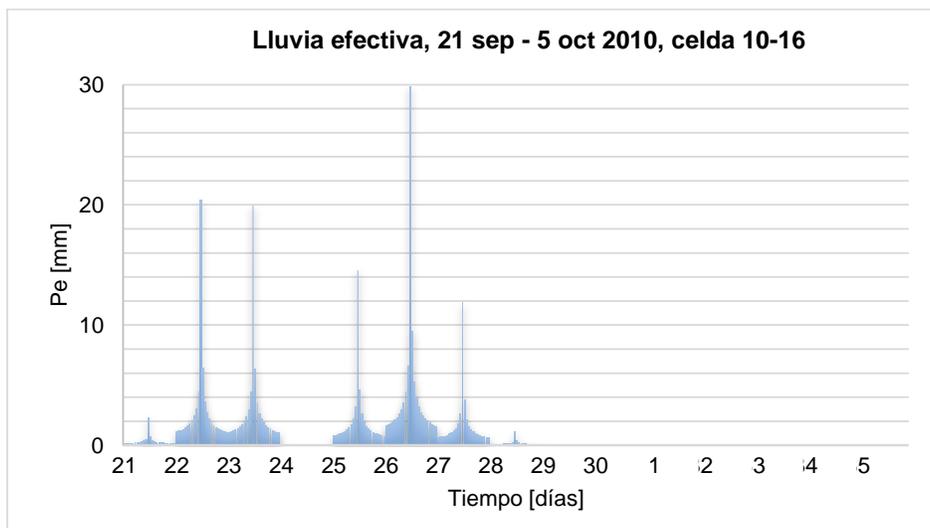


Figura 5-16: Hietograma de lluvia efectiva estimada por asimilación, celda 10-16, en intervalos de una hora y duración total de quince días.

En disposición de la lluvia efectiva estimada por asimilación y por interpolación IDW, distribuida espacial y temporalmente en el periodo del 21 de septiembre y el 5 de octubre de 2010, se inició el proceso de simulación para la transformación de la lluvia en escurrimiento directo hasta el sitio de la E.H. 29005 Las Perlas, con el empleo del módulo hidrodinámico de la herramienta Iber.

En el proceso de simulación de escurrimientos se utilizó la información topográfica digital generada por el INEGI, continuo de elevaciones mexicano 3.0 (celdas de 15 m de resolución), en particular los datos vectoriales de los estados de Oaxaca y Veracruz. En lo que se refiere al análisis del flujo sobre la superficie general de la subcuenca y dentro del cauce, se tomó en cuenta un conjunto de valores del coeficiente de rugosidad de Manning [21]: agricultura 0.04, asentamientos humanos 0.02, bosque 0.15, cuerpo de agua 0.04, palmar 0.03, pastizal 0.03, sabanoide 0.06, selva 0.15, vegetación 0.035, zona urbana 0.016 y cauce del río 0.025.

La Figura 5-17 presenta la variación de los escurrimientos superficiales obtenidos mediante simulación hidrodinámica, para los esquemas de lluvia estimada por asimilación y por interpolación IDW de los eventos del 2010, además del comportamiento en simultáneo de los gastos registrados instantáneamente a lo largo del tiempo en el sitio de la estación Las Perlas, datos recopilados del banco de datos BANDAS de la CONAGUA.

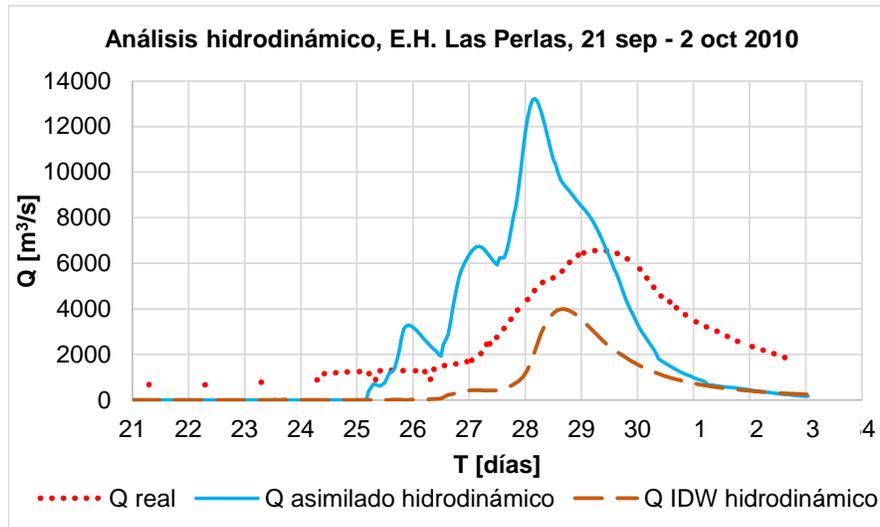


Figura 5-17: Escurrimiento superficial, real y simulado con análisis hidrodinámico, sitio de la E.H. Las Perlas.

Como puede observarse, existe cierta coincidencia en el tiempo de inicio y finalización de los escurrimientos simulados respecto de la avenida histórica, pero en la magnitud de los gastos estimados se manifiestan diferencias más importantes, debido especialmente a una falta de aproximación de las láminas de lluvia estimadas respecto del campo de lluvia real precipitado sobre la totalidad de la superficie analizada, produciéndose una subestimación de escurrimientos al considerar la precipitación calculada por interpolación IDW, así como una sobreestimación al considerarse las láminas obtenidas por asimilación.

Para concluir con la revisión de la calidad de las lluvias estimadas con el método de asimilación propuesto, también se practicó un procedimiento hidrológico de parámetros concentrados para la obtención de la avenida al sitio de la estación Las Perlas, de manera que pueda observarse si bajo dicho nivel de análisis hay evidencia del beneficio de aplicar los valores estimados con la metodología de asimilación de lluvias respecto de los calculados por interpolación IDW.

A partir de la información topográfica digital del INEGI utilizada para el análisis, se conoce que la subcuenca de estudio tiene un área de 8 993.11 km<sup>2</sup>, además de una longitud y pendiente media de cauce principal de 257 416.10 m y 0.00132, respectivamente, por lo que el tiempo de concentración calculado con la expresión de Kirpich es de 61.23 horas. Para el tiempo y gasto de pico se consideró:

Tiempo de retraso:  $t_r = 36.74$  h

Duración en exceso:  $d_e = 6.68$  h

Tiempo de pico:  $t_p = 40.08$  h

Gasto de pico:  $q_p = 46.00$  m<sup>3</sup>/s/mm

La forma del hidrograma unitario asociado a la duración en exceso se definió en apego a la propuesta por el U.S. Soil Conservation Service (Figura 5-18).

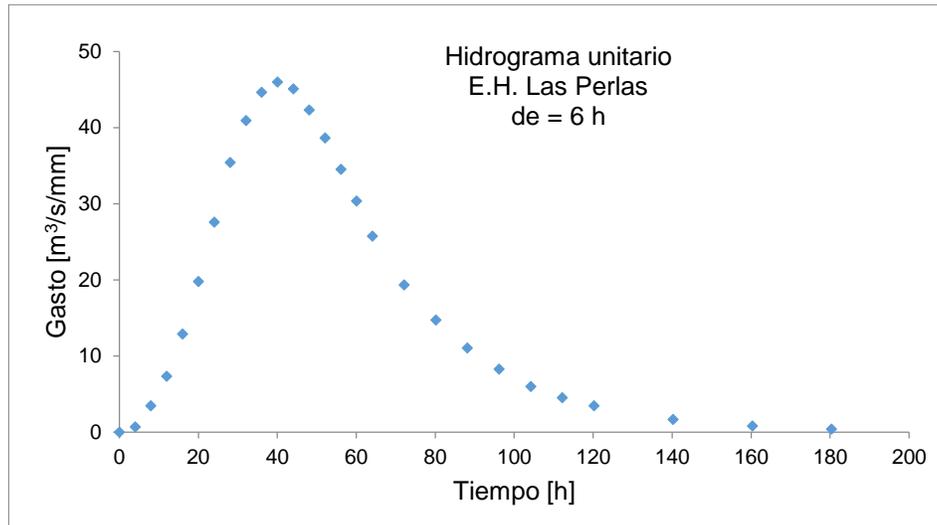


Figura 5-18: Hidrograma unitario asociado a una duración en exceso de seis horas, correspondiente a la subcuenca de la E.H. Las Perlas.

El análisis hidrológico se realizó considerando para cada día del tren de precipitaciones un solo valor de lluvia efectiva media sobre todas las celdas que definen el área de captación de la subcuenca, valores que se concentran en la Tabla 5.4.

Tabla 5.4: Láminas de lluvia efectiva media, subcuenca de la E.H. Las Perlas.

Día	21	22	23	24	25	26	27	28	29	30	1	2
Pe asimilada (mm)	6.02	28.74	37.23	23.49	86.18	140.27	116.70	53.20	0	0	0	0
Pe IDW (mm)	4.09	4.67	14.71	9.30	23.95	70.97	117.97	3.19	0	0	0	0

En función del número de celdas de la malla de estudio a las que les corresponden factores de convectividad R de 0.30, 0.45 y 0.65, dada la distribución de regiones de la Figura 5-3, para el cálculo de la intensidad de lluvia máxima de una hora se utilizó un factor de convectividad medio de 0.40. Asimismo, la distribución temporal de las láminas de precipitación de los trenes de lluvia efectiva media se definió con los factores de reducción K indicados en la Tabla 5.1, para R igual a 0.40.

Finalmente, los escurrimientos en el sitio de la E.H. Las Perlas se obtuvieron mediante la convolución del hidrograma unitario de la Figura 5-18 con los hietogramas de lluvia efectiva distribuida en intervalos de seis horas, de manera que la Figura 5-19 presenta la variación de los escurrimientos derivados de las lluvias definidas por asimilación y por interpolación IDW, respecto al comportamiento de la avenida real.

En atención a los resultados obtenidos, puede concluirse que las diferencias más importantes entre el registro histórico de gastos instantáneos en la estación hidrométrica y los valores obtenidos por simulación de los procesos de lluvia-escorrimento, se deben especialmente a la sobreestimación de las precipitaciones, de manera que el resultado que representa una mejor aproximación se obtuvo al considerar las lluvias estimadas con el método de interpolación IDW, porque la magnitud del área analizada en consideración de una cantidad reducida de estaciones pluviométricas localizadas principalmente en el perímetro de la subcuenca de captación, fueron condiciones adversas que provocaron una falta de aproximación en las estimaciones de lluvia en un mayor número de celdas, al utilizarse el procedimiento de asimilación propuesto.

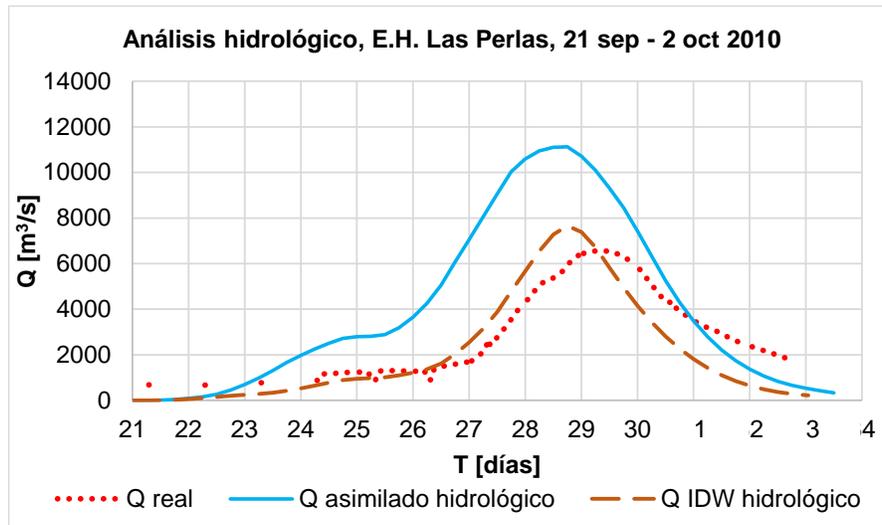


Figura 5-19: E scorrimento superficial, real y simulado con análisis hidrológico, sitio de la E.H. Las Perlas.

## 6. Conclusiones generales

De acuerdo con el registro histórico de las condiciones climatológicas que son propias de la República Mexicana, existe evidencia del comportamiento variable de la precipitación pluvial, tanto de su distribución espacial como temporal, asociado principalmente a la ubicación y características geográficas de su territorio.

Por otra parte, se sabe que la acumulación de lluvia es más grande en la región sur, donde los eventos de naturaleza convectiva y ciclónica son característicos (con actividad ciclónica más intensa en la costa del Pacífico), mientras que las condiciones en la zona norte son más favorables para un ambiente seco, y en la zona de la Península de Baja California la generación de lluvia está relacionada especialmente con eventos ciclónicos.

Dado que la temporada de lluvias inicia en mayo y finaliza en noviembre, considerando que la temporada de huracanes inicia oficialmente el 15 de mayo en el Océano Pacífico Oriental, y el 1° de junio en el Pacífico Central y el Océano Atlántico, para finalizar el 30 de noviembre en ambos océanos, se ha observado que más de dos terceras partes de la precipitación normal ocurre durante los meses de junio y septiembre, pero durante los meses de invierno (ajenos a la temporada de lluvias) la combinación de los frentes fríos con el ingreso de humedad de los litorales también puede suscitar la caída de precipitación.

La disponibilidad de registros pluviométricos sobre el territorio mexicano tradicionalmente ha posibilitado el análisis estadístico de la precipitación puntual, al tomar en cuenta que esta información en general se considera confiable, aunque dichos valores solo representan las condiciones de la región más próxima a los puntos de medición.

La lluvia precipitada sobre la superficie nacional (alrededor de 1 964 375 km<sup>2</sup>) se ha registrado principalmente en intervalos preestablecidos de 24 horas, utilizando un conjunto de aproximadamente 5 500 estaciones climatológicas, con información originada desde principios del siglo veinte, pero con el transcurso del tiempo gran cantidad de estaciones han dejado de operar o ya no existen, por lo que al 31 de diciembre de 2017 solo 3 079 estaciones se encontraban operando bajo una distribución espacial irregular y dispersa, con menor densidad de mediciones en la parte norte, noroeste y sureste del país. Es en estas condiciones de densidad y distribución que la red de observaciones resulta insuficiente para captar la estructura espacial de las tormentas, dado un distanciamiento promedio entre estaciones de 25 km.

Debido a que existen antecedentes del empleo de imágenes de reflectividad obtenidas con radares meteorológicos como elementos para la estimación de láminas de lluvia, con reportes de resultados relativamente satisfactorios en cuanto a la recuperación del comportamiento espacial de una tormenta histórica, para el cumplimiento de los objetivos de este trabajo de tesis se consideró el empleo de estos datos de percepción remota. Mediante una solicitud a la ventanilla única del

SMN, institución encargada de operar los instrumentos y resguardar la información de los 13 radares que conforman la red nacional, de cada fecha de interés fue posible recopilar hasta 96 imágenes PPI, con resolución temporal de 15 minutos y resolución espacial de 833 m.

Para efectuar un análisis robusto de las imágenes de reflectividad disponibles, de manera que pudieran obtenerse valores de lluvia distribuida de mejor calidad, condición importante para una valoración más objetiva del empleo de estos datos respecto de los calculados con otros procedimientos, la estimación de lluvia de radar se realizó siguiendo la metodología recomendada por Vilchis *et al.* (2011), quienes propusieron un conjunto de pasos ordenados para el manejo de las 96 imágenes consecutivas que representan los registros de un día, atendiendo aspectos de georreferenciación y conversión apropiada de la información numérica de las imágenes de reflectividad en láminas de lluvia acumulada.

Adicionalmente, como la disponibilidad de imágenes de radares meteorológicos que involucren eventos históricos ocurridos en cualquier región de la superficie nacional suele ser muy limitada, también se realizó la recopilación de observaciones meteorológicas de sensores satelitales, como complementos de la estimación de lluvias obtenidas por percepción remota.

En la actualidad pueden conocerse estimaciones generadas mediante diferentes algoritmos que combinan información registrada por diversos tipos de sensores, pero también existen estudios que señalan la obtención de resultados satisfactorios al utilizar los productos de la misión GPM con respecto a los productos CMORPH y PERSIANN, por ejemplo.

Para los propósitos de este trabajo se consideraron las lluvias estimadas con el algoritmo IMERG de la GPM, con resoluciones de  $0.1^\circ$  por  $0.1^\circ$ , en particular el producto denominado IMERG Final, el cual está disponible a 3.5 meses de ocurrido cada evento. Los archivos tipo ráster para las fechas y regiones de interés se obtuvieron del sistema de datos en línea GIOVANNI del Centro de Servicios de Información y Datos de Ciencias de la Tierra Goddard de la NASA (GES DISC).

Como el interés principal de este trabajo de tesis es proponer una metodología que permita mejorar la estimación del comportamiento espacial de un campo de lluvia diaria histórica en México, para complementar el déficit de registros pluviométricos que se consideran confiables pero solo son representativos de zonas puntuales, se optó por la combinación del registro de las redes de pluviómetros disponibles con los productos de percepción remota de sistemas satelitales y radares meteorológicos, en la consideración de que dichos productos estiman de manera apropiada la estructura espacial de los eventos de tormenta.

En el entendido de que el método empleado para la estimación del estado de la atmósfera es un factor principal en la exactitud que pueda alcanzarse a partir del conocimiento de un conjunto limitado de observaciones de la realidad, se consideraron experiencias compartidas en trabajos de investigación orientados a definir el estado de un sistema como una representación discreta de valores localizados sobre las celdas de una malla regular.

La asimilación de datos es la metodología para la estimación del estado (diagnóstico y pronóstico) de un sistema variable en el tiempo, a partir de datos observados y campos antecedentes, así como de un modelo basado en leyes físicas conocidas del sistema, por lo que desde un punto de vista estadístico el resultado corresponde con el estado esperado del sistema analizado dada la información disponible, de manera que puede evaluarse la magnitud de los errores asociados. Es así como la asimilación se ha convertido en una herramienta útil en la investigación meteorológica

para la observación y el registro permanente del estado del tiempo, con orientación a la predicción numérica del clima.

En general existen dos métodos para afrontar el problema de asimilación. El primer método, de asimilación secuencial (por filtros), está basado en un planteamiento recursivo, porque el procesamiento de datos se realiza en cada ocasión que hay información disponible, es decir, se analiza la evolución de un modelo para definir el estado de un sistema entre tiempos consecutivos, procurando la estimación de una serie de estados condicionados a las observaciones pasadas y presentes. El segundo método, de asimilación variacional (por suavizado), está basado en procedimientos para la determinación explícita de los campos al minimizar una función objetivo predefinida que describe la relación entre las variables observadas y simuladas, por lo que se requiere de algoritmos eficientes de aproximación para el cálculo del gradiente de la función con respecto a sus argumentos, siendo ésta una razón que no ha favorecido a su implementación.

De acuerdo con la bibliografía, existen diversos esquemas de asimilación de datos de tipo secuencial, los que se diferencian por el nivel de detalle del análisis requerido. Para un problema de diagnóstico, la aplicación de los esquemas de correcciones sucesivas resulta práctica porque son procedimientos iterativos donde el análisis no se realiza de manera óptima, pues la intención es suavizar las observaciones contenidas dentro de un radio de influencia en función de la distancia que existe entre ellas.

Las estimaciones se obtienen por una combinación de correcciones calculadas para cada celda del espacio de análisis, las cuales son proporcionales a las diferencias entre lo observado y los valores que se estimaron preliminarmente. La ponderación de las correcciones se realiza empíricamente al considerar que las observaciones más cercanas a una celda tienen más importancia.

Respecto a la estimación de los campos de precipitación diaria como objetivo principal de este trabajo de tesis, se consideró conveniente aprovechar el concepto de asimilación de datos relativo a un esquema secuencial hasta la etapa de diagnóstico, para determinar la distribución bidimensional de la lluvia por la combinación de diferentes bases de datos. Se optó por el empleo del esquema de correcciones sucesivas que considera la función Kernel Gaussiana propuesta por Barnes, por estar recomendado como modelo simple de aplicación operacional para el análisis de tormentas, sugerido para regiones de gran escala, con sensibilidad reducida a errores, que proporciona resultados de calidad apropiada si la densidad de observaciones no es tan baja. Para agilizar los cálculos correspondientes a la estimación de los campos de precipitación, como parte de las actividades desarrolladas en esta tesis se construyó una herramienta basada en el método de Barnes, utilizando el lenguaje de programación propio del sistema de cómputo numérico MATLAB R2017a.

Para la ejecución del programa se requieren: conjunto de observaciones de precipitación acumulada diaria (registros pluviométricos y estimaciones de satélite y radar); dimensiones del área de estudio, en función del tamaño de las celdas cuadradas que definen la malla regular que representa la región de interés; parámetro de potencia que controla la caída del peso respecto a la distancia en el método de interpolación IDW; y el parámetro de convergencia requerido en el segundo paso del procedimiento de Barnes, relacionado con el grado de detalle o nivel de convergencia del campo interpolado en relación con el campo de observaciones.

En la evaluación de la calidad de las estimaciones de lluvia de cada esquema de asimilación se utilizó un procedimiento de validación cruzada, al considerarse que la magnitud de los errores está relacionada con el grado de suavización alcanzado a consecuencia de la cantidad de puntos utilizados en el proceso de interpolación. Entonces, la evaluación del campo asimilado se realizó en la ubicación de un conjunto de observaciones inicialmente removidas para fines de comparación, tomando en cuenta algunos índices estadísticos útiles para la medición de la correspondencia entre eventos observados y estimados. Se procuró la remoción de hasta un 10% de los datos pluviométricos disponibles, de preferencia aquellos localizados cerca de los centroides de las celdas con mayor cantidad de información, para evitar una modificación notable de la densidad y distribución original de las observaciones.

La comparación de las lluvias asimiladas respecto de los valores obtenidos de manera independiente con cada base de datos analizada se realizó considerando las láminas correspondientes a la ubicación del centroide de las celdas de validación. Lo anterior se debe a que la asimilación de lluvia se realizó sobre los centroides de cada celda, y entonces la evaluación debe practicarse considerando que el dato de validación es representativo de la superficie total de cada celda, lo que incluye al centroide de la celda a la que pertenece.

Con la intención de que al hacer uso del procedimiento de asimilación se obtenga el campo de lluvia más aproximado a la realidad del evento histórico analizado, se efectuó la puesta a punto del modelo con la asignación de valores a los parámetros involucrados en el proceso, es decir, se evaluó la respuesta del modelo a través de la modificación ordenada de sus parámetros, dentro de un rango de valores convenientes, a partir de un conjunto de entradas conocidas, para la obtención de resultados propios de un conjunto de salidas también conocidas.

Para establecer las características de ciertos ejercicios de asimilación que permitieran una evaluación de la calidad de los resultados obtenidos con la aplicación del modelo propuesto, se consideraron algunos aspectos relacionados con la naturaleza de los datos y el manejo de la información utilizada en trabajos de investigación desarrollados con anterioridad, en los cuales es común el análisis de láminas acumuladas durante varios días o meses consecutivos, sobre regiones que en amplitud han superado naciones enteras, por lo que suelen emplearse datos satelitales de resolución espacial de  $0.25^\circ$  por  $0.25^\circ$ .

En el entendido de que el cálculo de lluvia por percepción remota es mejor en escalas temporales más grandes, en este trabajo se optó por el análisis de intervalos de lluvia de menor duración, estimando lluvias diarias o eventos de días consecutivos asociados directamente al desarrollo de escurrimientos superficiales, utilizando información satelital de mejor resolución, de  $0.1^\circ$  por  $0.1^\circ$ .

En la identificación de eventos de tormenta históricos sobre la república mexicana para la aplicación de la metodología se consideró: primero, analizar una región monitoreada extensamente mediante una red de registros pluviométricos densa; después, el análisis de regiones de diferentes dimensiones, con disponibilidad pluviométrica muy limitada, pero con información hidrométrica en la zona baja de la captación de lluvia.

En primer lugar, se optó por el análisis de tormentas registradas sobre la superficie de la Ciudad de México (CDMX), para lo cual se revisó el orden de magnitud de las lluvias medias máximas anuales registradas en la región, que resultó de 47 mm. Se analizaron cinco eventos del año 2008, por considerarse representativos dado el tamaño de algunas láminas involucradas y su distribución

espacial, y se procedió a la evaluación de distintos esquemas: integración de la totalidad de la información disponible; combinación de datos con redes de pluviómetros de densidad reducida, revisando la influencia de la distribución espacial de los pluviómetros; asimilación que considera solo una base de datos de percepción remota a la vez; y la asimilación de los datos de percepción remota de satélite y radar.

Cabe señalar que, con la construcción de mapas de la distribución espacial de la variable mediante una interpolación simple de las observaciones pluviométricas, fue posible identificar anticipadamente las zonas de acumulaciones máximas de lluvia, mismas que también pudieron visualizarse con el manejo de los productos de percepción remota considerados, lo que justificó su empleo en el proceso de asimilación.

Con un tamaño de celda para la asimilación que coincide con la resolución de los datos satelitales, se realizó el análisis de eventos de naturaleza principalmente convectiva en la CDMX.

En general, con la combinación de las tres bases de datos recomendadas, utilizando la totalidad de los registros, se obtuvieron campos de lluvia adecuados pues, aunque los errores *ME* indican que la combinación de lluvia presentó una ligera subestimación, también se observó un mejor desempeño y exactitud respecto al cálculo del *MAE* y la *RMSE*. Asimismo, con el cálculo de coeficientes *NSE* próximos a la unidad, se interpretó que los ajustes son de buena calidad, con coeficientes *CC* que advierten sobre una asociación lineal importante entre la lluvia integrada y los datos de validación. Respecto a los datos de percepción remota utilizados, se observó una sobreestimación de las lluvias derivadas de datos satelitales, así como una subestimación de las lluvias calculadas con imágenes de radar, aunque el orden de magnitud de los errores que les corresponden es similar.

En vista de que al emplear la totalidad de los registros pluviométricos se obtuvieron estimaciones ampliamente satisfactorias, con una red original de datos de hasta 84 estaciones, ésta se consideró muy vasta porque el espaciamiento promedio entre sus registros es de 7.0 km (45 km<sup>2</sup> por pluviómetro), entonces también resultó factible la evaluación de ejemplos con densidades menores: alta, media y baja, correspondientes a distanciamientos entre registros de 8.1, 10.0 y 13.9 km (respectivamente 60, 91 y 182 km<sup>2</sup> por pluviómetro).

Con el análisis de redes de diferentes densidades y distribuciones espaciales de pluviómetros relativamente uniformes, se observó consistencia en la capacidad del modelo de asimilación, al verificarse que la combinación de las tres bases de datos genera aproximaciones satisfactorias independientemente de la configuración de una red, especialmente si la densidad es alta o media. En cambio, la calidad de las estimaciones se reduce si la densidad es baja, con errores comparables a los obtenidos por interpolación IDW de los registros pluviométricos, aunque dichos resultados de asimilación son mejores que las estimaciones individuales del satélite o radar.

Como complemento, se verificó la conveniencia de procurar la aplicación del procedimiento de asimilación, aun cuando solamente se disponga de una base de datos adicional a la de registros pluviométricos, ya sean productos de satélite o de radar.

De estos ejemplos de asimilación con dos bases de datos se concluyó que los resultados son útiles por su aproximación respecto a la mejor estimación obtenida al combinar las tres bases de datos recomendadas. Así, los errores por asimilación de lluvias pluviométricas con datos de satélite o radar son semejantes en magnitud, observándose respectivamente errores de sobreestimación y

subestimación, y por ello, independientemente de la tecnología de percepción remota utilizada, pueden obtenerse resultados de asimilación de calidad similar, con resultados adecuados respecto a interpolaciones IDW de lluvias registradas con una red de pluviómetros de densidad media. Como detalle adicional, cabe enfatizar que la asimilación de los dos productos de percepción remota, sin información pluviométrica, genera estimaciones de baja calidad.

Adicionalmente, la metodología se aplicó sobre otras tres regiones donde la disponibilidad de datos fuera más escasa y su distribución espacial menos uniforme, para la asimilación de registros pluviométricos y datos de satélite, en el entendido de las dificultades que implican la recopilación por escasez y el manejo de registros de radar. Los resultados tanto para lluvias diarias como para láminas acumuladas durante varios días mostraron que el incremento del error de estimación de las lluvias asimiladas está relacionado principalmente con la falta de exactitud de las lluvias calculadas como productos de satélite, lo que es más recurrente en los campos de lluvia de magnitud reducida.

Por ejemplo, en las subcuencas de las estaciones hidrométricas de Teapa (Tabasco) y Cadereyta (Nuevo León), con distanciamiento promedio entre observaciones de 17.5 y 22.0 km (respectivamente 319 y 478 km<sup>2</sup> por pluviómetro), se observó que las estimaciones por interpolación IDW de datos pluviométricos pueden ser mejores para analizar días de precipitación reducida, si esta circunstancia contribuye a incrementar la inexactitud de las estimaciones con satélite. Al contrario, en los días en que la acumulación de lluvia es más significativa, las estimaciones por asimilación de las dos bases de datos son una mejor aproximación que las obtenidas con las bases individuales.

En el análisis de la subcuenca de la estación hidrométrica Las Perlas (Veracruz), con distanciamiento promedio entre observaciones de 45.8 km (2 170 km<sup>2</sup> por pluviómetro), se observaron órdenes de magnitud de lluvia semejantes entre los registros pluviométricos y las estimaciones con satélite, es decir, la falta de exactitud de los productos satelitales fue menos significativa, inclusive en días de precipitación reducida. Por esta razón se concluyó que en caso de que los datos de satélite guarden un orden de magnitud similar a los registros pluviométricos disponibles, la aproximación de la lluvia por asimilación mostrará invariablemente un mejor desempeño respecto a considerar los datos de manera independiente.

Debido a que el análisis de precipitación en condiciones de escasez de información se realizó sobre áreas de captación en cuyo punto de salida existen registros hidrométricos, fue posible complementar la revisión de la calidad de la lluvia distribuida espacialmente al considerarla como el insumo principal de un análisis de escurrimiento superficial por lluvia en cuenca propia, para su posterior validación respecto a los gastos registrados. Como elementos de comparación se consideró el análisis de los escurrimientos generados a partir de las lluvias calculadas por interpolación IDW.

El módulo hidrodinámico del programa Iber se utilizó como herramienta para obtener el hidrograma a la salida de las regiones de estudio, porque existen antecedentes satisfactorios del empleo de este programa. Como el volumen del escurrimiento superficial es función de la lámina de lluvia efectiva, y la aplicación apropiada de los modelos de infiltración requiere de información que generalmente no está disponible sobre cualquier región de México, la estimación de la lluvia efectiva en cada celda de las mallas estudiadas se obtuvo con el método de los números de escurrimiento.

Las condiciones de humedad antecedente como resultado de la precipitación consecutiva de varias láminas de lluvia diaria (tren de tormentas) se consideraron mediante el concepto del índice de precipitación antecedente, utilizando una modificación al método tradicional de cálculo de la lluvia efectiva, propuesta por Fuentes *et al.* 2014.

Al no disponer de registros de la variación temporal de la lluvia sobre la región y fechas de interés, para definir la distribución temporal de las láminas diarias de lluvia efectiva distribuidas espacialmente, se consideró la relación entre las precipitaciones de máxima intensidad de una hora de duración con las acumulaciones en 24 horas, utilizando los factores de convectividad para la República Mexicana propuestos por Baeza en 2007, así como los valores asociados a otras duraciones menores que 24 h, obtenidos por Luna y Domínguez en 2013, para establecer hietogramas de lluvia efectiva en un acomodo de bloque alterno, relacionados con eventos de condiciones más adversas.

Del análisis de escurrimientos al sitio de la E.H. Teapa al emplear las lluvias estimadas con el método de asimilación propuesto, se observó una coincidencia notable respecto a los tiempos de inicio y finalización de la avenida real, así como mayor semejanza en la magnitud del gasto máximo, pues los escurrimientos relacionados con las láminas definidas por interpolación IDW mostraron una sobreestimación más grande. Caso similar resultó el análisis de escurrimientos hasta la E.H. Cadereyta, donde se observó cierta coincidencia en el cálculo de la magnitud y el tiempo asociado al gasto máximo de la avenida histórica.

Cabe señalar que adicionalmente se realizó el cálculo de las avenidas sobre los puntos de registro hidrométrico a través de un procedimiento hidrológico de parámetros concentrados, observándose también un beneficio al utilizar los campos de lluvia generados por asimilación.

De estos ejemplos se concluyó que existe similitud en la magnitud y variación de los gastos obtenidos con un mismo esquema de lluvia (asimilado o IDW), analizado indistintamente con un modelo hidrodinámico o hidrológico, lo que aportó confiabilidad a los resultados. Así, se confirmó que las diferencias respecto a los registros hidrométricos son consecuencia de la sobreestimación de las precipitaciones, aunque la mejor aproximación al evento real se obtuvo con las lluvias definidas con el método de asimilación propuesto.

Para el caso de los escurrimientos correspondientes a la E.H. Las Perlas, se apreció cierta coincidencia en el tiempo de inicio y finalización de los escurrimientos simulados respecto de los históricos, pero se observaron diferencias importantes en la magnitud de los gastos. Los escurrimientos calculados a partir de las lluvias estimadas por asimilación resultaron de menor calidad respecto de los obtenidos con lluvias estimadas por interpolación IDW, resultado que se atribuye a la cantidad reducida de estaciones pluviométricas, localizadas principalmente en el perímetro del área de captación de grandes dimensiones, lo que ocasionó una falta de aproximación en la lluvia estimada en un número mayor de celdas sobre la totalidad de la superficie analizada.

Por lo expuesto anteriormente, puede concluirse que la recuperación del comportamiento espacial de una tormenta histórica es posible si se dispone de una red de pluviómetros de densidad y distribución suficientes para complementarse con las precipitaciones obtenidas como productos de satélites y radares meteorológicos. En una condición más desfavorable, donde la disponibilidad de pluviómetros sea muy escasa, los registros serían insuficientes para transmitir la información sobre la totalidad de las celdas de la malla, especialmente en los sitios más distantes a los registros

disponibles sobre una superficie de gran tamaño, y la lluvia asimilada sobre la mayoría de las celdas dependería principalmente de los datos obtenidos por percepción remota, los cuales suelen presentar errores de exactitud. Así, la variación de los gastos en la salida del área de captación de lluvia podría presentar diferencias notorias respecto de los registros aforados en una estación hidrométrica, porque en la mayoría de las celdas de captación de lluvia la asimilación se habría realizado en ausencia de registros pluviométricos. Entonces se considera que, si la cantidad de pluviómetros a instalarse en una región debe ser francamente limitada, su distribución espacial debería realizarse de manera uniforme sobre la totalidad de la superficie de captación, lo que ayudaría a transmitir la información de manera más generalizada.

De los resultados obtenidos en este trabajo se observó que, en disposición de una red de pluviómetros de densidad media, las estimaciones por asimilación sí permiten conocer una aproximación del comportamiento espacial histórico de una tormenta, y por ello se considera que el espaciamiento promedio entre pluviómetros debe ser de hasta 10 km (100 km<sup>2</sup> por pluviómetro), lo que implica la instalación de un pluviómetro por cada celda de 0.1 grados de resolución.

Se concluye que el modelo de asimilación presentado es de gran utilidad para estimar la distribución espacial de eventos de lluvia diaria, aunque de igual manera podría emplearse para definir láminas acumuladas en periodos más cortos, pero es posible que tales estimaciones presenten errores de mayor magnitud, porque el cálculo de lluvia por percepción remota es mejor en escalas temporales más grandes.

Por otra parte, respecto a los productos de radar meteorológico (con resolución de 833 m) para el análisis de asimilación (con resolución de 0.1 grados), la lluvia para cada celda de la malla se definió a partir de las láminas más próximas al centroide. Sin embargo, en el procedimiento de asimilación también podía considerarse la totalidad de los valores de lluvia contenidos en cada imagen captada por el radar, pero anticipadamente se observó que muchos de estos datos manifiestan diferencias significativas respecto a las láminas registradas en los pluviómetros más cercanos. De manera semejante, como dato representativo de cada celda podía considerarse a la lámina correspondiente al centroide de dicha celda, o incluso al promedio de todas las láminas contenidas dentro de la misma celda, pero como se mencionó, los errores de exactitud observados entre las lluvias como productos de radar respecto de los registros de pluviómetros disponibles suelen ser importantes, y la calidad del producto final de asimilación resultaría más afectada.

Para finalizar, no se descarta que la calidad de las estimaciones de lluvia obtenidas con la combinación de las tres bases de datos sugeridas pueda mejorar, si se utiliza un método de asimilación más robusto, en conocimiento de que el procedimiento de correcciones sucesivas aplicado tiene entre sus desventajas principales la consideración de una función de peso que depende exclusivamente de la distancia entre las observaciones y la celda de la malla de interpolación, de manera que existe una dependencia muy fuerte con la magnitud de los datos más próximos al punto de asimilación, lo que puede trascender cuando el espaciamiento entre observaciones es irregular. No obstante, se considera que los resultados alcanzados son suficientes para corroborar la bondad de la aplicación del modelo propuesto, lo que representa un avance en cuanto a la posibilidad de obtener insumos apropiados para la modelación de escurrimientos superficiales.

## Referencias

- [1] Alcaraz, N., 2016. Análisis de la distribución espacial de la lluvia, caso de aplicación: cuenca del río Santa Catarina, RH 24 Bravo – Conchos. Facultad de Ingeniería, Tesis de Licenciatura, UNAM.
- [2] Álvarez, J., 2011. Estimación de la distribución espacial de la precipitación en zonas montañosas mediante métodos geoestadísticos. Departamento de Ingeniería Civil: Hidráulica y Energética, E.T.S.I. Caminos, Canales y Puertos.
- [3] Aparicio, F., 1992. Fundamentos de Hidrología de superficie. Editorial Limusa, México.
- [4] Arellano, F., 2012. Estimación alternativa del potencial de tormentas a partir de información climatológica antecedente. Caso de estudio Noroeste de México. Facultad de Ingeniería, Programa de Maestría y Doctorado en Ingeniería, Tesis de Doctorado, UNAM.
- [5] Barnes, S., 1964. A Technique for Maximizing Details in Numerical Weather Map Analysis. *Journal of Applied Meteorology*, 3(4), 396-409.
- [6] Barnes, S., 1973. Mesoscale objective map analysis using weighted time-series observations. NOAA Technical Memorandum. ERL NSSL – 62, National Severe Storms Laboratory, Norman, Oklahoma 73069, 60 pp.
- [7] Barth, A., Alvera, A., Joassin, P., Beckers, J.-M., Troupin, C., 2008. Introduction to Optimal Interpolation and Variational Analysis. *Geohydrodynamics and Environment Research*.
- [8] Basualdo, A., 2015. Manual de buenas prácticas para la generación, el almacenamiento y la difusión de informática climática en instituciones y organismos del MERCOSUR. Instituto Interamericano de Cooperación para la Agricultura, San José, Costa Rica.
- [9] Becerra, L., 2009. Estimación de lluvia en el Distrito Federal utilizando datos de pluviógrafos y de radar meteorológico. Facultad de Ingeniería, Programa de Maestría y Doctorado en Ingeniería, Tesis de Maestría, UNAM.
- [10] Bergthórsson, P., Döös, B., 1955. Numerical Weather Map Analysis. *Tellus*, Vol. 7, núm 3, 329-340.
- [11] Berndtsson, R., Niemczynowicz, J., 1988. Spatial and temporal scales in rainfall analysis – Some aspects and future perspectives. *Journal of Hydrology*, 100, 293-313.
- [12] Bladé, E., Gómez, M., 2006. Modelación del flujo en lámina libre sobre cauces naturales. Análisis integrado en una y dos dimensiones. Centro Internacional de Métodos Numéricos en Ingeniería, Monografía CIMNE N°-97, España.

- [13] Bladé, E., Cea, L., Corestein, G., Escolano, E., Puertas, J., Vázquez-Cendón, E., Dolz, J., Coll, A., 2014. Iber: herramienta de simulación numérica del flujo en ríos. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, 30(1): 1-10.
- [14] Blayo, E., Cosme, E., Nodet, M., Vidard. A., 2011. *Introduction to Data Assimilation*.
- [15] Bocquet, M., 2014. *Introduction to the principles and methods of data assimilation in the geosciences. Lecture notes, Master M2 MOCIS & WAPE, École des Ponts ParisTech*.
- [16] Bouttier, F., Courtier, P., 1999. *Data Assimilation Concepts and Methods. Meteorological Training Course Lecture Series, ECMWF*.
- [17] Caicedo, F., 2008. *Asimilación de precipitación estimada por imágenes de satélite en modelos hidrológicos aglutinados y distribuidos, caso de estudio afluencias al embalse de Betania (Huila, Colombia). Facultad de Ingeniería, Trabajo de Grado, Pontificia Universidad Javeriana*.
- [18] Calvetti, L., Beneti, C., Antunes, R., Toshio, R., Noronha, T., Gomes, A., Herdies, D., Gonçalves, L., 2017. *Quantitative Precipitation Estimation Integrated by Poisson's Equation Using Radar Mosaic, Satellite, and Rain Gauge Network. ASCE, Journal of Hydrologic Engineering, Vol. 22, No. 5*.
- [19] Campos, D., 1987. *Procesos del Ciclo Hidrológico. Universidad Autónoma de San Luis Potosí*.
- [20] Chow, V., Maidment, D., Mays, L., 1994. *Hidrología aplicada. McGRAW-HILL, Colombia*.
- [21] Chow, V., 2004. *Hidráulica de canales abiertos. McGRAW-HILL, Colombia*.
- [22] Cisneros, H., Domínguez, R., 2001. *Aplicación del método kriging en la construcción de campos de tormenta en la Ciudad de México. Ingeniería hidráulica en México, vol. XVI, núm. 3*.
- [23] CONAGUA, 2018. *Atlas del Agua en México. Comisión Nacional del Agua*.
- [24] CONAGUA, 2018. *Estadísticas del Agua en México. Comisión Nacional del Agua*.
- [25] Cressman, G., 1959. *An Operational Objective Analysis System. Weather Bureau, Vol. 87, No. 10, 367-374*.
- [26] Daley, R., 1991. *Atmospheric Data Analysis. Cambridge University Press, Cambridge*.
- [27] Daley, R. 1992. *The Lagged Innovation Covariance: A Performance Diagnostic for Atmospheric Data Assimilation. American Meteorological Society, Vol. 120, No. 1, 178-196*.
- [28] de Luna, F., 2019. *Modelo hidráulico de flujo bidimensional para estimar el escurrimiento a partir de la precipitación. Facultad de Ingeniería, Programa de Maestría y Doctorado en Ingeniería, Tesis de Doctorado, UNAM*.
- [29] Domínguez, R., Esquivel, G., Méndez, B., Mendoza, A., Arganis, M., Carrizosa, E., 2008. *Manual del Modelo para pronóstico de escurrimiento. Instituto de Ingeniería, UNAM*.
- [30] Domínguez, R., Arganis, M., Carrizosa, E., Hincapié, C., Fuentes, G., Osnaya, J., 2017. *Estudio para regionalizar los gastos generados por avenidas máximas, como base para la elaboración de mapas de peligro por inundaciones fluviales en todas las cuencas de la República*

Mexicana. Instituto de Ingeniería, UNAM, elaborado para Centro Nacional de Prevención de Desastres.

[31] Errico, R., VukićEvić, T., Raeder, K., 1993. Examination of the accuracy of a tangent linear model. *Tellus A: Dynamic Meteorology and Oceanography*, 45:5, 462-477.

[32] Eugenio, F., Marcello, J., Marqués, F., 2013. Tutorial de teledetección espacial. Universidad de Las Palmas de Gran Canaria. España.

[33] Fuentes, O., Franco, V., de Luna, F., Vélez, L., Morales, H., 2014. Caracterización fluvial e hidráulica de las inundaciones en México, Organismo de Cuenca X, Golfo Centro. Instituto de Ingeniería, UNAM, elaborado para Comisión Nacional del Agua.

[34] Fuentes, O., Domínguez, R., Morales, H., Esquivel, G., 2015. Manual de Diseño de Obras Civiles, Cap. A.1.5 Relación entre precipitación y escurrimiento. Instituto de Ingeniería, UNAM, elaborado para Comisión Federal de Electricidad.

[35] Gandin, L. 1965, Objective analysis of meteorological fields. U.S. Dept. Commerce and National Science Foundation, Washington, D.C. (Original en Ruso, ob"ektivnyi analiz meteorologicheskikh polei, 1963).

[36] Gebregiorgis, A., and Hossain F., 2011. How Much Can A Priori Hydrologic Model Predictability Help in Optimal Merging of Satellite Precipitation Products? *Journal of Hydrometeorology* 12, 1287-1298.

[37] Gilchrist, B., Cressman, G., 1954. An Experiment in Objective Analysis. *Tellus*, Vol. 6, No. 4, 309-318.

[38] Guillet, T., Teyssier, R., 2011. A Simple Multigrid Scheme for Solving the Poisson Equation with Arbitrary Domain Boundaries. *Journal of Computational Physics*, Vol. 230, No. 12.

[39] Harms, D., 1985. Application of an objective analysis scheme to mesoscale observational network design. A thesis submitted to the Graduate Faculty of North Carolina State University.

[40] Hernández, J., 2005. Control Inverso adaptativo. Proyecto Fin de Carrera, ETSi, Universidad de Sevilla.

[41] Huang, X., Vedel, H., 2002. An introduction to data assimilation. Danish Meteorological Institute.

[42] Huffman, G., 2020. The transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG). In: *Global Precipitation Measurement*, NASA, pp. 1-5.

[43] Ide, K., Courtier, P., Ghil, M., Lorenc, A., 1997. Unified Notation for Data Assimilation: Operational, Sequential and Variational. *Journal of the Meteorological Society of Japan*, Vol. 75, No. 1B, 181-189.

[44] Jiménez, S., 2016. Asimilación de datos de satélite para la caracterización de la lluvia ciclónica en México. Facultad de Ingeniería, Programa de Maestría y Doctorado en Ingeniería, Tesis de Maestría, UNAM.

- [45] Juárez, J., Arganis, M., Domínguez, R., Esquivel, G., Bladé, E., Dolz, J., Sánchez, M., Sánchez, H., Corestein, G., 2014. Comparación del hidrograma de salida de una cuenca con un modelo hidráulico y un modelo distribuido. XXIII Congreso Nacional de Hidráulica, Puerto Vallarta, Jalisco, México.
- [46] Kalman, R., 1960. A New Approach to Linear Filtering and Prediction Problems. ASME, Journal of Basic Engineering, Vol. 82(1), 35-45.
- [47] Kalman, R., Bucy, R., 1961. New Results in Linear Filtering and Prediction Theory. ASME, Journal of Basic Engineering, Vol. 83(1), 95-108.
- [48] Koch, S., desJardins, M., and Kocin, P., 1981. The GEMPAK Barnes Objective Analysis Scheme. NASA, Technical Memorandum 83851, NASA/GLAS, Greenbelt, MD 20771, 56 pp.
- [49] Koch, S., desJardins, M., and Kocin, P., 1983. An Interactive Barnes Objective Map Analysis Scheme for Use with Satellite and Conventional Data. Journal of Climate and Applied Meteorology, Vol. 22, No. 9.
- [50] Lahoz, W., Khattatov, B., Ménard, R., 2010. Data Assimilation: Making Sense of Observations, Springer, Berlin.
- [51] Le Dimet, F., Talagrand, O., 1986. Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. Tellus, 38A, 97-110.
- [52] Li, M., Shao, Q., 2010. An improved statistical approach to merge satellite rainfall estimates and raingauge data. Journal of Hydrology, 385, 51-64.
- [53] Lorenc, A., 2003. The potential of the ensemble Kalman filter for NWP---a comparison with 4D-Var. Royal Meteorological Society, Vol. 129, No. 595, 3183-3203.
- [54] Lozoya, J., Domínguez, R., Arganis, M., Hincapié, C., Becerril, M., Labrada, P., 2017. Manual de Diseño de Obras Civiles, Cap. A.1.7 Tormentas de diseño. Instituto de Ingeniería, UNAM, elaborado para Comisión Federal de Electricidad.
- [55] Lujano, E., Felipe, O., Lujano, A., Quispe, J., 2015. Validación de la precipitación estimada por satélite TRMM y su aplicación en la modelación hidrológica del río Ramis Puno Perú. Rev. Investig. Altoandin, Vol 17 No 2.
- [56] Méndez, B., 2005. Aplicación hidrológica de los radares meteorológicos. Facultad de Ingeniería, División de Estudios de Posgrado, Tesis de Doctorado, UNAM.
- [57] Méndez, B., Domínguez R., Magaña, V., Caetano, E., Carrizosa, E. 2006. Calibración hidrológica de radares meteorológicos. Ingeniería hidráulica en México XXI (4), 43-64.
- [58] Montoya, G., Eslava, J., Calderón, A., Vega, F., 2000. Implementación del método de Gandin para interpolar datos de precipitación en Colombia. Meteorología Colombiana, No. 1, 25-30.
- [59] Nanding, N., Rico, M., and Han, D., 2015. Comparison of different radar-raingauge rainfall merging techniques. Journal of Hydroinformatics, Vol. 17, No. 3, 422-445.
- [60] NASA. Navier – Stokes Equations. Glenn Research Center, National Aeronautics And Space Administration. Recuperado de: <https://www.grc.nasa.gov/www/k-12/airplane/nseqs.html>

- [61] Nerini, D., Zulkafli, Z., Wang, L.-P., Onof, C., Buytaert, W., Lavado-Casimiro, W., Guyot, J.-L., 2015. A Comparative Analysis of TRMM-Rain Gauge Data Merging Techniques at the Daily Time Scale for Distributed Rainfall-Runoff Modeling Applications. American Meteorological Society, Vol. 16, No. 5, 2153-2168.
- [62] OMM, 2011. Guía de prácticas hidrológicas, Volumen I: Hidrología – De la medición a la información hidrológica, OMM-N° 168, Organización Meteorológica Mundial.
- [63] OMM, 2017. Guía del Sistema Mundial de Observación. Organización Meteorológica Mundial.
- [64] OMM, 2017. Guía de Instrumentos y Métodos de Observación Meteorológicos, OMM-N° 8, Organización Meteorológica Mundial.
- [65] Ortega, A., Perozo, M., Marrero, N., 2014. Análisis espacial y temporal de la precipitación máxima en una cuenca hidrográfica de Venezuela. 17ª Convención Científica de Ingeniería y Arquitectura.
- [66] Panofsky, H. 1949. Objective weather-map analysis. Journal of Meteorology. Vol. 6, No. 6, 386-392.
- [67] Rosengaus, M., Sánchez, J., 1990. Gilbert: ejemplo de huracanes de gran intensidad. Ingeniería Hidráulica en México, enero – abril.
- [68] Rosengaus, M., 1995. Fundamentos de radares meteorológicos: aspectos clásicos (primera de dos partes). Instituto Mexicano de Tecnología del Agua. Ingeniería Hidráulica en México. Vol. X, Núm. 1.
- [69] Rosengaus, M., 1995. Fundamentos de radares meteorológicos: aspectos modernos (segunda de dos partes). Instituto Mexicano de Tecnología del Agua. Ingeniería Hidráulica en México. Vol. X, Núm. 2.
- [70] Rosengaus, M., Jiménez, M., Vázquez, M., 2002. Atlas climatológico de ciclones tropicales en México, Centro Nacional de Prevención de Desastres.
- [71] Rozante, J., Moreira, D., de Gonçalves, L., and Vila, D., 2010. Combining TRMM and Surface Observations of Precipitation: Technique and Validation over South America. American Meteorological Society, Vol. 25, No. 3, 885-894.
- [72] Santos, C., 2018. Física del caos en la predicción meteorológica. Ministerio para la Transición Ecológica, Agencia Estatal de Meteorología, Madrid.
- [73] Sasaki, Y., 1958. An Objective Analysis Bases on the Variational Method. Journal of the Meteorological Society of Japan. Ser II. Vol. 36, No. 3, 77-88.
- [74] Sinha S., Narkhedkar, S., and Mitra, A., 2006. Barnes objective analysis scheme of daily rainfall over Maharashtra (India) on a mesoscale grid. *Atmósfera*, 19(2), 109-126.
- [75] Skiba, Y., 2008. Introducción a la Dinámica de Fluidos. Centro de Ciencias de la Atmósfera, UNAM.
- [76] SMN, 2012. Servicio Meteorológico Nacional: 135 años de historia en México. Comisión Nacional del Agua.

- [77] Soto, F., González, M., 2019. Análisis de métodos estadísticos para evaluar el desempeño de modelos de simulación en cultivos hortícolas. *Agronomía Mesoamericana*, Universidad de Costa Rica, Nota técnica.
- [78] Tintoré, J., Alonso, S., Gomis, D., 1991. Análisis objetivo y diagnóstico en fluidos geofísicos. *Física de la Tierra*, núm 3, 179-218. Ed. Univ. Compl. Madrid, España.
- [79] Tsuyuki, T., Miyoshi, T., 2007. Recent Progress of Data Assimilation Methods in Meteorology. *Journal of the Meteorological Society of Japan*, Vol. 85B, 331-361.
- [80] Uni-Koeln, 2006. Objective Analysis. Recuperado de:  
<http://www.uni-koeln.de/math-nat-fak/geomet/eurad/lehre/oa-ws0809/oa-ws0809-vl.pdf>
- [81] Velázquez, J., 2014. Calibración de un modelo de estimación de lluvia con imágenes de satélite, utilizando datos de estaciones climatológicas, para la región hidrológica número 30 de México. Facultad de Ingeniería, Tesis de Licenciatura, UNAM.
- [82] Vila, D., Gonçalves, L., Toll, D., Rozante, J., 2009. Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America. *American Meteorological Society*, Vol. 10, No. 2, 533-543.
- [83] Vilchis, I., 2009. Integración y análisis espacial de la estimación de precipitación con imágenes de radar meteorológico. Facultad de Ingeniería, Centro Interamericano de Recursos del Agua, Tesis de Maestría, UAEM.
- [84] Vilchis, I., Quentin, E., Bâ, K., Díaz, C., 2011. Estimación de precipitación diaria a través de un SIG con imágenes de radar meteorológico. *Tecnología y Ciencias del Agua*, vol. II, núm. 4, 167-174.
- [85] Vilchis, I., 2013. Estimación y validación de precipitación derivada de imágenes de satélite y radar para uso en modelos hidrológicos. Facultad de Ingeniería, Centro Interamericano de Recursos del Agua, Tesis de Doctorado, UAEM.
- [86] Walther, B., Moore, J., 2005. The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, Vol. 28, No. 6, 815-829.
- [87] Yang, R., and Xing, B., 2021. A comparison of the performance of different interpolation methods in replicating rainfall magnitudes under different climatic conditions in Chongqing Providence (China). *Atmosphere* 12(10), 1318, 1-24.
- [88] Yu, L., Leng, G., Python, A., Peng, J., 2021. A Comprehensive Evaluation of Latest GPM IMERG V06 Early, Late and Final Precipitation Products across China. *Remote Sensing*, 13, 1208, 1-22.
- [89] <https://app.conagua.gob.mx/sistemasdeagua/>
- [90] <http://ciclotrop.com/radares-meteoroloacutegicos.html>
- [91] <https://datos.gob.mx/>
- [92] <https://giovanni.gsfc.nasa.gov/giovanni/>

[93] <https://www.sacmex.cdmx.gob.mx/>

[94] <https://smn.conagua.gob.mx/es/climatologia/informacion-climatologica/informacion-estadistica-climatologica>

# Anexo 1. Metodología propuesta para Asimilación de tres bases de datos de lluvia

## Recopilación de información histórica de lluvia

- Respecto de un sitio y fecha de interés, para establecer el estado de diagnóstico de un campo de lluvia, por ejemplo, recuperar el comportamiento espacial de una tormenta histórica, realizar la recopilación de datos registrados en simultáneo.
- De los registros pluviométricos diarios considerar:
  - Es indiferente el instrumento empleado para medir la lluvia diaria, es decir, conviene la recopilación de datos de pluviómetros y pluviógrafos.
  - En el armado de la red pluviométrica es apropiado tomar en cuenta estaciones de registro localizadas incluso afuera de la región definida por el partaguas de la zona por analizar, porque la Asimilación implica el intercambio de información entre celdas vecinas de una malla, y la mejor estimación para las celdas perimetrales puede requerir de láminas registradas en zonas distantes de la región central de la malla.
  - Revisar los registros diarios para la identificación de valores atípicos. Respecto de los atípicos distribucionales, tomar en cuenta los valores extremos anuales, los valores medios y las medidas de dispersión. Respecto de los atípicos espaciales, generalmente es de utilidad verificar la precipitación simultánea de láminas de un mismo orden de magnitud en las estaciones próximas.
- De los registros de sensores satelitales considerar:
  - Se recomienda el empleo de la lluvia generada como producto de la misión satelital GPM, para latitudes de 90°N a 90°S, por su gran resolución (0.1°x0.1°) y amplia disponibilidad (lluvia desde el año 2000), que puede obtenerse a través de internet desde la base de datos GIOVANNI de la NASA, mediante una plataforma de manejo intuitivo, sin costo alguno.
  - De acuerdo con las necesidades del análisis, optar por alguno de los productos satelitales: IMERG Early, para el análisis de eventos en tiempo casi-real; IMERG Late, disponible doce horas después de ocurrido un evento; IMERG Final, disponible tres y medio meses después, siendo el producto de máxima calidad.
- De los registros de radar meteorológico considerar:
  - Verificar la existencia de algún radar meteorológico próximo a la zona de análisis, a una distancia de menor que su radio de cobertura. En su caso, realizar una solicitud de datos a la ventanilla única del SMN, para conocer sobre la disponibilidad de imágenes PPI registradas en la fecha de interés.
  - Se recomienda aplicar la metodología presentada por Vilchis *et al.* (2011), para efectuar un análisis robusto de imágenes crudas de reflectividad, que permita disponer de estimaciones de lluvia aproximada a las tormentas reales.
  - Como consideración práctica, realizar la Asimilación de lluvia sobre los centroides de cada celda de la malla regular (0.1°x0.1°), definida en la descarga electrónica de archivos ráster del sitio GIOVANNI. Entonces, con la lluvia de radar (resolución de 833 m, por ejemplo), obtener un valor representativo para cada celda, y se recomienda el promedio de las nueve láminas de resolución de 833 m más cercanas a cada centroide.
- Verificar la asociación lineal que debería existir entre los conjuntos de datos de pluviómetros y de percepción remota. Con la construcción de diagramas de dispersión puede observarse la relación de los conjuntos alrededor de la línea de error nulo. Pueden

identificarse condiciones de subestimación o sobreestimación, e incluso establecerse que la aportación de la variable secundaria (percepción remota) será menos significativa si la correlación de los conjuntos es pobre.

- Identificar las zonas de acumulación de lluvia con la construcción de mapas de la distribución espacial de las tormentas, mediante una interpolación simple de las observaciones pluviométricas. Así verificar que la distribución de las acumulaciones máximas por interpolación esté representada en los productos de percepción remota, lo que justifica su empleo en el proceso de Asimilación.

### **Asimilación de datos de precipitación**

- Aplicación de la metodología de Asimilación de datos de lluvia secuencial por correcciones sucesivas, considerando la Técnica de Análisis Objetivo de Barnes, Capítulo 3.
- Para aplicar la metodología de Asimilación de tres bases de datos, se recomienda el empleo de una herramienta que agilice los cálculos, como el programa que se proporciona en el Anexo 2, el cual requiere: conjunto de observaciones de precipitación acumulada diaria (registros pluviométricos y estimaciones de satélite y radar), dimensiones del área de estudio, y parámetros de potencia y de convergencia.
- Como primer paso de la metodología, realizar una estimación del estado del campo de lluvia distribuida sobre los centroides de la malla de análisis, utilizando el conjunto de observaciones disponibles y la función de peso inicial.
- Como segundo paso de la metodología, corregir el estado antecedente de la lluvia, aplicando la función de peso modificada sobre las discrepancias (innovaciones) entre las observaciones de estados verdaderos y estados estimados.
- En ausencia de datos de radar, realizar la Asimilación de registros pluviométricos con datos de origen satelital, en conocimiento de que el resultado será una aproximación respecto del mejor resultado, especialmente si los datos de satélite guardan un orden de magnitud similar a los registros pluviométricos.
- Las interpolaciones IDW de datos pluviométricos pueden ser mejores para analizar días de precipitación reducida, si esta circunstancia contribuye a incrementar la inexactitud de las estimaciones de percepción remota.
- Si la densidad de la red de pluviómetros es baja la calidad de la Asimilación se reduce, hasta errores comparables con los obtenidos por interpolación IDW, e inclusive con resultados menos aceptables si la distribución los pluviómetros es muy irregular.

### **Calibración del modelo de Asimilación**

- En el modelo de Asimilación propuesto intervienen algunos parámetros que deben ajustarse con la intención de obtener estimaciones de lluvia de mejor aproximación, por lo que se requiere practicar un procedimiento de calibración. En conocimiento de la distribución espacial de eventos de lluvia históricos sobre la región de interés, considerar lo siguiente:
  - El procedimiento de calibración respecto de los datos de un evento de lluvia conocido se realizará mediante la asignación de valores apropiados a los parámetros de convergencia  $\gamma$  y de potencia  $\beta$ , tomando en cuenta que  $\gamma$  influye en la Asimilación de cada celda de la malla, al participar en el suavizado de la función de peso modificada, mientras que  $\beta$  solo influye en la estimación de las celdas periféricas de la malla, donde con valores más grandes de este parámetro se incrementa la influencia de las observaciones más cercanas.

- Es conveniente iniciar la calibración suponiendo un valor fijo  $\beta$  igual a 2 unidades, de aplicación común en la literatura, y obtener estimaciones de la lluvia considerando variaciones de  $\gamma$  entre 0 y 1, para identificar el valor de  $\gamma$  que proporciona un campo de lluvia con errores de menor magnitud.
- Como segunda etapa de calibración, practicar la variación del parámetro  $\beta$ , en el intervalo de 1 a 10 unidades, manteniendo fijo el valor de  $\gamma$  identificado anteriormente, para establecer la combinación de valores  $\gamma$  y  $\beta$  que mejora la estimación del campo de lluvia correspondiente al evento analizado.
- El procedimiento ha de repetirse para el análisis de otros eventos históricos de lluvia ocurridos en la región de estudio, con la intención de identificar posibles combinaciones de valores  $\gamma$  y  $\beta$  que permitan obtener resultados de calidad, en el entendido de que existe un orden de magnitud de los parámetros que puede considerarse representativo de la región.

### **Cálculo de errores de Asimilación**

- Para evaluar la calidad de la estimación de un campo de lluvia de un evento estimado por Asimilación de datos, realizar un procedimiento de validación cruzada. Antes de la Asimilación, remover hasta un 10% de los datos pluviométricos disponibles, procurando evitar una modificación drástica de la densidad y distribución original de la red de registros, de manera que dichos valores se consideren como referencias para la medición de errores.
- Para la medición matemática de errores de las lluvias estimadas por Asimilación, se recomienda emplear diferentes criterios estadísticos, como son: el error medio ( $ME$ ), el error absoluto medio ( $MAE$ ), la raíz del error cuadrático medio ( $RMSE$ ), el coeficiente de eficiencia de Nash-Sutcliffe ( $NSE$ ) y el coeficiente de correlación ( $CC$ ), Capítulo 3.

## Anexo 2. Programa para Asimilación de datos

En la primera parte del código se concentra la información requerida como datos de entrada, además de una lista de las variables involucradas en los cálculos de los ciclos o rutinas que conforman el programa para Asimilación de datos en lenguaje MATLAB R2017a.

```
clc
disp('Método para calcular precipitación pluvial por método de Barnes')
disp('Elaboró: María Asucena Rodríguez Ramirez, Semestre: 2022-1,
Programa de Doctorado en Ingeniería, UNAM')
%disp('Previamente debiste haber importado la matriz de Excel de las
coordenadas UTM y Estaciones Pluviométricas')
%DATOS:
%Tres bases de datos (puntuales, radar, satélite)
%L, Longitud de celda, está en función de la resolución de los datos de
satélite
%Ae, Área de estudio
%No, Número de observaciones puntuales que se van a utilizar
%dn, Espaciamiento entre observaciones
%c, Parámetro de escala de longitud

%VARIABLES que intervienen en el proceso de programación:
%N, número de observaciones que hay en cada una de las celdas
%d, distancia que existe entre el centroide y los datos contenidos en la
celda de análisis
%WM, función de peso por cada observación
%S1, numerador de la ecuación de la Etapa 1
%S2, denominador de la ecuación en la Etapa 2
%SG1, corresponde a la primera estimación (Etapa 1), aplicando el método
de Barnes
%SWM2, numerador en la división de la Etapa 2
%SWM2i, denominador en la división de la Etapa 2
%EPER, observaciones periféricas contenidas en las celdas de la malla
%acumper, acumulada periférica
%D, distancia que existe entre las observaciones contenidas en la celda
de análisis con respecto a los centroides de las tres celdas adyacentes
%DM, numerador de la división para calcular LAMBDA
%DMA, denominador de la división para calcular LAMBDA
%LAMBDA=DM/DMA
%Sg, valores obtenidos de la interpolación, utilizando IDW, aplicado solo
para celdas periféricas
%Ixy, valores obtenidos de la interpolación, utilizando método Bilineal,
solo para celdas interiores
%SG2, corresponde a la segunda estimación (Etapa 2), aplicando el método
de Barnes
%E, E=(EPLUV), tabla de información sobre las observaciones puntuales (5
columnas: CLAVE, coordenadas en X, Y, Precipitación y numeración
progresiva)
%C, C=(CUTM), tabla con las coordenadas de los centroides X, Y de cada
una de las celdas de la malla
%C2, C2=(CUTM2), tabla con las coordenadas de los centroides de las
celdas de la malla (presentadas en forma vertical, 2 columnas)
%gama, variable que representa un parámetro de convergencia
%beta, variable involucrada en el método de interpolación IDW
(Ponderación de la Distancia Inversa)
```

Posteriormente, se solicita información para realizar cálculos iniciales como son el número de renglones y columnas que componen la malla, la superficie de la misma, y el número total de observaciones de lluvia involucradas en la ejecución del programa.

```

%Longitud de celda es l
ciudad=input('Teclea el nombre de la cuenca a estudiar:', 's');
lx=input('Teclee la longitud de la malla en la dirección horizontal en
km:');
ly=input('Teclee la longitud de la malla en la dirección vertical en
km:');
opcion='n';
while(opcion=='n')
l=input('Teclee la longitud de la celda en km:');
cv=round(lx/l);
ch=round(ly/l);
disp('El cálculo se va a hacer a tantas celdas horizontales por
verticales')
disp([ch cv])
opcion=input('Desea continuar el cálculo con los datos previos o teclear
otros:', 's');
end

format short g
Ae=ch*cv*1*1;
disp('Área de estudio km2:')
disp(Ae)

ep=input('Teclee la cantidad de estaciones pluviométricas:');
No=ch*cv*2+ep;
disp('Número de observaciones:')
disp(No)

dn=(Ae/No)^.5;
disp('Espaciamiento entre observaciones:')
disp(dn)

c=5.052^.5*2*dn/pi;
disp('Parámetro de escala de longitud:')
disp(c)

E=table2array(EPLUV);
C=table2array(CUTM);
SG1=zeros(ch,cv);
acum=0;
acumper=0;
N=zeros(ch,cv);

```

El programa consta de nueve ciclos para llegar al resultado final, es decir, las láminas de lluvia distribuidas espacialmente sobre la superficie de una malla, en unidades de mm, convenientemente localizadas en el centroide de cada celda que la conforma, y que provienen de la combinación de las tres bases de datos consideradas: red de pluviómetros, satélite y radar.

En el primer ciclo se calcula el número de observaciones contenidas en cada celda de la malla. Los resultados de esta variable (N) son presentados en una matriz, que se muestra durante la ejecución del programa.

```

%CICLOS PARA CONOCER EL NÚMERO DE OBSERVACIONES EN CADA CELDA
for (x=1:ch)
    np=0;
    npp=0;
    for (y=1:cv)
        for (i=1:No)%CAMBIE EP POR NO
            if abs(E(i,2)-C(x,2*y-1))/1000<=(1/2) && abs(E(i,3)-
C(x,2*y))/1000<=(1/2)
                np=np+1;
            end
        end
        N(x,y)=np-npp;
        if x==1 || x==ch || y==1 || y==cv
            acumper=acumper+np-npp;
        end
        acum=acum+(np-npp);
        npp=np;
    end
end
disp('Número de observaciones en cada una de las celdas')
disp(N)
%disp(acumper);

```

En el segundo ciclo se identifican las observaciones contenidas únicamente en las celdas periféricas de la malla, información requerida en el proceso de interpolación IDW.

```

%OBTENER MATRIZ DE ESTACIONES PERIFÉRICAS
EPER=zeros(acumper,9);
j=0;
for (x=1:ch)
    for (y=1:cv)
        for (i=1:No)%CAMBIE EP POR NO
            if abs(E(i,2)-C(x,2*y-1))/1000<=(1/2) && abs(E(i,3)-
C(x,2*y))/1000<=(1/2)
                if x==1 || x==ch || y==1 || y==cv
                    j=j+1;
                    EPER(j,1)=E(i,2);%X DE LA ESTACIÓN
                    EPER(j,2)=E(i,3);%Y ESTACIÓN
                    EPER(j,3)=E(i,4);%PRECIPITACIÓN
                    EPER(j,4)=x;%RENGLÓN DE LA CELDA
                    EPER(j,5)=y;%COLUMNA DE LA CELDA
                    EPER(j,6)=N(x,y);%ESTACIONES EN ESA CELDA
                    EPER(j,7)=C(x,2*y-1);%CENTROIDE CELDA EN X
                    EPER(j,8)=C(x,2*y);%CENTROIDE CELDA EN Y
                    EPER(j,9)=E(i,5);%NÚMERO DE ESTACIÓN
                end
                % acum=acum+(np-npp);
                %npp=np;
            end
        end
    end
end
end

```

En el tercer ciclo se determina la distancia D que existe entre las observaciones contenidas en cada celda periférica y sus correspondientes celdas adyacentes, información requerida en el método de interpolación IDW.

```

%VOY A UBICARME EN CADA ESTACION PERIFÉRICA Y VOY A SACAR SU DISTANCIA
%CONTRA SUS VECINAS
D=zeros(acumper*3,10);
z=1;
for(i=1:acumper)
    for(x=1:ch)
        for(y=1:cv)
            %if (EPER(i,4)==1 && EPER(i,5)==1) || (EPER(i,4)==1 &&
EPER(i,5)==cv) || (EPER(i,4)==ch && EPER(i,5)==1) || (EPER(i,4)==ch &&
EPER(i,5)==cv)%esquina
                if abs(EPER(i,7)-C(x,2*y-1))/1000<=1 && abs(EPER(i,8)-
C(x,2*y))/1000<=1
                    if (x~=EPER(i,4) || y~=EPER(i,5))
                        %if (EPER(i,4)==1 && EPER(i,5)==1) || (EPER(i,4)==1 &&
EPER(i,5)==cv) || (EPER(i,4)==ch && EPER(i,5)==1) || (EPER(i,4)==ch &&
EPER(i,5)==cv)%esquina
                            d=((EPER(i,1)-C(x,2*y-1))^2+(EPER(i,2)-C(x,2*y))^2)^.5;
                            if(d~=0)
                                D(z,1)=d;%distancia
                                D(z,2)=x;%x contra la que se compara
                                D(z,3)=y;%y contra la que se compara
                                D(z,4)=EPER(i,9);%estación
                                D(z,5)=EPER(i,4);%celda donde estoy en renglón
                                D(z,6)=EPER(i,5);%celda donde estoy en columna
                                D(z,7)=EPER(i,7);%centroide celda x
                                D(z,8)=EPER(i,8);%centroide celda y
                                D(z,9)=C(x,2*y-1);%centroide celda contra x
                                D(z,10)=C(x,2*y);%centroide celda contra Y
                                z=z+1;
                            end
                        end
                    end
                end
            end
        end
    end
end

ld=length(D);
beta=input('Teclea el valor de beta:');
C2=table2array(CUTM2);
DMA=zeros(acumper,1);
DM=zeros(ld,6);
LAMBDA=zeros(ld,1);
Sg=zeros(acumper,4);
dma=0;
k=1;
for(i=1:ld)
    DM(i,1)=(1/D(i,1))^beta;
    DM(i,2)=D(i,2);
    DM(i,3)=D(i,3);
    DM(i,4)=D(i,4);
    DM(i,5)=D(i,5);
    DM(i,6)=D(i,6);
    dma=dma+DM(i,1);
    if(i==ld)
        DMA(k,1)=dma;
        DMA(k,2)=i;
    end
end

```

```

end
if (i+1<=ld)
if (D(i,4)~=D(i+1,4))
    DMA(k,1)=dma;
    DMA(k,2)=i;
    k=k+1;
    dma=0;
end
end
end
end

```

En el cuarto ciclo se calcula el peso de cada observación, en función de la distancia que la separa respecto al centroide de la celda de análisis, en la primera etapa del método de Barnes.

```

%CICLO PARA OBTENER LOS WM
r=ep+2*ch*cv;
WM=zeros(r,1);
for (x=1:ch)
    for (y=1:cv)
        for (i=1:r) %SOLO SACARLO PARA LAS DE LA CELDA
            if abs(E(i,2)-C(x,2*y-1))/1000<=(1/2) && abs(E(i,3)-
C(x,2*y))/1000<=(1/2)
                dm=((C(x,2*y-1)-E(i,2))^2+(C(x,2*y)-E(i,3))^2)^.5;
                e=- (dm/(c*1000))^2;
                wm=exp(e);
                WM(i,1)=wm;
                %disp([wm e dm C(x,2*y-1)-E(i,2) C(x,2*y)-E(i,3) x y])
            end
        end
    end
end
end
%disp(WM)

```

En el quinto ciclo se obtiene la primera estimación del método de Barnes, SG1, para cada una de las celdas de la malla, y los resultados están disponibles en un arreglo matricial, al finalizar la ejecución del programa.

```

%CÁLCULO DE SG1
for (x=1:ch)
    for (y=1:cv)
        n=N(x,y);
        s1=2;
        s2=0;
        for (i=1:n)
            for (j=1:ep+2*ch*cv) %SOLO SACARLO PARA LAS DE LA CELDA
                if abs(E(j,2)-C(x,2*y-1))/1000<=(1/2) && abs(E(j,3)-
C(x,2*y))/1000<=(1/2)
                    s1=s1+WM(j,1)*E(j,4);
                    s2=s2+WM(j,1);
                end
            end
        end
        SG1(x,y)=s1/s2;
    end
end
end
end

```

En el sexto ciclo, como parte de la segunda etapa del método de Barnes, se obtienen valores de lluvia que son estimaciones de las observaciones de las celdas periféricas, mediante interpolación IDW, considerando las lluvias SG1 de las celdas adyacentes, obtenidas en la primera etapa.

```
%CÁLCULO DE LA LAMBDA
k=1;
sg=0;
LAMBDA=zeros(ld,1);
for (i=1:ld)
    lambda=DM(i,1)/DMA(k,1);
    LAMBDA(i,1)=lambda;
    x=DM(i,2);
    y=DM(i,3);
    LAMBDA(i,1)=lambda*SG1(x,y);%ATERRIZAR A LA CELDA
    sg=sg+lambda*SG1(x,y);
    if (i==ld)
        Sg(k,1)=sg;
        Sg(k,2)=DM(i,5);
        Sg(k,3)=DM(i,6);
        Sg(k,4)=DM(i,4);
    end
    if (i+1<=ld)
        if (D(i,4)~=D(i+1,4))
            Sg(k,1)=sg;
            Sg(k,2)=DM(i,5);
            Sg(k,3)=DM(i,6);
            Sg(k,4)=DM(i,4);
        end
        k=k+1;
        sg=0;
    end
end
end
```

En el séptimo ciclo, dentro de la segunda etapa del método de Barnes, se obtienen valores de lluvia que son estimaciones de las observaciones de las celdas interiores, mediante interpolación Bilineal, considerando las lluvias SG1 de las celdas adyacentes, obtenidas en la primera etapa.

```
%CÁLCULO BILINEAL
I=zeros(acum,6);
j=1;
for (i=1:ch*cv)
    if C2(i,4)==1 || C2(i,4)==ch || C2(i,5)==1 || C2(i,5)==cv
        x=C2(i,4);
        y= C2(i,5);
        %disp([x y])
        %disp ('periférica')
    else
        x=C2(i,4);
        y= C2(i,5);
        %disp ('intermedia')
        for (k=1:No)%PLUVIÓMETROS cambie ep por No
            if abs(E(k,2)-C2(i,1))/1000<=(1/2) && abs(E(k,3)-
C2(i,2))/1000<=(1/2)
                Iarr=(C2(i+cv,1)-E(k,2))/(C2(i+cv,1)-C2(i-ch-1,1))*SG1(x-1,y-
1)+(E(k,2)-C2(i-ch-1,1))/(C2(i+cv,1)-C2(i-ch-1,1))*SG1(x-1,y+1);
                Iab=(C2(i+cv,1)-E(k,2))/(C2(i+cv,1)-C2(i-ch-1,1))*SG1(x+1,y-
1)+(E(k,2)-C2(i-ch-1,1))/(C2(i+cv,1)-C2(i-ch-1,1))*SG1(x+1,y+1);
```



```

        if abs(E(i,2)-C(x,2*y-1))/1000<=(1/2) && abs(E(i,3)-
C(x,2*y))/1000<=(1/2)
            dm=((C(x,2*y-1)-E(i,2))^2+(C(x,2*y)-E(i,3))^2)^.5;
            e=-((dm/(c*1000))^2)/gama;
            wm1=exp(e);
            %for(j=1:No)%cambie acumper por acum
            %if (Sg(j,4)==E(i,5))
            for(k=1:acum)
                if(I(k,6)==E(i,5))
                    break;
                end
            end
            WM1(i,1)=wm1*(E(i,4)-I(k,3));
            sWM2=sWM2+ WM1(i,1);
            WM1i(i,1)=wm1;
            sWM2i=sWM2i+WM1i(i,1);
            WM1(i,2)=x;
            WM1(i,3)=y;
            WM1i(i,2)=x;
            WM1i(i,3)=y;
            %end
            %end
        end
    end
    SWM2(x,y)=sWM2;
    SWM2i(x,y)=sWM2i;
    sWM2=0;
    sWM2i=0;
end
end
end

```

Por último, en el noveno ciclo se concluye con la segunda etapa del método de Barnes, y se obtiene el valor SG2 para cada una de las celdas de la malla de estudio, con la presentación de los resultados de lluvia en un arreglo matricial.

```

%CÁLCULO DE SG2
SG2=zeros(ch,cv);
for(x=1:ch)
    for(y=1:cv)
        SG2(x,y)=SG1(x,y)+SWM2(x,y)/SWM2i(x,y);
    end
end

disp("SG2 RESULTADO FINAL de lluvia, en mm")
disp(SG2)

```

## Anexo 3. Metodología para estimación de lluvia con imágenes de radar

**Obtención de lluvia acumulada diaria a través del uso del SIG Idrisi, mediante la metodología propuesta por Vilchis *et al.* (2011).**

Como un sistema para el análisis y seguimiento de datos geoespaciales, TerrSet es un software desarrollado por Clark Labs, y se define como una constelación integrada de aplicaciones para el seguimiento y modelado de sistemas terrestres.

Por su parte, IDRISI es un software integrado a TerrSet que ofrece un conjunto de herramientas analíticas para el análisis de información geográfica, y está orientado a datos de tipo ráster para el procesamiento de imágenes de satélite. Es uno de los SIG más difundidos en el mundo debido a su capacidad de análisis, planificación y gestión del territorio, al contener más de 300 módulos para la visualización y el análisis de información, entre ellos, los módulos analíticos agrupados en secciones: búsqueda selectiva, operadores matemáticos, análisis de distancia, operadores de contexto, estadísticas, tratamiento digital de imágenes, etc.

Con respecto a la información necesaria para efectuar el proceso de obtención de imágenes de lluvia espacialmente distribuida sobre una malla uniforme, acumulada a lo largo de un día (imagen en formato .tif), se dispuso de dos carpetas: la primera con 96 imágenes PPI de reflectividad de un día, proporcionados por el SMN; la segunda carpeta integrada por Rutinas o Macros (Figura A.3-1), mediante las cuales se ejecutan los pasos de la metodología propuesta por Vilchis *et al.* (2011).

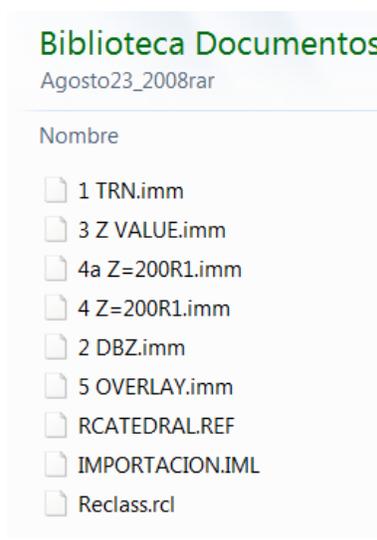


Figura A.3-1: Rutinas o Macros ejecutadas en el SIG Idrisi.

Dentro de la aplicación TerrSet se selecciona el proyecto de trabajo, y a continuación se realiza la importación de las imágenes PPI del radar dentro del SIG Idrisi, las cuales tienen una estructura definida para facilitar su manejo y control: nombre, fecha, tiempo UTC y producto PPI.

La importación de los archivos se realiza a través del buscador del SIG Idrisi, con la palabra Run Macro. Al desplegarse la ventana correspondiente y seleccionar "IMPORTACION", se procede a la ejecución del proceso con la opción "Run Macro", para importar las imágenes contenidas en el proyecto elegido (Figura A.3-2).



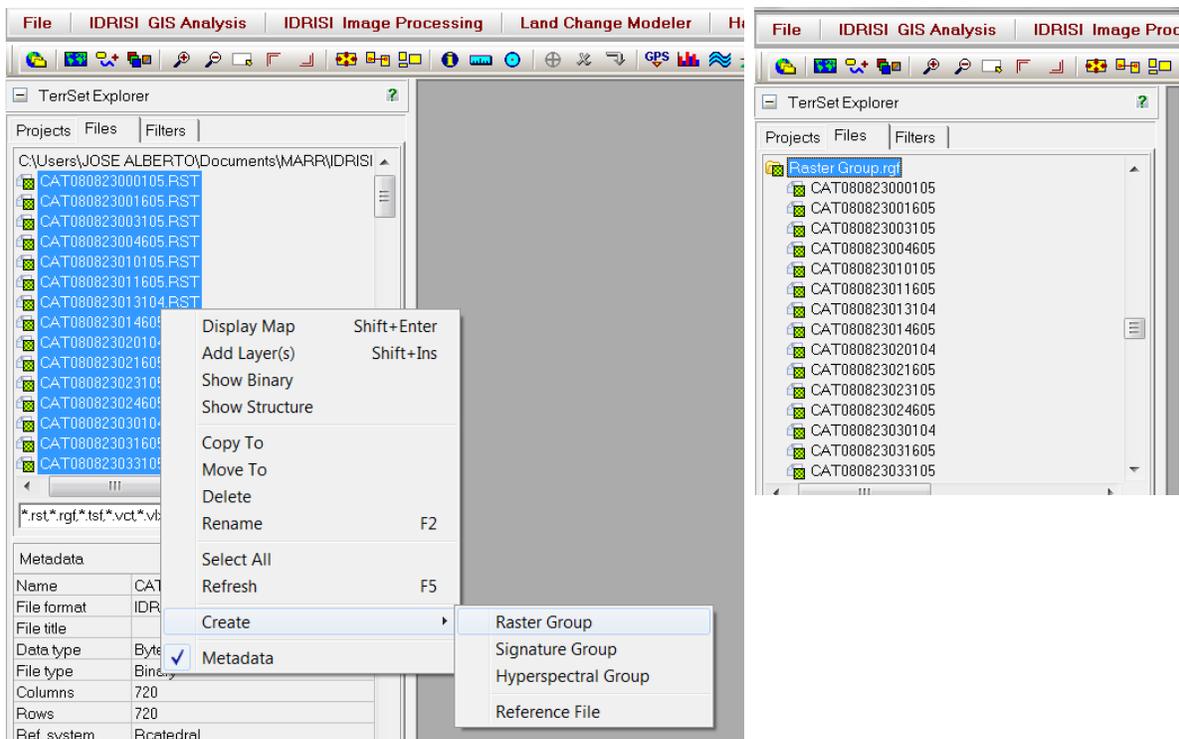


Figura A.3-4: Creación del Raster Group.

Como consecuencia de la dirección de escaneo que siguió el radar, las imágenes originales de reflectividad presentan un giro de 180° (Figura A.3-5), y por ello se requiere de su transposición.

Dentro del SIG Idrisi se selecciona el ícono Macro Modeler , desplegándose una ventana donde puede abrirse la rutina correspondiente para su ejecución (Figura A.3-6); con esto se logra el giro necesario respecto al estado original, generándose un nuevo archivo de salida para cada una de las imágenes (Figura A.3-7).

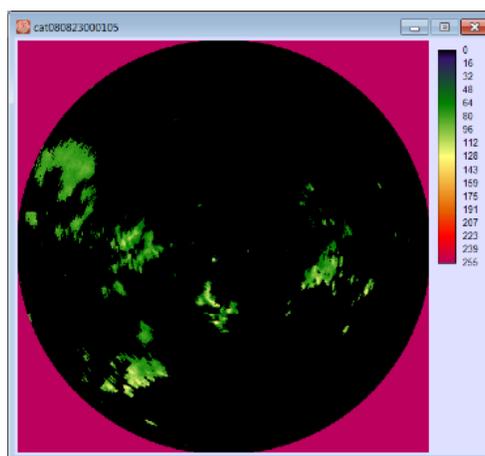


Figura A.3-5: Imagen original.

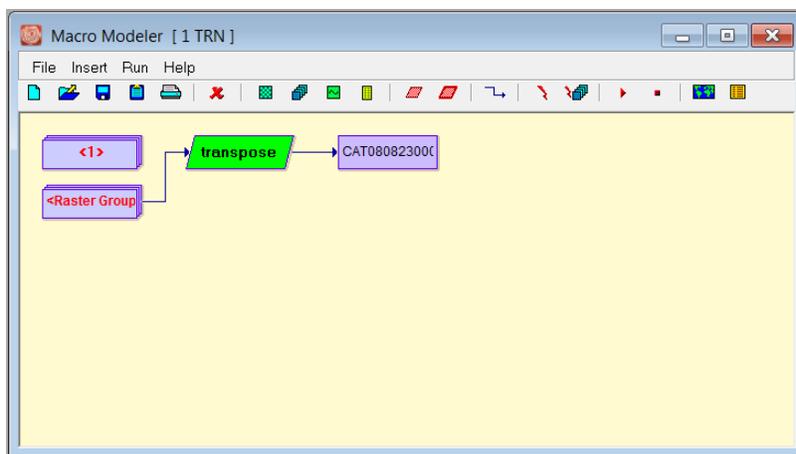


Figura A.3-6: Ejecución del proceso para la transposición de las imágenes originales.

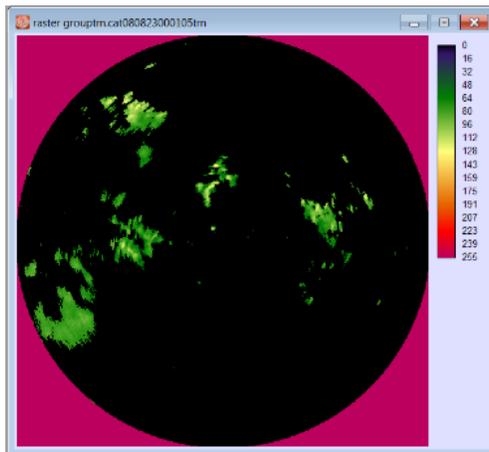


Figura A.3-7: Imagen girada 180°.

En una etapa posterior se realiza la conversión de los valores N contenidos en las imágenes, los cuales se encuentran en un rango de 0 a 255 unidades. Los valores en dBZ (decibelios de reflectividad) se calculan aplicando la ecuación  $\text{dBZ} = 0.5N - 32$ .

Para llevar a cabo lo anterior, dentro del SIG Idrisi se selecciona el ícono Macro Modeler , donde se solicita la opción correspondiente (2 DBZ.imm, en la Figura A.3-1), como se aprecia en la Figura A.3-8, y con ello se obtiene la imagen en decibelios de Z (dBZ), Figura A.3-9.

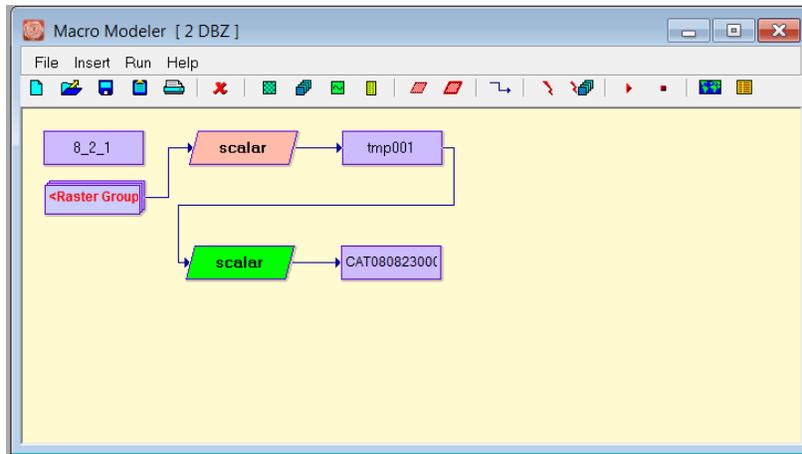


Figura A.3-8: Ejecución de rutina para obtener los decibelios de reflectividad (dBZ).

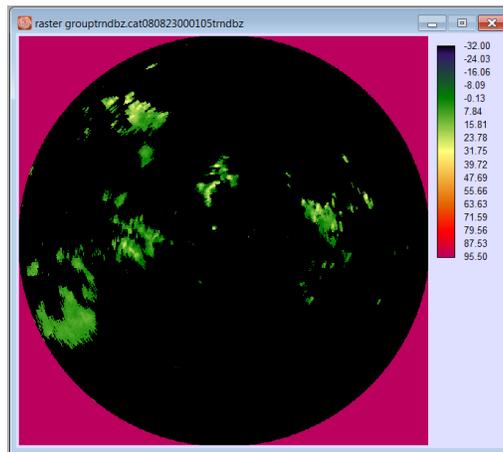


Figura A.3-9: Imagen modificada para la obtención de decibelios de Z (dBZ).

Posteriormente, se obtienen los valores de reflectividad (Z) para cada una de las imágenes. Para ello se selecciona el ícono Macro Modeler , y se ejecuta la rutina que corresponde al cálculo de la reflectividad, Figura A.3-10.

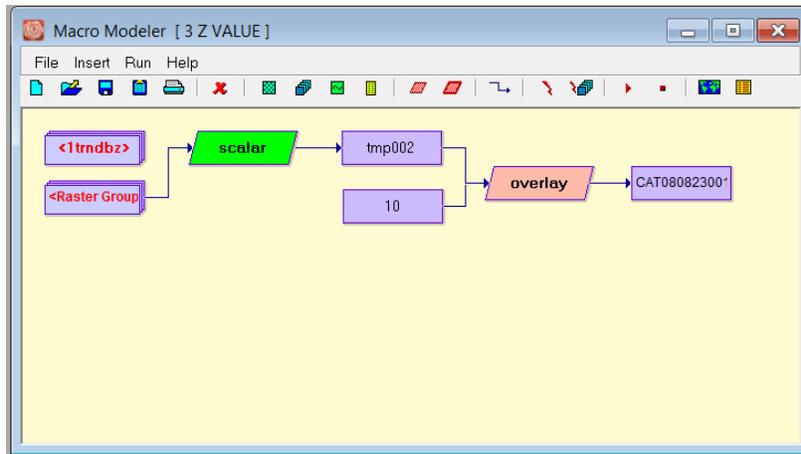


Figura A.3-10: Ejecución de rutina para obtener la reflectividad.

Para el cálculo de la intensidad de lluvia, nuevamente se selecciona el ícono Macro Modeler , y se abre la rutina que corresponde al cálculo de la intensidad de lluvia. En este paso puede eliminarse del ruido contenido en las imágenes, representado en color magenta, por medio el comando Reclass, Figura A.3-11.

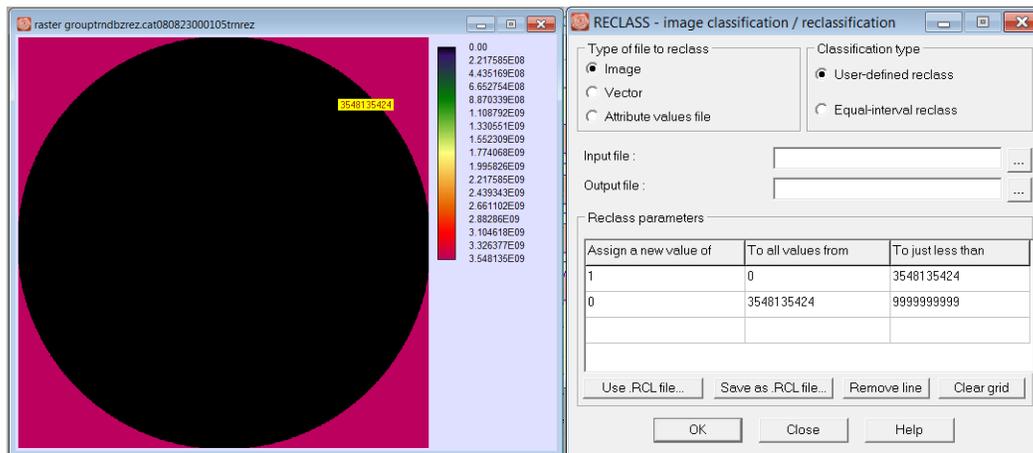


Figura A.3-11: Eliminación del ruido de la imagen.

Para la obtención de la intensidad de lluvia a partir de las imágenes de reflectividad, se aplicaron dos relaciones de Z-R. Las lluvias estratiformes se analizaron con la ecuación propuesta por Marshall-Palmer, y las lluvias convectivas mediante la ecuación de NEXRAD. Se realiza la ejecución de la rutina como se muestra en la Figura A.3-12. Los resultados arrojan intensidades de lluvia en mm a cada 15 minutos.

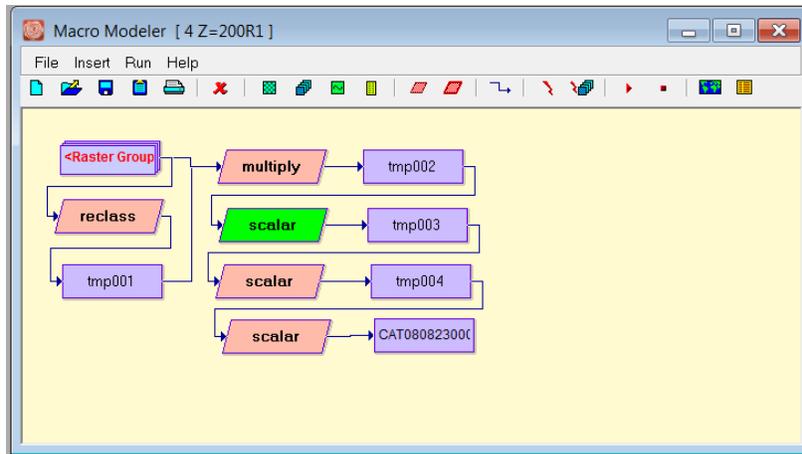


Figura A.3-12: Ejecución de rutina para el cálculo de la intensidad de lluvia.

Para la obtención de la lluvia acumulada en un día, es necesario sobreponer las imágenes y generar una sola imagen ráster en el SIG Idrisi. Para llevar a cabo lo anterior, se selecciona el ícono Macro Modeler , y se abre la rutina correspondiente, Figura A.3-13. En la Figura A.3-14 se muestra la imagen con la lluvia acumulada en un día.

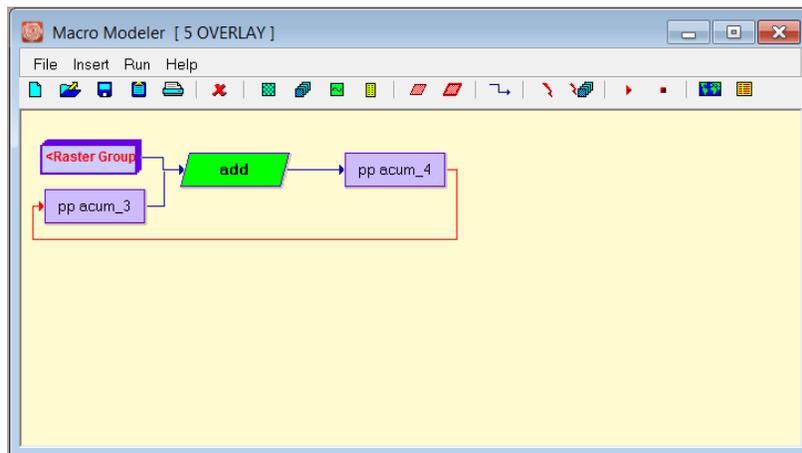


Figura A.3-13: Ejecución de rutina para la obtención de lluvia acumulada en un día.

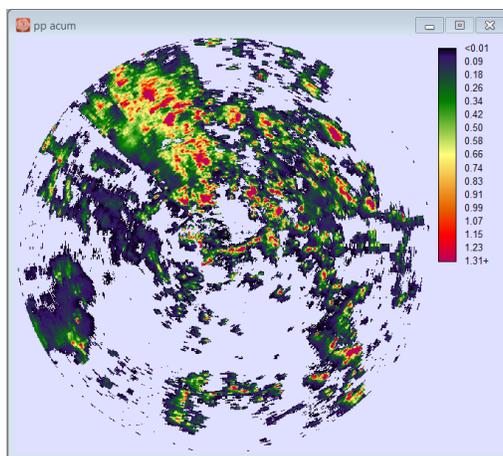


Figura A.3-14: Lluvia acumulada en un día.

La imagen de la Figura A.3-14 requiere del cambio en su sistema de referencia a coordenadas geográficas, por lo que a través del buscador del SIG Idrisi se solicita el comando “Project”, para después proporcionar la información necesaria, Figura A.3-15.

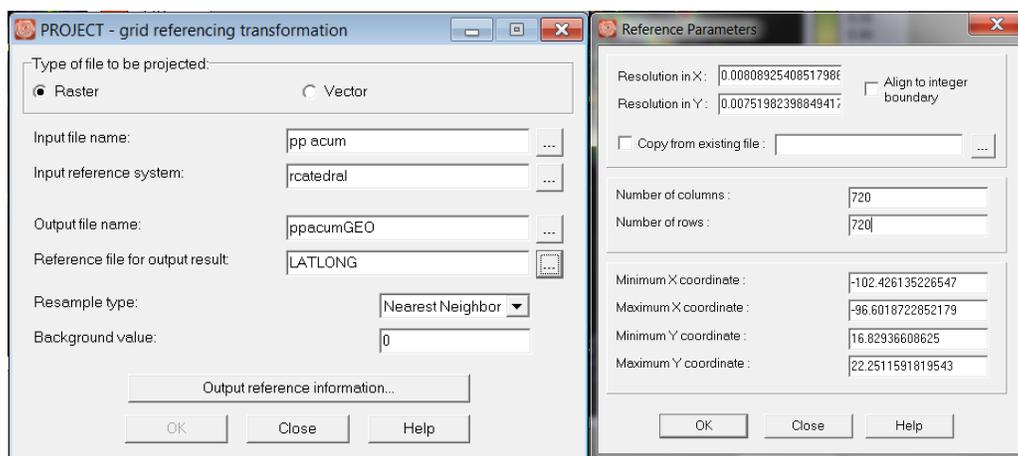


Figura A.3-15: Datos requeridos para cambiar el sistema de referencia de la imagen de lluvia acumulada.

La imagen en coordenadas geográficas se aprecia en la Figura A.3-16.

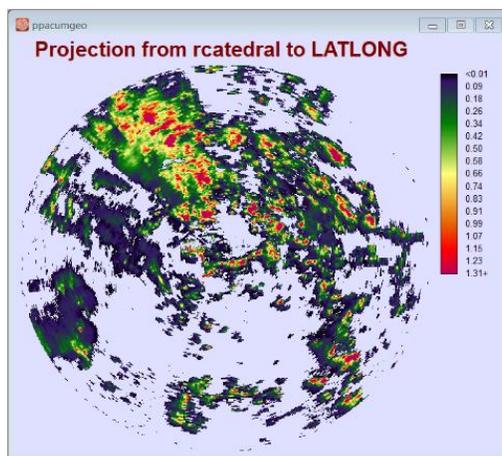


Figura A.3-16: Imagen de radar en coordenadas geográficas.

Para el manejo de la información de la Figura A.3-16, es conveniente el empleo de una plataforma como ArcGIS. Para continuar, es necesario que desde el SIG Idrisi se practique la conversión de la imagen al formato .tif, mediante el comando GeoTiff, ver Figura A.3-17.

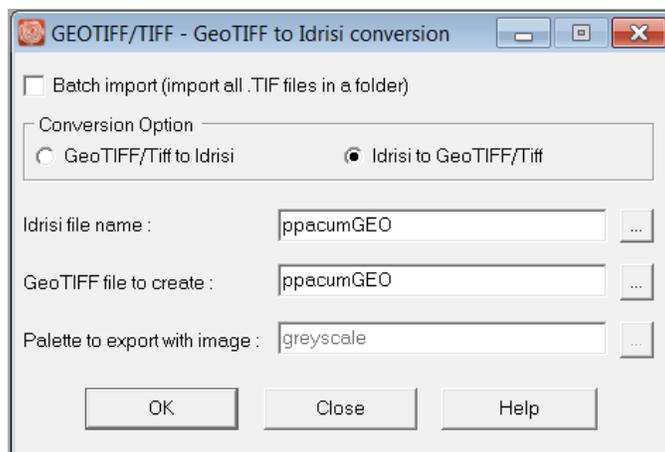


Figura A.3-17: Conversión de imagen a formato .tif, dentro del SIG Idrisi.

Finalmente, la Figura A.3-18 es una representación de la lluvia acumulada en un día, en formato .tif.

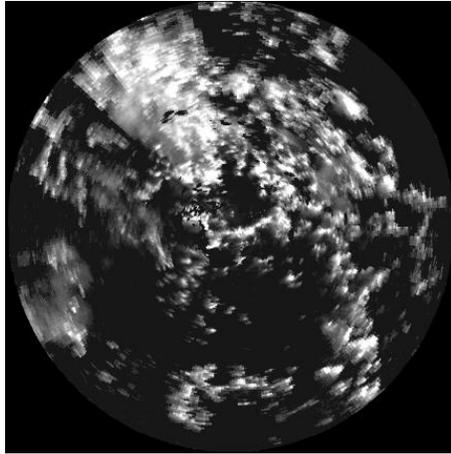


Figura A.3-18: Imagen de lluvia acumulada en un día, en formato .tif.

## Anexo 4. Tablas

Tabla A.4.1: Registros pluviométricos, eventos de 2008 en CDMX, en mm.

NOMBRE	CLAVE	COORD.	COORD.	Precipitación, en mm				
		X UTM	Y UTM	Jul 17	Ago 2	Ago 5	Ago 7	Ago 25
CALVARIO 61	9004	484319.4	2123302.1	0.0	0.0	31.1	3.3	8.4
COLONIA AMERICA	9010	478827.4	2146483.1	2.5	0.0	23.6	7.3	11.7
COLONIA SANTA URSULA COAPA	9014	484445.6	2134397.8	2.0	24.0	9.4	36.0	20.0
DESVIACION ALTA AL PEDREGAL	9020	480855.4	2133694.3	4.0	13.5	43.0	2.5	8.5
EL GUARDA	9022	481800.6	2115712.6	6.0	23.0	24.0	4.0	0.0
GRAN CANAL KM. 06+250	9029	490409.0	2153573.5	14.4	3.5	3.3	21.5	45.6
PLAYA CALETA 454 COLONIA MARTE	9036	489733.4	2144567.9	9.1	7.0	19.7	56.0	0.0
SAN FRANCISCO TLALNEPANTLA	9041	486479.7	2122593.4	6.0	14.0	16.0	4.0	7.0
SAN JUAN DE ARAGON	9043	491691.1	2152312.6	11.4	3.4*	9.2	12.0	31.5
SANTA ANA TLACOTENCO	9045	499708.0	2120621.3	0.4	15.0	14.0	3.0	0.0
CENTRAL TACUBAYA	9048	479438.8	2145529.5		6.6	14.5	15.5	
TLAHUAC	9051	499620.5	2129903.7	2.0	12.0	21.0	8.0	3.0
PUENTE LA LLAVE	9068	494459.5	2148315.8	0.0	9.3	3.5	0.5	5.6
COLONIA EDUCACION	9071	486111.6	2137777.5	5.7*		9.0	0.0	18.4
CHALMITA	1	484734.2	2162357.4	3.6*	3.6	16.5	19.6*	35.1*
REMEDIOS	2	483693.6	2157820.2	6.1	2.5	6.6	11.9	27.4
LINDAVISTA	3	487313.6	2155048.6	13.7	4.1	2.3	18.8	26.7
COYOL	4	490543.5	2154466.7	5.8	1.3		21.6	38.4
GENERADORA 101	5	489600.4	2151595.1	17.3		16.3	15.0	24.9
ROSARIO	6	479971.6	2157274.5	2.0	3.8	30.5	9.1	57.4
CAMPAMENTO MECOAYA	7	480806.4	2154719.0	4.1	4.3	16.8	19.3	1.8
NUEVA STA. MARÍA	8	482626.1	2152769.9	2.5	3.8	2.5	3.0	5.8
SAN JOAQUÍN	9	476038.2	2149181.3	2.0	5.3	11.7	10.4*	72.1
TIZOC	10	482083.5	2149483.6	7.1*	1.3*	16.3	17.5*	26.4*
GDF	11	486071.1	2148578.4	4.3	1.0	13.7	17.3	9.1
MARCOS CARRILLO	12	485935.8	2145472.8	6.9	0.5	15.2	29.0	7.9
LOPEZ MATEOS	13	492732.9	2147422.4	12.2	1.0	12.2	9.7	3.0
CHURUBUSCO LAGO	14	495019.3	2149762.7	2.5*	2.3	11.2	5.1*	1.0*
TRIANGULO	15	477078.3	2144802.5	3.3	4.3	2.3	4.1	16.5
SANTA FE	16	472138.5	2140942.8	3.3*	8.1*	17.0	4.3	66.5*
TANQUE LIENZO	17	476795.9	2138219.2	2.0	4.1	3.0	2.0	21.3
UNIVERSIDAD	18	480972.2	2139317.1	4.3	4.1*	17.3	13.5	29.0
RADIO COMUNICACIÓN	19	482383.4	2145025.7	4.8	1.8	31.8	29.2	15.7
MUNICIPIO LIBRE	20	486584.8	2141295.7	5.8	2.3	13.2	48.0	21.3
CARTERO	21	471518.0	2142222.4	3.0	5.8	18.5	5.3	83.6
LA VENTA	22	467343.0	2137760.7	3.6	16.5	19.1	2.8	29.5

NOMBRE	CLAVE	COORD.	COORD.	Precipitación, en mm				
		X UTM	Y UTM	Jul 17	Ago 2	Ago 5	Ago 7	Ago 25
ZARCO	23	462919.7	2133556.0		25.9	29.0	6.9	32.5
SAN FRANCISCO	24	474874.2	2135343.4	3.6	6.9	44.2	2.8	9.4
RÍO MAGDALENA	25	472197.0	2132804.2	6.9	32.8	25.4	3.6	18.8
MONTE ALEGRE	26	470553.4	2126285.7	25.4	18.3	21.6	14.0	5.6
XOTEPINGO	27	484810.2	2137665.5			12.4	17.8	16.8*
ACULCO	28	489518.8	2142786.5	10.7	3.6	38.1	38.4	11.2
EJERCITO DE ORIENTE	29	496097.1	2142840.1	34.8	5.1	14.2	16.5	4.8
PLANTA CERRO DE LA ESTRELLA	30	491884.0	2138136.5	12.4	8.1	26.4*	15.2	24.4
LA CALDERA	31	502614.1	2138384.9	21.8	18.5	36.1	19.3	
PLANTA SANTA CATARINA	32	497228.8	2134225.1	8.6	0.5	24.9	0.5	
TLÁHUAC	33	498648.1	2132125.0	8.4	15.2	41.4	3.8	
BOSQUE DE TLALPAN	34	479493.5	2133199.1	3.3	5.8*	23.1*	2.8	4.1
VILLA COAPA	35	486698.0	2132554.5	6.6	18.3	13.5	7.9	2.3
SAN PEDRO MARTIR	36	482129.9	2129934.9	4.1*	15.0	34.0	2.8*	4.8*
AJUSCO	37	477857.2	2125094.3	3.3	4.3	33.5	23.1	
TOPILEJO	38	486278.9	2123338.5	1.0	5.6	10.7	2.3	3.3
CASETA FORESTAL (PARRES)	39	481955.0	2116023.4	4.3	19.3	27.2	5.1	6.1
GAVM SUR	40	489627.4	2136066.5	8.9	14.2	25.9	5.6	36.3
NATIVITAS	41	490119.1	2128169.4	4.1	13.0	17.8	2.8	7.4
SAN LUIS TLAXIALTEMALCO	42	496949.8	2129588.7		16.5	45.7		
MILPA ALTA	43	497118.3	2121540.3	9.7	4.6	15.7	6.1	6.9
BARRIENTOS	44	480623.2	2165128.1	5.1	8.9*	7.9*	19.1	39.6
CHICONAUTLA I	46	495681.1	2166561.1		0.5	7.9	30.2	12.7
CHICONAUTLA II	47	488630.1	2157898.2	17.3	2.0	1.3	30.5	58.2
PALMAS	48	473233.4	2148423.7	5.8	6.9	7.4*	11.2	33.8
EL VENADO	49	464872.7	2143319.6	10.2	9.4	14.2	11.7	41.9
100 METROS	50	485652.0	2152929.1	7.4	4.8	10.2*	9.4	28.4
PEÑON	51	491368.9	2149666.1	5.3	1.5	26.7	11.2	7.4
DEPORTIVA	52	489899.7	2146051.2	13.7	2.3	6.6*	19.1	9.9
REBOMBEO LA QUEBRADORA	53	497822.9	2139018.4	45.2*	6.9*	53.8	9.9*	
CAIDA DEL BORRACHO	54	468647.3	2145854.1	6.6	6.1	18.8	15.7	39.4
YAQUI	55	469880.7	2141099.9	2.0	9.9	14.2*	3.8*	51.6
DESIERTO DE LOS LEONES	56	467362.0	2135513.1	4.1	32.5	11.2	5.8	28.4
SAN BARTOLO AMEYALCO	57	471200.6	2137390.9	4.1	14.7	21.1	3.3	
CARCEL DE MUJERES	58	500201.1	2140247.1	20.6	10.2	25.7	12.4	
TETELCO	59	503007.8	2124185.5	1.8	7.1	15.2	9.1	1.3
HUAYATLA	60	472182.2	2134594.9	6.4	7.4	10.2	2.5	6.6
SAN SALVADOR CUAUTENCO	61	490511.9	2121916.0	7.9	11.9	14.2	4.8	5.1
SIERRA DE GUADALUPE	62	487182.9	2164141.0	8.1				

NOMBRE	CLAVE	COORD. X	COORD. Y	Precipitación, en mm				
		UTM	UTM	Jul 17	Ago 2	Ago 5	Ago 7	Ago 25
SANTA ANA TLACOTENCO	63	503040.3	2116570.1	1.0			23.1	
CHAPINGO	64	512041.4	2154600.9	0.3		14.2	3.0	
PRESA MADIN	65	471946.8	2158670.0	26.4			28.4	
LA GRANDE	68	508678.9	2164886.0	1.8		19.1	20.8	
PLANTA CHAPULTEPEC	73	478783.4	2147773.0	2.3	11.9	20.1	13.7	22.4
DEL VALLE	74	482894.5	2142726.9	2.0	1.0	11.9*	26.4*	8.9
ESTADIO AZTECA	75	484026.9	2134087.3	4.6	20.1	19.8	18.8	2.5
TARANGO	77	477663.1	2140777.8	3.8	5.6	13.0	6.6	28.7
PLANTA ABASOLO	78	479646.0	2126786.7	0.3	0.3	0.3	0.3	

\*Lámina de lluvia considerada en el proceso de validación.

Tabla A.4.2: Criterios estadísticos para calibración de parámetros, julio 17 de 2008.

Beta 2, Gamma variable					
Gamma	ME, mm	MAE, mm	RMSE, mm	NSE	CC
0.1	3.153	8.301	10.144	0.501	0.802
<b>0.3</b>	<b>-0.272</b>	4.847	7.026	0.761	0.970
0.4	-0.923	4.442	6.808	0.775	0.973
0.5	-1.324	4.408	6.753	0.779	0.972
0.7	-1.778	4.395	6.768	0.778	0.968
0.9	-2.026	4.401	6.812	0.775	0.964

Gamma 0.3, Beta variable					
Beta	ME, mm	MAE, mm	RMSE, mm	NSE	CC
1	-0.277	4.842	7.026	0.761	0.970
<b>2</b>	<b>-0.272</b>	4.847	7.026	0.761	0.970
3	-0.264	4.855	7.027	0.761	0.970
5	-0.243	4.876	7.031	0.760	0.971
7	-0.225	4.894	7.034	0.760	0.971
9	-0.215	4.904	7.036	0.760	0.971
10	-0.212	4.907	7.036	0.760	0.971

Gamma 0.4, Beta variable					
Beta	ME, mm	MAE, mm	RMSE, mm	NSE	CC
1	-0.927	4.439	6.808	0.775	0.973
2	-0.923	4.442	6.808	0.775	0.973
3	-0.917	4.449	6.809	0.775	0.973
5	-0.896	4.469	6.811	0.775	0.973
7	-0.877	4.488	6.814	0.775	0.973
9	-0.865	4.500	6.816	0.775	0.974
10	-0.861	4.504	6.816	0.775	0.974

Gamma 0.5, Beta variable					
Beta	ME, mm	MAE, mm	RMSE, mm	NSE	CC
1	-1.326	4.406	6.753	0.779	0.971
2	-1.324	4.408	6.753	0.779	0.972
3	-1.319	4.414	6.753	0.779	0.972
5	-1.299	4.433	6.755	0.779	0.972
7	-1.279	4.453	6.758	0.779	0.972
9	-1.266	4.466	6.759	0.778	0.972
10	-1.262	4.470	6.760	0.778	0.972

Tabla A.4.3: Criterios estadísticos para calibración de parámetros, agosto 25 de 2008.

Beta 2, Gamma variable					
Gamma	ME, mm	MAE, mm	RMSE, mm	NSE	CC
0.1	-4.200	12.203	14.968	0.533	0.922
0.3	-2.874	10.043	12.871	0.655	0.907
0.4	-2.629	10.006	12.910	0.653	0.885
0.5	-2.463	10.271	13.046	0.645	0.866
0.7	-2.258	10.632	13.323	0.630	0.840
0.9	-2.138	10.858	13.536	0.618	0.823

Gamma 0.3, Beta variable					
Beta	ME, mm	MAE, mm	RMSE, mm	NSE	CC
1	-2.968	10.137	12.881	0.654	0.910
2	-2.874	10.043	12.871	0.655	0.907
3	-2.790	9.959	12.865	0.655	0.903
5	-2.665	9.876	12.863	0.655	0.897
7	-2.594	9.947	12.865	0.655	0.894
9	-2.558	9.983	12.867	0.655	0.892
10	-2.547	9.994	12.867	0.655	0.892

Gamma 0.4, Beta variable					
Beta	ME, mm	MAE, mm	RMSE, mm	NSE	CC
1	-2.722	9.912	12.903	0.653	0.889
2	-2.629	10.006	12.910	0.653	0.885
3	-2.544	10.090	12.921	0.652	0.881
5	-2.419	10.216	12.942	0.651	0.874
7	-2.348	10.287	12.957	0.650	0.871
9	-2.311	10.323	12.966	0.650	0.869
10	-2.300	10.334	12.969	0.650	0.868

Gamma 0.5, Beta variable					
Beta	ME, mm	MAE, mm	RMSE, mm	NSE	CC
1	-2.556	10.178	13.027	0.646	0.871
2	-2.463	10.271	13.046	0.645	0.866
3	-2.379	10.355	13.067	0.644	0.862
5	-2.253	10.481	13.104	0.642	0.855
7	-2.182	10.552	13.129	0.641	0.851
9	-2.145	10.589	13.142	0.640	0.849
10	-2.134	10.600	13.146	0.640	0.849