



UNIVERSIDAD NACIONAL DE COLOMBIA

VALIDACIÓN DE LA CAPACIDAD DEL MODELO WRF “WEATHER RESEARCH AND FORECASTING” PARA PRONOSTICAR LLUVIA INTENSA, USANDO EL MÉTODO ORIENTADO A OBJETOS Y TABLAS DE CONTINGENCIA

Mauricio Jiménez García

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Geociencias
Bogotá, Colombia

2014

VALIDACIÓN DE LA CAPACIDAD DEL MODELO WRF “WEATHER RESEARCH AND FORECASTING” PARA PRONOSTICAR LLUVIA INTENSA, USANDO EL MÉTODO ORIENTADO A OBJETOS Y TABLAS DE CONTINGENCIA

Mauricio Jiménez García

Tesis de investigación presentada como requisito parcial para optar al título de:
Magister en Ciencias Meteorológica

Director (a):

Ph.D. Gerardo de Jesús Montoya Gaviria

Línea de Investigación:

Predicción numérica del tiempo y el clima

Grupo de Investigación:

Grupo de investigación en meteorología

Universidad Nacional de Colombia

Facultad de Ciencias, Departamento de Geociencias

Bogotá, Colombia

2014

A mis padres

Por estar siempre a mi lado, dándome voces de aliento y llenándome de motivación cada día para alcanzar las metas que me propongo. Y a cada uno de los ángeles que Dios ha puesto día a día en mi camino, quienes con su apoyo incondicional hicieron que este trabajo se hiciera realidad.

Agradecimientos

A la *Fuerza Aérea Colombiana* por darme la confianza y brindarme la oportunidad de seguir creciendo profesionalmente, y además por apoyarme con sus equipos, instalaciones e información meteorológica para desarrollar esta investigación.

Al Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) por brindarme la información meteorológica para el desarrollo de esta investigación.

A los docentes de la Maestría en Meteorología, quienes moldearon con conocimiento mi proyecto académico.

A todos mis compañeros de la Maestría en Meteorología, por todo su apoyo y colaboración.

Resumen

El presente trabajo de investigación analizó la capacidad del modelo WRF "***Weather Research and Forecasting***" para simular precipitaciones intensas en regiones con terreno complejo como Colombia. Para ello, se realizó una comparación entre las técnicas de estadística tradicional (TET) y las técnicas orientadas a objetos (TOO). En las TET, se utilizan las simulaciones de la precipitación generadas por el WRF en puntos de grilla o áreas de 5 X 5 km. Para la TOO, se estudian las precipitaciones más intensas que superen al menos 60 mm/día en áreas de 50 X 50 km. En un área de ese tamaño se encuentran 100 puntos de grilla, donde se pueden presentar fuertes tormentas. Cuando la simulación del WRF tiene similitud con lo observado por la Red Pluviométrica Nacional del IDEAM (RPN), la información puede servir para alertar oportunamente la ocurrencia de precipitaciones intensas que puedan generar inundaciones o desastres naturales.

El estudio fue realizado durante cinco estaciones lluviosas entre los años 2009 y 2013. En este periodo de tiempo se recopilaron 393 datos, que sirvieron para la medición de las precipitaciones máximas. Se evidenció que en zonas bajas se generan resultados satisfactorios, mientras que en regiones elevadas hay menos precisión generalmente por falta de información.

A través de la TOO y la información obtenida en las estaciones se determinó que los mejores resultados se presentaron en el piedemonte llanero, especialmente en los departamentos de Casanare y Meta. Otros lugares con resultados satisfactorios fueron las zonas del valle del Magdalena, entre los departamentos de Caldas, Cundinamarca, Antioquia, Santander y Boyacá, así como el oriente del Tolima. También fue favorable en la parte baja del Chocó, en la Costa Pacífica.

Caso contrario ocurrió en el sur del Chocó sobre la Cordillera Occidental, en los parques Chingaza y Sumapaz, en el suroriente de Cundinamarca, el noroccidente de Antioquia, la Cordillera Oriental en Norte de Santander, Santander y Boyacá. Allí, los resultados del modelo fueron deficientes. Esto no es para menos, pues existen zonas en las que no hay estaciones de la RPN y no hay datos observados para hacer comparaciones. "Por eso, se debe mejorar la red para comprobar la calidad y precisión del modelo, sin ellos no se define qué tan buenos pueden ser".

Palabras clave: Técnicas Orientadas a Objetos, Técnicas de Estadística Tradicional, Validación de modelos, Precipitaciones extremas, modelo WRF.

Abstract

This study evaluates the ability of the WRF "Weather Research and Forecasting" model to simulate heavy precipitations in regions with complex terrain such as Colombia. For this, a comparison between traditional statistical techniques (TET) and object-oriented techniques (TOO) was performed. In TET, WRF simulated precipitation on grid boxes of 5x5 km are used. For TOO, intense precipitations exceeding at least 60 mm / day in areas of 50X50 km are considered. Such area would include 100 grid points. Whenever the WRF output is similar to that observed by the National rainfall network IDEAM (RPN), the timely information could help alert the occurrence of heavy precipitations that may cause flooding or natural disasters.

The study was conducted for five rainy seasons between 2009 and 2013. During this period of time 393 data points were collected, providing maximum precipitation data. Low-elevation areas generated satisfactory results, however less precision is obtained at higher elevated areas, usually due to lack of information.

Based on TOO and station information this study found that the best results occurred in the piedmont plains, especially in the departments of Casanare and Meta. Other places with satisfactory results are the areas of the Magdalena Valley, between the departments

of Caldas, Cundinamarca, Antioquia, Santander and Boyacá and eastern Tolima. Results are also positive at the lower Chocó region, on the Pacific Coast.

On the contrary, WRF forecasts are worse in southern Chocó on the western Cordillera, in the Chingaza and Sumapaz parks, in the south-east of Cundinamarca, the north-west of Antioquia, the Oriental Cordillera in Norte de Santander, Santander and Boyacá. There, the model results were poor, most probably because these areas have no RPN stations and no observed data for comparisons. "Therefore, the network must be improved to ensure the quality and accuracy of the model, without them it is not possible to determine how good they are."

Keywords: Object-OrientedTechniques, Traditional-StatisticalTechniques, Model validation, extreme rain, WRF model.

Contenido

	Pág.
Resumen	IX
Lista de figuras	XV
Lista de tablas	XVII
Introducción	1
1 Antecedentes y Justificación	3
1.1 Planteamiento del problema	7
1.2 Objetivos.....	9
1.2.1 Objetivo General	9
1.2.2 Objetivos Específicos.....	9
2 Proceso Físico de la Precipitación y Modelamiento Numérico	11
2.1 Proceso Físico de la Precipitación	11
2.1.1 Origen de la Precipitación	12
2.2 Procesos Generadores de la Precipitación en Colombia	14
2.2.1 Escala Inter-Decadal.....	15
2.2.2 Escala Interanual	15
2.2.3 Escala Anual	16
2.2.4 Escala Intra-Anual.....	18
2.2.5 Escala Diurna.....	19
2.3 Precipitaciones Extremas	19
2.4 Modelamiento Numérico de la Dinámica Atmosférica	20
2.5 Modelos Mesoescalares	22
2.5.1 Modelo Mesoescalar WRF “ <i>Weather Research and Forecasting</i> ”	23
3 Verificación de Pronósticos y Tablas de Contingencia	25
3.1 Delineamientos Generales.....	25
3.1.1 Qué es la Verificación	25
3.1.2 Propósito de la Verificación	25
3.1.3 Modelos de Verificación	26
3.1.4 Factorización de la Distribución Conjunta de Pronósticos y Observaciones	27
3.1.5 Atributos Escalares del Desempeño de un Pronóstico	29
3.2 Variables Continuas.....	31
3.3 Medidas de Precisión Escalar en Variables Continuas	31
3.4 Pronósticos Categóricos de Predictandos Discretos (Eventos Categóricos) ..	33
3.4.1 Tablas de Contingencia de 2X2	33

3.4.2	Atributos Escalares de las Tablas de Contingencia de 2X2	36
3.5	Pronósticos de Probabilidad	41
4	Técnicas Orientadas a Objetos	43
4.1	Conceptualización de la Técnica de Verificación Orientada a Objetos	43
4.2	Desarrollo Histórico	44
4.2.1	Método de Verificación Basada en Objetos	45
4.2.2	Otras Aproximaciones de la Técnica Orientada a Objetos	49
4.3	Técnica Orientada a Objetos Utilizada en el Presente Trabajo de Tesis	50
5	Metodología.....	52
5.1	Área de Estudio.....	52
5.1.1	Descripción Físico-Geográfica	54
5.1.2	Topografía	54
5.1.3	Hidrografía.....	56
5.1.4	Clima	58
5.2	Análisis y Control de Calidad de los Datos Observados	60
5.2.1	Datos Pluviométricos	60
5.2.2	Análisis y Control de Calidad	62
5.3	Verificación Método de Interpolación.....	67
5.4	Simulaciones del Modelo WRF.....	70
5.4.1	Condiciones Marco	70
5.4.2	Ajustes del Modelo WRF	71
5.4.3	Uso del Modelo WRF en la investigación.....	73
5.5	Metodología de Validación	75
5.5.1	Técnicas de Estadística Tradicional con Variable Continua	76
5.5.2	Técnicas de Estadística Tradicional con Variable Discreta	77
5.5.3	Técnicas Orientadas a Objetos con Variable Discreta	79
6	Análisis de los Resultados.....	82
6.1	Técnicas de Estadística Tradicional con Variable Continua.....	85
6.2	Técnicas de Estadística Tradicional con Variable Discreta.....	88
6.3	Técnicas Orientadas a Objetos con Variable Discreta	93
7	Conclusiones	101
A.	Anexo: Estación con datos de precipitación iguales en 5 días consecutivos .	105
B.	Anexo:NameList de la configuración del modelo WRF de la FAC	113
C.	Anexo: Resultados de las técnicas de validación implementadas	117
	Bibliografía.....	135

Lista de figuras

	Pág.
Figura 1-1: Diferencia entre el promedio de precipitación modelada y observada. (En verde sobrestimación, magenta subestimación, topografía en líneas de contorno cada 200m).....	8
Figura 3-1: Modelo general para la verificación de pronóstico meteorológico.....	27
Figura 3-2: Tabla de contingencia de 2X2, clasificación de eventos dicotómicos	34
Figura 3-3: Tabla de contingencia de 2X2	35
Figura 3-4: Caso de Finley en la predicción de tornados en los EE.UU.....	38
Figura 4-1: Comparación de la precipitación, (A) observada, (B) pronosticada modelo alta resolución, (C) baja resolución.	44
Figura 4-2: Proceso de suavizado de la precipitación.....	47
Figura 4-3: Se resalta la reducción del área del objeto, (A) observado área mayor, (B) modelado área mayor, (C) observado área menor y (D) observado área menor.	51
Figura 5-1: Dominio del modelo WRF de la FAC, con resolución espacial de 5 Km.	53
Figura 5-2: Topografía del área de estudio.....	56
Figura 5-3: Hidrografía del área de estudio	57
Figura 5-4: Promedio anual de la precipitación (izquierda), promedio anual precipitaciones máximas en 24 horas (derecha).....	59
Figura 5-5: Distribución espacial de la estaciones en el dominio del WRF.	62
Figura 5-6: Estaciones con que no cumplen el test de calidad	63
Figura 5-7: Análisis de datos	65
Figura 5-8: Análisis de datos, en círculo rojo los días en los que existe duplicidad de datos en 5 días consecutivos.	66
Figura 5-9: Datos de precipitación atípicos (rojo) y datos no atípicos (verde).	67
Figura 5-10: Análisis iterativo del método de Cressman	69
Figura 5-11: Interpolación de Cressman para el día 30 de junio de 2013.	69
Figura 5-12: Diagrama de flujo para la técnica de estadística tradicional variable continua.	77
Figura 5-13: Diagrama de flujo para la técnica de estadística tradicional variable discreta.	79
Figura 5-14: Diagrama de flujo para la técnicas orientadas a objetos variable discreta. .	81
Figura 6-1: Estaciones dentro del dominio (puntos verdes) y regiones con falencia de datos (círculos rojos).	83

Figura 6-2: Diferencia del promedio de la precipitación modelada y observada, en magenta subestimación, en verde sobrestimación, en círculos rojos aéreas con falacia de datos observados.....	84
Figura 6-3: Resultado BIAS (izq.) y MAE (der.), TET con variable continua	86
Figura 6-4: Resultado MSE (izq.) y RMSE (der.), TET con variable continua.....	87
Figura 6-5: Resultados BIAS, TET con variable discreta.....	90
Figura 6-6: Resultados POD y FAR, TET con variable discreta	91
Figura 6-7: Resultado del CSI, TET con variable discreta	92
Figura 6-8: Resultado del ETS, TET con variable discreta.	93
Figura 6-9: Resultados de las tablas de contingencia con técnicas orientadas a objetos	94
Figura 6-10: Resultado de BIAS, TOO con variable discreta.....	96
Figura 6-11: Resultados POD (izq.) y FAR (der.), TOO con variable discreta.	97
Figura 6-12: Resultados CSI, TOO con variable discreta.	98
Figura 6-13: Resultados ETS, TOO con variable discreta.	99

Lista de tablas

	Pág.
Tabla 5-1: Coordenadas del dominio No. 3 del WRF de la FAC	53
Tabla 5-2: Estaciones meteorológicas usadas en el estudio.....	61
Tabla 5-3: Fechas en las que no se cuenta con información de simulaciones del WRF .	75
Tabla 6-1: Ganancia en exactitud de la TET contra la TOO	100

Introducción

En la actualidad, los efectos del cambio climático han venido alterando drásticamente el bienestar del ser humano dentro del planeta [52]. Debido a sus efectos se ha alterado la frecuencia, intensidad y longitud de muchos eventos meteorológicos extremos (como lluvias torrenciales, sequías, tornados, inundaciones, etc.), causando miles de muertes, e impactando negativamente a la sociedad, al medio ambiente y todas las actividades económicas del estado [5], por ello, se ha generado la necesidad de tomar medidas que permitan adaptarnos a sus efectos, entre ellas, la implementación de modelos numéricos de predicción meteorológica del tiempo, que nos permiten mediante la modelación de algunas variables meteorológicas, generar alertas oportunas ante eventos extremos.

Es así, como en diferentes regiones del mundo se han venido desarrollando modelos numéricos como medidas de prevención, que permiten simular la dinámica atmosférica de mesoescala, pronosticando fenómenos meteorológicos como la precipitación. Actualmente se cuenta con modelos que ajustan las difíciles condiciones topográficas y permiten modelar el comportamiento de la atmósfera de tal manera que se asemejan a las condiciones meteorológicas predominantes, generando aproximaciones en la detección y alertas oportunas de precipitaciones intensas que puedan afectar a las poblaciones o infraestructura.

La efectividad en el uso de modelos en Colombia, está asociada a un modelo que ajuste las complejas condiciones topográficas, debido a que nuestro país está surcado casi que en su totalidad por tres ramificaciones de la cordillera de los Andes. Una herramienta que cuenta con esta capacidad, es el modelo WRF “**Weather Research and Forecasting**”, el cual representa los últimos avances en modelamiento de las variables meteorológicas, y es considerada una herramienta eficiente a la hora de simular los diferentes procesos atmosféricos con buena resolución. Este modelo permite simular la dinámica atmosférica en de mesoescala, resuelve a través de parametrizaciones en radiación, desarrollo de

nubes, capa límite planetaria y otros esquemas físicos, procesos de microescala y mesoescala en la atmósfera a partir de datos globales que ofrecen los modelos de baja resolución [25].

Aunque en la actualidad el modelo WRF es considerado una buena herramienta para generar pronósticos meteorológicos de la precipitación, estudios realizados muestran que algunas regiones pueden estar propensas a una sobre o subestimación[14]. Aunque el WRF predice valores de precipitación comparables con datos reales, éste produce pronósticos de precipitaciones débiles en los valles y muy intensas en las laderas. Los pronósticos son una herramienta de gran utilidad para los pronosticadores del tiempo, por ello resulta indispensable realizar una extenuante verificación de sus simulaciones que permita brindar medidas de exactitud y detección de deficiencias, para que a su vez se realicen las correcciones necesarias que proporcione mejor calidad del pronóstico.

Este proyecto de investigación tiene como fin identificar las regiones en las que el modelo WRF está sobrestimando o subestimando la precipitación, se busca realizar una validación de los datos de precipitación modelados con los datos pluviométricos de las estaciones meteorológicas ubicadas en superficie, implementando una metodología de validación denominada técnica orientada a objetos y tablas de contingencia [27, 10, 21]. Para esto, el dominio es dividido en áreas de 50 X 50 Km (objetos), comparando en cada objeto, el máximo de precipitación observada con la pronosticada, siempre y cuando supere un umbral de 60 mm/día. Estas comparaciones son realizadas por medio de tablas de contingencia.

1 Antecedentes y Justificación

En las diferentes regiones del mundo las consecuencias del cambio climático son cada vez más evidentes. Los diferentes factores antrópicos han generado cambios drásticos en la variabilidad climática de algunas regiones del planeta [34, 52, 11], esto a su vez ha incrementado la frecuencia e intensidad de muchos eventos extremos, tales como inundaciones, sequías, tormentas y temperaturas extremas.

El incremento de las precipitaciones extremas alrededor del planeta ha traído consigo pérdidas humanas, daños estructurales y grandes pérdidas económicas. Según la Organización para la Cooperación y el Desarrollo Económicos (OCDE), sólo en los Estados Unidos, estos fenómenos pueden llegar a causar 6 mil millones de dólares en daños y matar cerca de 140 personas al año. A nivel mundial se puede observar claramente sus efectos, por ejemplo, a finales de 2011 Tailandia fue azotada por una fuerte inundación provocada por continuas lluvias torrenciales [46], más de 700 personas fallecieron, y alrededor de 13 millones de personas se vieron afectadas. A finales de noviembre de 2008, en Brasil al menos 84 personas murieron y cerca de 54.000 fueron desplazadas del estado brasileño de Santa Catarina como consecuencia de las fuertes lluvias que azotaron al sur del país [20].

En Colombia, se encuentran ejemplos similares, en el 2010, una lluvia histórica inundó a Barrancabermeja, el río de Oro ocasionó la muerte de 25 personas, la desaparición de cinco más y la destrucción de 5.000 casas en Girón y Bucaramanga [8]. El 7 de octubre de 2014, en Ibagué, un torrencial aguacero que se prolongó durante toda la tarde dejando inundaciones, cortes en el suministro de energía, caídas de árboles, vehículos aplastados y grandes pérdidas económicas [29].

Según Stocker [52], es necesario adoptar medidas efectivas y oportunas para enfrentar eventos extremos como los mencionados anteriormente, ya que la información disponible muestra que la ocurrencia de estos eventos, en términos de intensidad y frecuencia viene

en aumento. Una manera de contrarrestar o mitigar los efectos de las lluvias extremas es mediante el uso de los modelos numéricos de predicción del tiempo, ellos permiten identificar con antelación el desarrollo de precipitaciones intensas, para alertar oportunamente y evitar o mitigar daños devastadores como los mencionados.

En diferentes regiones del mundo el uso del modelo WRF se ha vuelto cada vez más frecuente debido a los buenos resultados obtenidos en las predicciones meteorológicas, en algunos casos se utilizó el modelo del WRF para alimentar un modelo hidrológico. Con el paso del tiempo se ha intensificado el uso del WRF entre la comunidad científica, [56] es posible mostrar que el WRF pronostica más acertadamente el número de tormentas intensas que otros modelos, mejorando la predicción del área de lluvia y la localización de los centros de precipitación. Golvez J.M. [24], analizó el uso del modelo WRF en estudios meteorológicos en América del Sur, este modelo fue usado para estudiar la circulación y los procesos que generan tormentas sobre el lago Titicaca, logrando así, una buena representación de los procesos de la mesoescala en la región.

Con el paso del tiempo se ha intensificado el uso del WRF entre la comunidad científica, generando avances y mejoras en distintas versiones del mismo. En varias investigaciones se han realizado comparaciones entre los diferentes modelos usados para pronosticar eventos de precipitación, [17] compara los modelos WRF y BRAMS utilizando anidados de alta resolución, donde partir del análisis de precipitación de todo un periodo, se puede concluir que, mediante el uso de métodos de alta resolución se puede representar máximos de precipitación en casos de precipitación intensa, que no se logra detectar con el modelo de baja resolución. Comparativamente, el WRF mostró mayor habilidad que el BRAMS en la resolución utilizada en todos los umbrales considerados. También, se comparó los resultados de pronóstico del WRF y del modelo ETA, constatando que el primero presenta cierta utilidad y ventajas en la representación de los procesos convectivos, e incluso se encontró una configuración en el WRF que arrojó mejores resultados que los obtenidos con el ETA [58].

Actualmente en Colombia, el modelo WRF ha sido implementado por diferentes Universidades, entre ellas la Universidad Nacional de Colombia, principalmente con fines investigativos. El IDEAM (Instituto de hidrología, Meteorología y Estudios Ambientales),

ha implementado el modelo con fines operativos. Igualmente la Fuerza Aérea Colombiana en el año 2007 puso en funcionamiento el modelo WRF para pronóstico operacional utilizando tres dominios; con resoluciones de 45 Km (sur de Centro América y norte de Sur América), 15 Km (territorio nacional) y 5 Km (Región Andina).

Aunque el WRF modela bastante bien los procesos convectivos de mesoescala y genera predicciones meteorológicas relativamente acertadas, ha mostrado una tendencia a sobrestimar la precipitación, especialmente en regiones con topografía compleja, tal como menciona Colle [14], quien encontró que el modelo MM5 (considerada la versión inicial del WRF) a 12 Km de resolución, sobrestima la precipitación a lado barlovento y subestima a sotavento en regiones montañosas o con laderas empinadas. Lo que genera la necesidad de profundizar en la validación de los resultados arrojados por el WRF en regiones con topografía compleja como Colombia.

Algunos estudios en el país muestran los resultados de la validación del WRF, donde se la validación de los datos de precipitación modelados con el modelo MM5, con datos obtenidos de 408 estaciones meteorológicas del IDEAM durante el mes de octubre del año 2006 [33]. La comparación entre los datos observados y simulados se hizo con el esquema Reisner2 estándar y con el esquema de microfísica mejorado. Los resultados obtenidos en este trabajo mostraron que el esquema mejorado presenta mejor desempeño que el esquema estándar Reisner2, sin embargo, el corto periodo de la muestra de datos usados no permite hacer conclusiones más robustas.

Cáceres [9], realizó simulaciones numéricas de la precipitación con el modelo WRF en una escala espacial de 10 Km, sobre el sur del territorio nacional, la muestra seleccionada para dicho análisis fue el periodo de lluvia del 1 de enero al 31 de mayo de 2007 con los datos de 150 estaciones distribuidas en el área de estudio. La validación de los datos observados con los pronosticados se realizó en tres partes: la validación cualitativa mediante el uso de series temporales de tiempo, la validación cualitativa mediante la interpolación espacial y finalmente la validación cuantitativa mediante el uso de índices de correlación [9]. Aunque, según el autor, la validación cualitativa es satisfactoria, el BIAS indica una sobrestimación de la precipitación y un índice de correlación relativamente bajo (0.24).

Tanto en los dos trabajos mencionados como en la investigación de Uribe [55], se usó para la validación, la técnica de estadística estándar en puntos de grilla, es decir la sobreposición en puntos de grilla, del dato observado contra el pronosticado por el modelo.

Las técnicas estadísticas estándar en puntos de grilla se han convertido en el método de validación más utilizado para los pronósticos cuantitativos de la precipitación. Sin embargo, esta técnica de puntos de grilla falla a medida que aumenta la resolución de la cuadrícula y la convección se resuelve de forma explícita [12]. Este error se produce porque las técnicas penalizan gravemente a los pequeños desplazamientos de elementos convectivos con coeficientes de correlación bajo o negativo, alto error cuadrático medio, y valores estadísticos pobres para categorizar (lluvia / no lluvia). Por esta razón, se han desarrollado métodos alternativos como las Técnicas Orientadas a Objetos [19, 3, 16]. En estas técnicas, los "objetos", que son fácilmente distinguibles por un observador o pronosticador humano, se identifican, tanto en el pronóstico como en las observaciones [16].

Las técnicas orientadas a objetos han sido ampliamente utilizadas en estudios previos. Por ejemplo, fue utilizado para comparar los sistemas convectivos de mesoescala simulados, con parametrización de la convección a 10 Km de resolución, y microfísica explícita a 4 Km de resolución [18]. También fue utilizado en estudios separados con el fin de caracterizar y clasificar los eventos de precipitación [28, 57].

Como se mencionó anteriormente, la validación de las simulaciones del WRF por medio de coeficientes de estadística tradicional en punto de grilla (como sesgo, error absoluto medio, error cuadrático medio, etc.) no muestran buenos resultados debido a que penalizan los pequeños desplazamientos de los sistemas simulados. Es decir si se pronostica acertadamente un sistema, pero tiene un ligero desplazamiento, la estadística tradicional no toma en cuenta este acierto y arroja un no acierto, generando peores resultados. Por el contrario en las técnicas orientadas a objetos se intenta verificar varias regiones o patrones de precipitación, localizando atributos espaciales que caracterizan a esos patrones, tales como: tamaño, orientación, tiempo de vida, propagación, etc. Luego se estima el grado de correspondencia entre el atributo pronosticado y simulado, es de

esta manera en la que a pesar de que existan pequeños desplazamiento en la simulación las técnicas orientas a objetos muestran acierto y por ende mejores resultados.

El propósito del presente trabajo de investigación es presentar soluciones a las falencias que se han presentado en anteriores trabajos [9, 33, 55] realizando un análisis detallado de la precipitación modelada por el WRF en Colombia, empleando la técnica orientada a objetos, lo que permiten obtener una representación más significativa de la distribución espacial de la precipitación.

1.1 Planteamiento del problema

Los modelos numéricos de predicción del tiempo, en particular el WRF, muestran sensibilidad a las condiciones topográficas complejas (la orografía afecta en gran medida los resultados de las simulaciones). Tal como menciona Colle [14], los modelos sobreestiman la precipitación en zonas montañosas, sobreestimando a lado barlovento y subestimando a sotavento. Está sobreestimación sobre el territorio colombiano, puede ser aún más compleja debido a que el mismo es surcado por tres cadenas montañosas y a que la precipitación puede estar asociada con masas de aire provenientes del este sobre la Amazonia y/o del oeste sobre el litoral Pacífico.

Lo mencionado se puede observar en la Figura 1-1, en la que se presenta la diferencia entre el promedio de la precipitación modelada por el WRF y la observada por las estaciones en superficie, notándose la fuerte sobreestimación de la precipitación sobre las cordilleras. Este ejemplo muestra claramente lo mencionado por Colle [14]. De ahí radica la importancia de verificar los resultados que arrojan los modelos de predicción meteorológica sobre territorio colombiano.

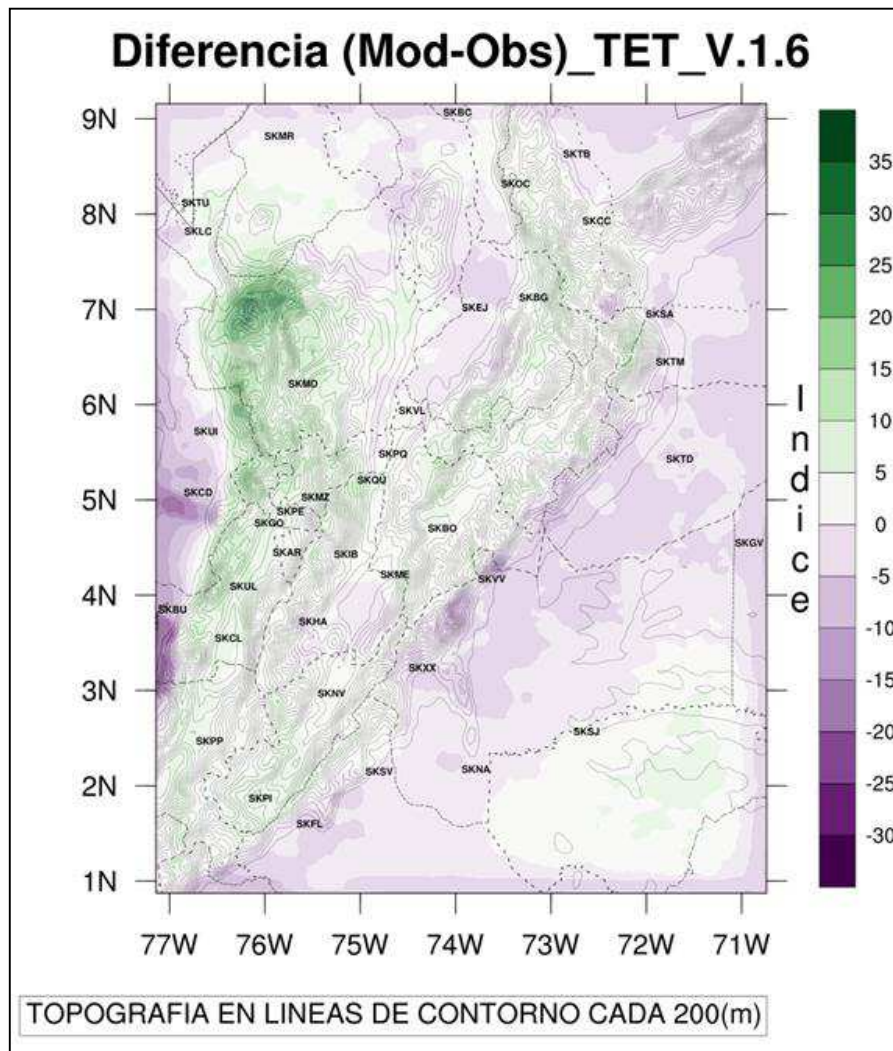


Figura 1-1: Diferencia entre el promedio de precipitación modelada y observada. (En verde sobrestimación, magenta subestimación, topografía en líneas de contorno cada 200m).

La verificación de los Pronósticos Cuantitativos de la Precipitación (*Quantitative Precipitation Forecast* "QPF") usando técnicas de estadística tradicional, penaliza severamente los pequeños desplazamientos de los sistemas simulados, por el contrario las técnicas orientadas a objetos permiten realizar una mejor comparación, que no penaliza estos desplazamientos. Además, la validación mediante técnicas de estadística tradicional, se dificulta a medida que la resolución del modelo se incrementa, impidiendo una adecuada verificación en modelos de mesoescala como el WRF, que corren a una resolución de pocos kilómetros [12].

En este trabajo se busca comparar el método de validación de estadística tradicional en puntos de grilla con la técnica orientada a objetos, para determinar en qué regiones el modelo WRF subestima y sobrestima la precipitación. Surgen entonces las siguientes preguntas:

- ¿Cómo se aplica la técnica orientada a objetos para validar la precipitación?
- ¿Cómo implementar la técnica de estadística tradicional en puntos de grilla para validar la precipitación?
- Al comparar el método orientado a objetos con la estadística tradicional en puntos de grilla, ¿cuál es la ganancia en exactitud del uno contra el otro?
- Y por último, ¿en qué regiones el modelo WRF sobreestima y en qué regiones subestima la precipitación?

1.2 Objetivos

1.2.1 Objetivo General

Validar la capacidad del modelo WRF “*Weather Research And Forecasting*” para pronosticar lluvia intensa, determinando en qué regiones el modelo sobrestima y en qué regiones subestima la precipitación, mediante el uso técnicas orientadas a objetos y tablas de contingencia.

1.2.2 Objetivos Específicos

- Analizar los resultados de la implementación de la técnica orientada a objetos para validar la precipitación modelada por el WRF.
- Analizar los resultados de la implementación de la técnica de estadística tradicional en puntos de grilla para validar la precipitación modelada por el WRF.

- Determinar la ganancia en exactitud del método orientado a objetos y la estadística tradicional en puntos de grilla.

- Determinar las regiones en las cuales el modelo WRF sobreestima o subestima la precipitación.

2 Proceso Físico de la Precipitación y Modelamiento Numérico

El estado físico de la atmosfera puede describirse mediante datos meteorológicos, que pueden ser variables o fenómenos que caracterizan las condiciones de la atmosfera en un lugar y tiempo dados (por ejemplo: temperatura, presión, viento, humedad, precipitación, etc.) [44].

Cuando se quiere predecir el estado de la atmosfera en un instante futuro, es necesario estudiar no solo dichas variables o fenómenos, sino también los procesos físicos que en ella ocurren. Por ello, en el siguiente capítulo se realiza una descripción de los procesos físicos que generan la precipitación y los sistemas que permiten su desarrollo en Colombia. Además, se realiza una breve descripción del modelamiento numérico de la dinámica atmosférica, como preámbulo al estudio de los métodos de verificación con los que se evaluarán las precipitaciones generadas por el modelo de mesoescala WRF.

2.1 Proceso Físico de la Precipitación

En la atmosfera ocurren procesos físicos a nivel microescalar. El agua se halla en las primeras capas de la atmosfera, pero generalmente en forma de vapor, imposible de distinguir a nuestra vista. Solo llega a ser visible por la formación de nubes, cuya presencia y forma suministran frecuentemente detalles sobre la evolución del tiempo que se nos aproxima.

La formación de la mayoría de las nubes resulta del movimiento ascendente del aire húmedo que se expande a causa de la disminución de la presión con la altitud y por el consiguiente enfriamiento adiabático. Es en este momento en el que una parte del vapor de agua se condensa para formar la nube [44].

La condensación se inicia con el proceso de nucleación (formación de un núcleo de 0.1 - 1.0 μm en diámetro) el cual continúa creciendo por adherencia del vapor de agua sobre él, hasta la formación de una gota de nube o cristal de hielo estable. Si las condiciones dinámicas lo permiten, las gotas de nube continúan creciendo por colisión y coalescencia, hasta formar gotas de lluvia, cristales de hielo, granizo, etc. [45]. Los procesos que conducen a la formación de nubes y precipitación tales como; nucleación, crecimiento por condensación y coalescencia, ruptura espontánea, por colisión, etc., se denominan *procesos microfísicos* mientras que los que tienen que ver con el movimiento del aire (corrientes ascendentes, estabilidad, etc.) que favorece la formación de nubes y de lluvia se llaman *procesos dinámicos*[51].

Para que la condensación tenga lugar, el aire debe estar cercano a su punto de saturación, es decir cercano a una humedad relativa del cien por ciento, el aire se satura de dos formas: adicionando vapor de agua a través de evaporación, o enfriando el aire hasta alcanzar el punto de rocío. El proceso de enfriamiento en la atmósfera recibe el nombre de *ascenso convectivo*.

La lluvia se forma dentro de la nube en la activación y sucesivo crecimiento de núcleos de condensación. La precipitación se forma cuando algunas gotas o partículas de hielo de una nube originalmente monodispersa, aumentan de tamaño, vencen la resistencia del aire, comienzan a caer debido a la fuerza de gravedad, y finalmente se sedimentan con una velocidad terminal en la superficie de la tierra.

2.1.1 Origen de la Precipitación

Existen dos mecanismos que dan origen a la precipitación: La formación de gotas de lluvia por coalescencia entre ellas mismas y el crecimiento de cristales de hielo por el proceso llamado mecanismo de Wegener-Bergeron-Findeisen [63], ambos procesos pueden actuar simultáneamente o formar precipitación de manera independiente.

- **Crecimiento por Coalescencia** En este proceso, la lluvia se forma por coalescencia (posibilidad de que dos o más gotas de nube se unan en un único cuerpo) entre gotas de mayor tamaño con gotas de menor tamaño dentro de la

nube. El mecanismo por el cual, dentro de una nube inicialmente monodispersa aparecen algunas gotas privilegiadas de mayor tamaño, es aún materia de discusión. Se supone que algunas de estas gotas pudieron tener su origen en núcleos súper gigantes de condensación. La formación de gotas de precipitación por choques estocásticos (no deterministas) entre gotas de nube, causados por la turbulencia, es otro de los mecanismos que se propone para explicar la ampliación del espectro original de gotas de nube. Una vez algunas gotas de precipitación se han formado, éstas continúan creciendo por colisión y coalescencia con otras de menor tamaño, es decir unas crecen a expensas de las otras.

- **Proceso Wegener-Bergeron-Findeisen** A medida que la nube continúa el ascenso, su tope puede sobrepasar la isoterma de 0°C , dando lugar a una nube mixta, es decir; una nube que contiene tanto una fase líquida como sólida. No todas las gotas de una nube se congelan al pasar la isoterma de 0°C . Las gotas que a pesar de encontrarse en un ambiente con temperatura por debajo de cero grados, continúan existiendo en la fase líquida, reciben el nombre de gotas sobreenfriadas "supercooled". Es común encontrar en una nube, gotas sobre enfriadas a temperaturas de -15°C e inferiores. Cuando en las proximidades de cristales de hielo coexisten gotas sobre-enfriadas, el vapor de agua sobre estas últimas migra hacia los cristales de hielo, por encontrar allí una tensión de saturación menor. Esto conlleva a que el cristal de hielo crezca más rápido a expensas de la gota sobre enfriada. En esto consiste, el crecimiento inicial de cristales de hielo en nubes mixtas y que dan inicio a la precipitación por el proceso de Wegener-Bergeron-Findeisen. Una vez los cristales de hielo han adquirido un tamaño suficiente, continúan creciendo por colisión con otros cristales y gotas sobre-enfriadas [63].
- **Influencia de los Aerosoles en la Precipitación** La mayoría de los aerosoles en la atmósfera tienen la capacidad de actuar como núcleos de condensación (NC). La cantidad y el tamaño de los NC, tiene un gran impacto en la precipitación. En general, el aire continental tiene más alta concentración de NC que el marítimo y por eso, las nubes continentales tienen mayor concentración de gotas que las marítimas [54]. Para una misma cantidad de agua líquida (relación de mezcla) la

existencia de una gran cantidad de aerosoles que sirven de NC, conducen a la formación de un espectro de gotas pequeñas, mientras que pocos NC conducen a un espectro de pocas, pero gotas de mayor tamaño. La relación entre el número de NC y el tamaño de las gotas se explica de la siguiente manera:

De acuerdo al tercer momento de la distribución de gotas por tamaños, La relación entre el contenido de agua líquida q_l , el radio medio R y la concentración o número de gotas por unidad de volumen N se escribe:

$$q_l = \frac{\rho_w}{\rho} \frac{4}{3} \pi R^3 N \quad (2.1)$$

donde, ρ es densidad del aire, ρ_w es la densidad del agua. Observe que la relación entre el radio medio y la concentración es inversa. Entonces, si q_l permanece constante; para una disminución de N , R aumenta y viceversa.

Esto significa que las gotas de nubes marítimas son más propicias a crecer por coalescencia que las de nubes continentales en un mismo intervalo de tiempo. Cambios en el número y tamaño de los NC puede reforzar o por el contrario inhibir la formación de precipitación. La precipitación puede ser inhibida donde los NC son abundantes como en la ciudades, incendios forestales como en la amazonia, mientras que la adición de pocos pero grandes NC puede reforzar la precipitación [2].

2.2 Procesos Generadores de la Precipitación en Colombia

Los procesos generadores de la precipitación en Colombia, son considerados sistemas que van desde la escala inter-decadal hasta la escala diurna, por esta razón se toma referencia la investigación realizada por Poveda [41], en donde se realiza un análisis climatológico de la precipitación en Colombia.

Colombia está situado en el trópico Americano, “Como una isla entre tres océanos” [49], teniendo en cuenta que además del Caribe y del Pacífico, Colombiano recibe la

influencia de la circulación atmosférica de la cuenca Amazónica. La presencia de la barrera orográfica que constituyen los tres ramales de la cordillera de los Andes, induce la formación de climas locales y regionales de alta complejidad y dificultad de predicción a distintas escalas de tiempo. Entre los Andes y la cuenca Amazónica existe una retroalimentación positiva, debido a que los vientos alisios del sureste transportan gran cantidad de humedad hacia los Andes, contribuyendo a la precipitación sobre el piedemonte Andino. A continuación se realiza una síntesis de los elementos y mecanismos físicos de la variabilidad climática de Colombia.

2.2.1 Escala Inter-Decadal

A escala inter-decadal la precipitación en Colombia está dominada por efecto del cambio climático. En Colombia los resultados muestran que las tendencias de las series de precipitación son positivas en algunas regiones y negativas en otras, y además se han identificado cambios en la fase y la amplitud del ciclo anual y semi-anual, es decir corrimiento de las temporadas lluviosas dentro del año, con importantes implicaciones en la agricultura y en generación de energía hidro-eléctrica [37].

Además, en esta escala encontramos la llamada Oscilación Década del Pacífico (PDO) y la Oscilación del Atlántico Norte (NAO), que tienen una influencia significativa sobre el clima en Colombia [37].

2.2.2 Escala Interanual

La escala interanual está fuertemente controlada por eventos El Niño/Oscilación del Sur (ENSO, por sus siglas en inglés), la Oscilación del Atlántico Norte (NAO) y Oscilaciones Cuasi-Bienales de los vientos estratosféricos (QBO). Generalmente durante El Niño se presenta una disminución en la precipitación en Colombia, así como una disminución en la humedad del suelo y en la actividad vegetal. En general los efectos son más fuertes y más inmediatos sobre el occidente que sobre el oriente del país [43]. Durante La Niña ocurren anomalías contrarias, con eventos de precipitación muy intensos que generan crecidas de los ríos, avalanchas e inundaciones de planicies aluviales.

Los mecanismos físicos por los cuales se presenta déficits en las precipitaciones en Colombia durante El Niño, se pueden resumir como [42]:

- Debilitamiento de la corriente en chorro del Chocó y de la corriente en chorro ecuatorial de los 700hPa.
- Reducción en número e intensidad de las ondas tropicales del este sobre el Atlántico tropical norte.
- Desplazamiento de la ZCIT “Zona de Convergencia Intertropical” hacia el sur-este de su posición normal, como consecuencia del establecimiento de una celda de Handley anómala (en posición y sentido de circulación) sobre la región, que presenta movimiento descendentes sobre el trópico americano.
- Alteración del balance de humedad atmosférico, en particular en los niveles bajos de la atmosfera, ya que se presenta déficit de humedad de suelo durante El Niño en Colombia [41].

Según los autores [64], al parecer no existe relación de las Oscilaciones Cuasi-Bienales de los vientos estratosféricos (QBO) con las precipitaciones en Colombia a pesar de la fuerte componente cuasi-bienal existente en los registros de lluvia del país.

2.2.3 Escala Anual

La escala anual se encuentra dominada por fenómenos físicos como la migración de la ZCIT “Zona de Convergencia Intertropical” asociada a la dinámica trans-ecuatorial de advección de humedad por los vientos alisios del este, la actividad de chorro del Choco y los sistemas convectivos de mesoescala [41].

- **Migración de la Zcit:** En Colombia se presentan dos temporadas lluviosas (Abril-Mayo y Octubre-Noviembre), y dos secas (Diciembre-Enero y Junio-Agosto) como resultado del doble paso de la ZCIT en el territorio. En lugares extremos de la oscilación de la ZCIT, el ciclo es uni-modal. Su posición media se encuentra

desplazada alrededor de los 5° y 10°N [40], la ZCIT constituye el lugar preferente para la convergencia superficial de los vientos alisios, y es una región preferente de bajas presiones atmosféricas superficiales, ascenso de vientos húmedos, desarrollo de mecanismos de convección profunda, alta nubosidad u divergencia en la alta atmosfera[41].

- **El Chorro del Chocó:** La distribución de la precipitación anual sobre la región tropical de América del Sur no puede ser explicada únicamente en términos de la advección de humedad por los vientos alisios del este, desde el Caribe y la cuenca Amazónica. La corriente de vientos en chorro del occidente de Colombia, o chorro del Chocó transporta grandes cantidades de humedad desde el Océano Pacífico hacia el interior del país. El chorro del Chocó interactúa con la topografía de los Andes occidentales y con los vientos alisios del este para favorecer la convección profunda y que contribuye a producir grandes cantidades de precipitación sobre el occidente del país. El chorro del Chocó exhibe un marcado ciclo anual, siendo más fuertes en los meses de septiembre-noviembre y más débiles en los meses de febrero-marzo, ello conduce a explicar por qué la segunda temporada lluviosa septiembre-noviembre es más intensa que la primera abril-mayo en muchas partes del occidente de Colombia [41]. Su comportamiento durante un evento El Niño es de debilitamiento y se intensifica durante La Niña [41].
- **Sistemas Convectivos de Meso-Escala:** Muchos de los eventos más intensos de precipitación en Colombia se desarrollan en Sistemas Convectivos de Meso-escala (SCM). Una de las zonas cuyo ambiente atmosférico es más favorable para la formación de eventos con SCM está al este del Pacífico y sobre la Costa Pacífica Colombiana. Sobre la región Andina en Colombia, el valle del Río Magdalena y el Norte de Antioquia presentan la mayor cantidad de SCM. La cuenca Amazónica y la región del Orinoco en Colombia también presentan altas frecuencias de SCM. Las zonas con menos densidad de eventos están al norte de Colombia y Venezuela y al occidente de Perú [41].
- **Balance de Humedad Atmosférica:** El análisis de flujo de humedad que entra y sale de Colombia se realizó mediante un volumen de control definido entre 77,5°O

y 67,5°O y entre 2,5°S y 10°N, y en la vertical de 1000 hPa a 300 hPa, el balance neto de humedad es de 1914 mm \pm 131 mm, la precipitación promedia es de 2898 mm/año. También se encontró que el mayor transporte de humedad hacia Colombia se da en los trimestres diciembre-febrero y marzo-mayo con valores promedios de 607 y 505 mm/año [41].

2.2.4 Escala Intra-Anual

Los fenómenos que mayor influencia tienen a esta escala son las oscilaciones de 30-60 días, las ondas tropicales del este, los huracanes sobre el Caribe y el Pacífico oriental norte, y en las épocas de invierno en el hemisferio sur cuando se presentan frentes fríos que transportan grandes contenidos de humedad y pueden alcanzar la geografía Colombiana. Además, los fenómenos de la interacción suelo-atmósfera a través del control de la evaporación por la humedad del suelo y las precipitaciones recicladas.

- **Oscilaciones de 30-60 Días:** Las variaciones intra-anual asociadas con la oscilación Madden-Julian, también llamadas oscilación intra-anual, exhiben variabilidad en la escala de tiempo intra-estacional, con periodos de 30-60 días, dominante en la zona trópic y asociada con la celda de circulación orientada en sentido zonal y con propagación hacia el este. Las anomalías en forma de ondas pueden ser trazadas desde el Océano Índico al Pacífico central y vía Hawái hasta América Central y el norte de América Latina [41]. Se ha encontrado que la fase de la oscilación Madden-Julian afecta fuertemente en ciclo diurno de las precipitaciones en Colombia, durante su fase convergente, favorece las precipitaciones y por el contrario, la fase subsiguiente las inhibe [41].
- **Ondas Tropicales del Este:** Son perturbaciones de verano y otoño (junio-noviembre), sobre el Atlántico tropical norte. Se trata de ondas atmosféricas con circulación ciclónica, las cuales se desplazan desde el occidente de África hacia el Atlántico, con un periodo de 4-8 días y una longitud de onda de 2000 km. Las ondas del este constituyen la génesis de los huracanes sobre el Atlántico Norte y el mar Caribe, se reducen durante El Niño. Estas ondas pasan por la geografía Colombiana en su camino hacia el Océano Pacífico, y constituyen una fuente

importante de variabilidad climática y de precipitaciones, en particular en la zona de la costa del Caribe, el valle del río Magdalena y el oriente del país [41].

2.2.5 Escala Diurna

El ciclo diurno de temperaturas, resultante del ciclo de insolación superficial es un factor fundamental en la formación, desarrollo e intensidad de la convección profunda, mecanismo importante en la génesis de la formación lluvia tropical. En Colombia, el ciclo diurno sigue siendo investigado, pero mediante aproximaciones preliminares se puede concluir: el ciclo diurno de las lluvias en Colombia es altamente variable en el espacio, las estaciones exhiben picos uni-modales o bi-modales (semi-diurnos), con una influencia muy fuerte de factores locales, el periodo comprendido entre las 08:00 a.m. y las 11:00 a.m., es el de menor probabilidad de lluvia en todos los Andes de Colombia, las estaciones localizadas sobre la vertiente occidental de la cordillera central presentan ciclo diurno uni-modal con precipitaciones en la madrugada, el ciclo diurno es variable durante los distintos meses del año, de manera que la uni-modalidad en una estación de registro durante ciertas épocas del año puede convertirse en bi-modalidad durante otras épocas. Los mecanismos físicos que explican tal comportamiento variable del ciclo diurno tienen que ver con el ciclo diurno de insolación, con los procesos de la interacción suelo-atmosfera, con la dinámica diurna de los Sistemas Convectivos de Meso-Escala, y con la circulación de los vientos locales y de escala sinóptica, dentro y sobre los valles y vertientes Andinas [41].

2.3 Precipitaciones Extremas

Según [5] las precipitaciones extremas son episodios lluviosos de gran intensidad, capaces de desencadenar procesos geomorfológicos de gran magnitud e importantes consecuencias en pérdidas económicas y humanas, estos efectos devastadores son mayores en áreas de montaña, de gran energía topográfica e inestabilidad geomorfológica. Se caracterizan por presentar un sistema acoplado de convergencia de vientos en superficie, divergencia de vientos en capas medias y altas, e intensificación de movimiento vertical ascendente. Este escenario atmosférico en presencia de altos porcentajes de humedad y poca cizalladura del viento, facilita la generación de nubes de

gran desarrollo vertical [47]. En Colombia la complejidad orográfica favorece el desarrollo de tormentas convectivas de gran intensidad, apoyados por los diferentes procesos (mencionados en la sección 2.2) que favorecen e intensifican aún más las precipitaciones sobre algunas regiones del país.

2.4 Modelamiento Numérico de la Dinámica Atmosférica

La predicción meteorológica consiste en determinar anticipadamente los valores de una serie de variables meteorológicas, como temperatura, humedad, presión, dirección e intensidad del viento, precipitación, etc., que pueden llegar a afectar a determinada región. Actualmente las predicciones se realizan utilizando modelos de predicción numérica que emplean las ecuaciones de las leyes físicas que describen el comportamiento de la atmósfera. La solución de estas ecuaciones es algo compleja, debido a que no existe una solución matemática exacta que permita determinar los valores futuros de las variables, por tal razón es necesario recurrir a las técnicas de modelamiento numérico [25].

Formalmente, los modelos de predicción meteorológica se basan en resolver de forma numérica estas ecuaciones, para ello recurren a realizar aproximaciones a un conjunto de ecuaciones diferenciales referidas a la conservación de la masa, la energía y el momento lineal en la atmósfera. Partiendo de condiciones iniciales obtenidas mediante un proceso de asimilación de datos, a partir de un modelo de circulación general y un conjunto de observaciones en un determinado momento de tiempo (t), el modelo resuelve las ecuaciones para encontrar valores a las variables pronosticadas en un tiempo futuro ($t + \Delta t$). Para ello es necesario hacer uso de las ecuaciones en derivadas parciales, que traducen las leyes generales de la física que rigen la dinámica de la atmósfera terrestre [25].

La atmósfera es un fluido en movimiento sometido a procesos de intercambio de energía. En consecuencia, se necesita describir adecuadamente el movimiento de las partículas de aire sometidas a un conjunto de fuerzas e introducir el balance energético una vez que se tiene en cuenta los intercambios que tienen lugar en el medio atmosférico, tanto en el medio como entre la atmósfera y la superficie terrestre. Para describir su

comportamiento se utilizan las ecuaciones generales de la mecánica de fluidos y de la termodinámica. Además, el movimiento de las partículas de fluidos atmosféricos debe ser estudiado teniendo en cuenta que nuestro planeta es un sistema de referencia no inercial. Por lo tanto, para que se verifiquen las leyes de la mecánica clásica, es necesario introducir algunos términos denominados fuerzas aparentes [25].

El movimiento de una partícula de fluido atmosférico viene dada por:

$$\frac{d\vec{V}}{dt} = \sum_i \frac{\vec{F}_i}{m} = -\frac{1}{\rho} \vec{\nabla} p + \vec{g} - 2\vec{\Omega} * \vec{V} + \vec{F}_r \quad (2.2)$$

donde,

$\left(\frac{d\vec{V}}{dt}\right) \rightarrow$ Aceleración por unidad de masa

$\left(\sum_i \frac{\vec{F}_i}{m}\right) \rightarrow$ Fuerza por unidad de masa

$\left(-\frac{1}{\rho} \vec{\nabla} p\right) \rightarrow$ Gradiente de presión

$\vec{g} \rightarrow$ Gravedad

$\left(-2\vec{\Omega} * \vec{V}\right) \rightarrow$ Fuerza de Coriolis

$\vec{F}_r \rightarrow$ Fuerza de rozamiento

El término de la izquierda representa la fuerza por unidad de masa que actúa sobre una burbuja de aire, medida desde un sistema de referencia no inercial (la Tierra). Los términos de la derecha señalan dicha burbuja: la fuerza debido al gradiente de presión, resultante de la fuerza gravitatoria y céntrica, la fuerza de Coriolis, que junto a la anterior, son los dos términos ficticios de la ecuación, y el término asociado a la fuerza de rozamiento. Para resolver esta ecuación es necesario introducir una serie de simplificaciones en función de las escalas espacio-temporal utilizado en cada caso [25].

Desde el punto de vista termodinámico hay que tener en cuenta los intercambios de calor asociados a los cambios de fases del agua (vapor, líquido y estado sólido) que afectan a la formación y desarrollo de las nubes y a los intercambio de energía con la superficie de

la tierra, los océanos y la troposfera. Otro factor a tener en cuenta es el hecho de que la atmósfera es un fluido gaseoso que sufre variaciones de densidad como consecuencia de la dinámica a la que se encuentra sometida [25].

Las ecuaciones matemáticas que describen todos estos procesos no son lineales y, por tanto, las soluciones no se pueden obtener de forma analítica. Para resolverlas es preciso recurrir al cálculo de numérico, que proporciona una solución aproximada. Para obtener estas soluciones aproximadas, el primer paso consiste en discretizar el área geográfica objeto de estudio. Para ello se define una rejilla (*grid*) tridimensional que la abarque. La distancia entre dos puntos de la rejilla se conoce como resolución del modelo. El siguiente paso consiste en resolver las ecuaciones, utilizando cálculo numérico, en todos los puntos de la grilla [25].

2.5 Modelos Mesoescalares

Los modelos mesoescalares son aquellos que permiten estudiar y pronosticar fenómenos meteorológicos, incluyendo los mesoescalares, es decir, aquellos que ocurren a una escala espacial relativamente pequeña. Entre los primeros modelos se encuentran los hidrostáticos pues consideraban despreciables los movimientos en la escala vertical frente a los movimientos en la horizontal. Esto ocurre con modelos que simulan a escala sinóptica, válidos para establecer el comportamiento de las masas de aire, frentes polares, etc. En estos casos se mantiene la condición hidrostática y la presión está completamente determinada por las masas de aire extendidas, quedando inhibidos los movimientos del aire en la vertical en las hipótesis iniciales. Sin embargo, cuando la escala de la rejilla del modelo se reduce, la resolución espacial aumenta y es posible incorporar movimientos en la vertical, y por tanto, la hipótesis hidrostática, deja verificarse. Se puede decir que esto permite incorporar algunos factores meteorológicos como la convección, ondas, etc. que son de tipo mesoescalar [25].

Entre tanto, hay dos grandes tipos de modelos numéricos: hidrostáticos, como el *High Resolution Limited Area Model* (HIRLAM), y no hidrostáticos, como el *Mesoscale Atmospheric Simulation System* (MASS), *Regional Atmospheric Modeling System* (RAMS), *PennState-NCAR* (MM5), o la versión avanzada de este último denominada

Weather Research and Forecasting (WRF). Actualmente, los modelos mesoescalares de área limitada permiten realizar simulaciones de la atmósfera bajo condiciones no hidrostáticas, y resultan muy útiles para el estudio de fenómenos de precipitaciones intensas y severas [25].

Todos los modelos de simulación mesoescalares hacen uso de distintos esquemas de parametrizaciones para resolver aquellos fenómenos cuya escala espacial es inferior a la resolución del modelo, entre ellos, parametrización de cúmulos, microfísica, capa límite, etc. [25].

2.5.1 Modelo Mesoescalar WRF “*Weather Research and Forecasting*”

El modelo de mesoescala WRF es considerado un modelo de última generación, diseñado para ser aplicado tanto en pronósticos operativos del tiempo, como en investigación de fenómenos meteorológicos, permitiendo integrar a la comunidad científica y los centros de pronóstico de todo el mundo. Es un modelo de simulación de la atmósfera de área limitada, no hidrostático, sensible a las características del terreno y diseñado para predecir la circulación atmosférica a escala sinóptica, mesoescalar y regional. Resuelve a través de parametrizaciones en radiación, desarrollo de nubes, capa límite planetaria y otros esquemas físicos, procesos de la micro y mesoescala en la atmósfera a partir de datos globales que ofrecen los modelos de baja resolución. Como ya se mencionó, es un modelo no hidrostático, lo que le permite ser usado en la escala de unos pocos kilómetros, aunque también permite correr bajo condiciones hidrostáticas. Actualmente es usado en diferentes regiones del mundo por ser un modelo robusto y de fácil acople a las diferentes condiciones de cada región, tales como topografía, vegetación, superficies de agua y uso del suelo. El WRF ha sido desarrollado en colaboración entre, *The National Oceanic and Atmospheric Administration* (NOAA), *The National Center for Atmospheric Research* (NCAR) y otras organizaciones [25].

Tal como menciona [25], cuenta con diferentes módulos que se dividen distintas funcionalidades:

GEOGRID: Permite configurar el área geográfica del área de estudio, tomando en cuenta topografía, usos del suelo y sus características físicas en función de la época del año.

UNGRIB: Este módulo prepara los datos de inicialización de modelo y las condiciones de contorno, para ello toma datos de un modelo que tenga mayor cobertura espacial como el GFS (*Global ForecastSystem*).

METGRID: Realiza la interpolación horizontal de los datos meteorológicos extraídos del modelo usado para inicialización al tamaño de grilla establecido para correr el modelo.

REAL: Realiza la interpolación vertical de los niveles de presión hasta el sistema de coordenadas normalizado sigma del WRF.

WRF: Este módulo contiene las ecuaciones con las que se realizan los cálculos para generar las predicciones, en él también se localizan las parametrizaciones que de los fenómenos físicos cuya escala espacio-temporal es inferior a la resolución fijada en las simulaciones.

Las salidas que proporciona el modelo WRF contienen información pronosticada para un intervalo de tiempo determinado de diferentes variables meteorológicas. Los ficheros que componen estas salidas cumplen con las especificaciones de un formato conocido como netCDF, cuya extensión en los ficheros generados es “.nc”. Al obtener las salidas del modelo WRF, es necesario contar con algunas herramientas de pos-procesamiento que nos permitan visualizar mediante gráficas, tablas, mapas o imágenes visuales los resultados arrojados. Estos resultados pueden ser visualizados con cualquier herramienta capaz de reconocer formato netCDF. Algunas de ellas son GrADS, RIP4, Vis5D, VAPOR y NCL. Sin embargo en la actualidad el más usado es el NCAR *CommandLanguage*(NCL) debido a su versatilidad y a que cuenta con funciones específicas para trabajar con las salidas del WRF. El NCL es un lenguaje interpretado, diseñado especialmente para el análisis y visualización de información científica, que permite leer ficheros netCDF, generar graficas de alta resolución y además permite a los usuarios un gigantesco número de opciones de visualización [25].

3 Verificación de Pronósticos y Tablas de Contingencia

3.1 Delineamientos Generales

3.1.1 Qué es la Verificación

Es la valoración y cuantificación de la relación de coincidencia entre conjuntos de pronósticos y observaciones [50]. Las actividades de verificación son útiles si ellas conducen a toma de decisiones con relación al producto verificado. Esta decisión debe generar cambios en la forma como se elaboran los pronósticos o conducir a una decisión de “no hacer nada” y considerar que el producto es satisfactorio. Por cuanto se asume que las observaciones son una representación precisa de la realidad, por lo general se requiere asumir que un punto de observación representa adecuadamente a los eventos dentro un área determinada.

El proceso de verificación involucra medición de la relación entre un pronóstico o conjunto de pronósticos y las correspondientes observaciones del predictando. Por eso cualquier método de verificación de pronóstico envuelve necesariamente la comparación entre pares coincidentes de pronósticos y observaciones a las cuales ellos pertenecen.

3.1.2 Propósito de la Verificación

Los propósitos de la verificación son principalmente de 2 tipos: administrativos y científicos. El propósito administrativo puede ser; convencer al gobierno (congreso) sobre la necesidad de establecer y mejorar el sistema de pronóstico, actualización de equipos, etc. La verificación con fines administrativos debe hacerse de forma continua para

monitorear la calidad de los pronósticos, y mejorar con el transcurrir del tiempo, la calidad de los pronósticos.

El propósito científico es identificar las fortalezas y debilidades del pronóstico con suficiente detalle, para determinar qué acciones hay que tomar para mejorar la calidad del pronóstico, es decir, suministrar información para la investigación y el desarrollo.

3.1.3 Modelos de Verificación

Como se puede ver en la Figura 3-1, un modelo de verificación comienza por el conjunto de datos de pares de pronósticos/observaciones (paralelogramo rojo en la Figura 3-1) a ser verificados, luego se define el propósito de la verificación (administrativo o científico). De acuerdo con el propósito de la verificación, el conjunto de datos puede ser estratificado para alcanzar el propósito deseado. Estratificación (también se podría llamar clasificación) significa separar los eventos del conjunto en uno o varios grupos de acuerdo a una regla de selección y luego, llevar a cabo la verificación en cada grupo por separado. La estratificación puede ser externa o interna. Cuando la regla de selección es independiente del elemento a ser verificado, la estratificación se llama externa. Un ejemplo de estratificación externa es el tiempo, diurna, estacional, etc., por ejemplo, el pronóstico de la precipitación en Canadá, es más acertado en invierno que en verano, debido a que la variabilidad de pequeña escala es más difícil de pronosticar. Si se desea pronosticar eventos extremos, la estratificación es interna, en ella la separación entre eventos extremos y no extremos se hace utilizando valores del predictando. La estratificación también se puede hacer de acuerdo a las observaciones o de acuerdo al pronóstico. De acuerdo con las observaciones significa definir categorías conforme con los valores observados, luego se define el tipo de verificación (categórico o de variable continua) y finalmente, las medidas de la verificación (tablas de contingencia para el análisis categórico; ploteo de dispersión, error absoluto medio, etc., para la verificación por variable continua). [50]

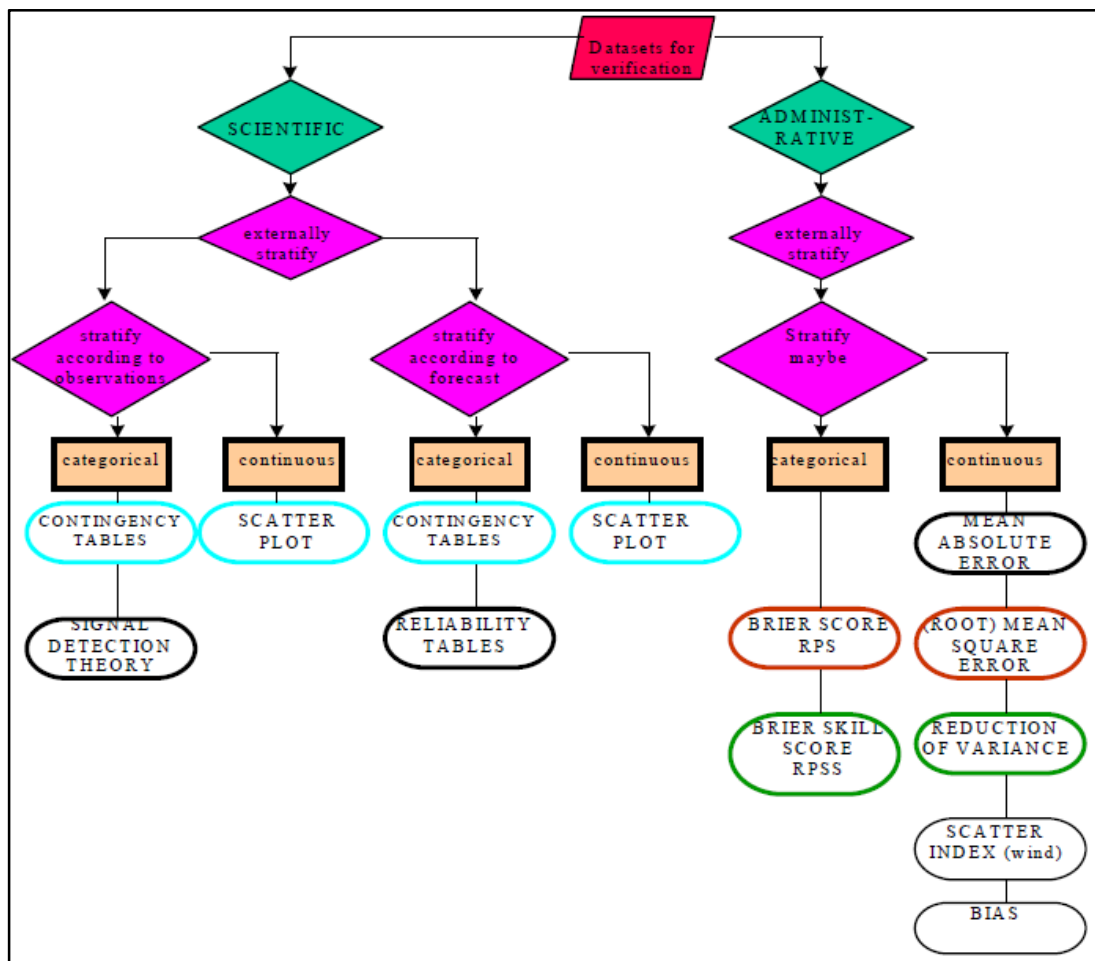


Figura 3-1: Modelo general para la verificación de pronóstico meteorológico

Fuente: [50]

3.1.4 Factorización de la Distribución Conjunta de Pronósticos y Observaciones

Un conjunto de datos de verificación dado, consiste en una colección de pares observación/pronóstico, cuyo comportamiento está ligado a la caracterización en términos de las frecuencias relativas de las posibles combinaciones de resultados observación/pronóstico.

De interés fundamental con respecto a la verificación de los pronósticos, es la distribución conjunta de los pronósticos y las observaciones. En términos prácticos, ambos, los pronósticos y las observaciones son variables discretas. Esto quiere decir, que si los pronósticos y las observaciones no son cantidades discretas, entonces se redondean operacionalmente a un conjunto de valores finitos.

Denotando el pronóstico por y_i , el cual puede tomar alguno de los valores I por y_1, y_2, \dots, y_I , y la correspondiente observación por o_j , la cual puede tomar alguno de los valores J , o_1, o_2, \dots, o_J . Entonces la distribución conjunta de la observación y el pronóstico es dado por:

$$p(y_i, o_j) = p\{y_i, o_j\} = Pr\{y_i \cap o_j\}; i = 1 \dots I, j = 1 \dots J \quad (3.1)$$

Esto es una función de distribución de probabilidad bi-variada discreta, que asocia una probabilidad a cada una de las $I \times J$ combinaciones posibles de pronósticos/observaciones.

Aún en los casos simples, para los cuales $I = J = 2$, esta distribución conjunta es difícil de usar directamente. La distribución conjunta se puede factorizar de dos maneras, las cuales son informativas sobre el problema de la verificación. Desde un punto de vista para el pronóstico, la forma más familiar e intuitiva es:

$$p(y_i, o_j) = p(o_i|y_j)p(y_i); \quad i = 1 \dots I, j = 1 \dots J \quad (3.2)$$

La cual ha sido llamada, la factorización de calibración – refinamiento. Una parte de esta factorización consiste en un conjunto de I distribuciones condicionales, $p(o_j|y_j)$ cada una de las cuales consiste de probabilidades para todos los J resultados o_j dado un pronóstico y_j . Esta distribución condicional especifica qué tan a menudo cada posible evento ocurre en esas ocasiones cuando es emitido un único pronóstico, o que tan bien cada pronóstico y_j es calibrado. La otra parte de la factorización es la distribución incondicional (marginal) $p(y_j)$ la cual especifica la frecuencia relativa de uso de cada uno

de los valores y_j , o con qué frecuencia es usado cada uno de los I posibles pronósticos. [59]

La otra posible factorización de la distribución conjunta de los pronósticos y observaciones es la factorización de tasa de probabilidad-base:

$$p(y_i, o_j) = p(y_i|o_j)p(o_j); \quad i = 1 \dots I, j = 1 \dots J \quad (3.3)$$

Aquí la distribución condicional $p(y_i|o_j)$ expresa las probabilidades de que cada uno de los pronósticos tenga valores permisible y_i habría sido emitido con anterioridad a los eventos meteorológicos observados o_j . Aunque este concepto puede parecer lógica inversa, puede revelar información útil acerca de la naturaleza de la actuación de los pronósticos. La distribución incondicional $p(o_j)$ consiste simplemente en las frecuencias relativas de los J eventos meteorológicos o_j en la verificación de los datos seleccionados, y es usualmente llamado muestra climatológica.

3.1.5 Atributos Escalares del Desempeño de un Pronóstico

Los atributos de un pronóstico dan información sobre uno o varios aspectos de la calidad de un producto de pronóstico. Por eso, un sistema de verificación incluye el cálculo de algunas medidas (índices) escogidas para describir los atributos que son más pertinentes para alcanzar el objetivo de la verificación. No existe una única medida de verificación que suministre información completa sobre localidad del producto de pronóstico.

El desempeño del pronóstico se evalúa usando una o varias medidas escalares (o sea, unidimensionales) o atributos. Muchos de estos índices han sido derivados de la experiencia. Al reducir la dimensionalidad, alguna información de la distribución conjunta inevitablemente es descartada. Algunos de los aspectos escalares o atributos de la calidad de un pronóstico son:

- **Precisión** “*accuracy*” es un término que indica el nivel de concordancia entre el pronóstico y las observaciones. La diferencia entre el valor observado y el

pronosticado se llama error, entre más pequeña sea esta diferencia mayor es la exactitud. La dificultad respecto a la exactitud consiste en definir su escala o sea los límites de su aceptación. Por ejemplo; un pronóstico estadístico para lluvia engelante (tormenta de hielo), puede ser útil para fines estadísticos pero no con suficiente exactitud para un pronosticador operativo.

- **Sesgo** “*bias*” El sesgo o sesgo incondicional, o sesgo sistemático, mide la correspondencia entre el pronóstico medio y el valor observado medio del predictando. Es diferente a la exactitud, la cual mide la correspondencia media entre pares individuales de pronósticos y observaciones. Por ejemplo, pronósticos de temperatura que son consistentemente muy cálidos, o pronósticos de precipitación que son consistentemente muy húmedos presentan sesgo, independientemente de si son razonablemente exactos o muy inexactos.
- **Habilidad** “*skill*” La habilidad o precisión relativa se refiere a la exactitud de un pronóstico dado, relativo a la exactitud de pronósticos producidos por un procedimiento estándar, por ejemplo; pronóstico climatológico, de persistencia o de azar (“*Chance*”).

La habilidad del pronóstico SS “*SkillScore*” se interpreta como un mejoramiento porcentual sobre el pronóstico de referencia. Sea A la medida de la exactitud del pronóstico a verificar y A_{ref} , la medida de la exactitud del pronóstico de referencia.

Entonces SS se define como:

$$SS_{ref} = \frac{A - A_{ref}}{A_{perf} - A_{ref}} \quad (3.4)$$

Donde, A_{perf} es la medida de la exactitud que se alcanzaría con un pronóstico perfecto [50].

3.2 Variables Continuas

La verificación de variables continuas típicamente suministra estadísticas sobre cuánto difieren los valores del pronóstico de las observaciones y con ello, el cómputo de medidas relativas contra sistemas de pronóstico de referencia. Algunos de los parámetros que se verifican de esta forma son la temperatura, la dirección y velocidad del viento, la precipitación acumulada, etc. Su comportamiento puede ser diferente. Cuando la temperatura se muestra bastante suave (pulida) o uniforme, la velocidad del viento es esporádica y la precipitación intermitente.

La primera forma de verificar puede ser un ploteo de dispersión del pronóstico contra las observaciones. Antes que ser una medida, este ploteo es una manera de visualizar la correspondencia entre el pronóstico y las observaciones. También permite ver posibles valores atípicos.

En una manera similar al “*scatterplot*”, plotear series de tiempo, pronóstico versus observaciones, fácilmente permite ver posibles valores atípicos, tendencias y relaciones dependientes en el tiempo [59].

3.3 Medidas de Precisión Escalar en Variables Continuas

MEAN ERROR (BIAS): El error medio o sistemático (sesgo), se calcula mediante la diferencia media entre el pronóstico y las observaciones, donde n es el número total de pronósticos comparados. Debido a que el ME carece de términos en valor absoluto o al cuadrado, se permite la cancelación de errores individuales positivos y negativos, poniendo en evidencia solamente el exceso, es decir el error sistemático.

$$ME = \left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - o_i) \quad (3.5)$$

El sesgo es el más simple y más familiar de las puntuaciones y puede proporcionar información muy útil sobre el comportamiento local de un parámetro de tiempo

determinado (por ejemplo, temperatura máxima cerca de la costa o la temperatura mínima sobre el suelo cubierto de nieve). El rango del ME va desde menos infinito hasta infinito, y una puntuación perfecta es, $ME = 0$. Sin embargo, es posible llegar a una puntuación perfecta para un conjunto de datos con errores grandes, si hay errores de compensación de un signo inverso. El ME no es una medida de precisión, ya que no proporciona información sobre la magnitud de los errores de pronóstico.

Una medida sencilla para compensar los posibles errores positivos y negativos que puede generar el ME es mediante el cálculo del MAE.

MEAN ABSOLUTE ERROR (MAE): El error absoluto medio es el promedio del valor absoluto de la diferencia entre el pronóstico y la observación, donde n es el número total de comparaciones pronosticadas. Es error absoluto porque tiene en cuenta tanto errores positivos como negativos.

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |y_i - o_i| \quad (3.6)$$

El rango del MAE es de cero al infinito y, como en el ME, una puntuación perfecta es igual $MAE = 0$, y aumenta a medida que las discrepancias entre las predicciones y observaciones se hacen más grandes. Podemos interpretar el MAE como una magnitud típica del error de pronóstico en un conjunto de datos de verificación dado, y por tanto, es una medida escalar de precisión de los pronósticos. Es recomendable ver siempre el ME y el MAE simultáneamente.

MEAN SQUARED ERROR (MSE): El error cuadrático medio, se calcula elevando al cuadrado la diferencia media del pronóstico y las observaciones, donde n es el número total de comparaciones pronosticadas, el **RMSE** (ecuación 3.8) es la raíz cuadrada del **MSE** (ecuación 3.7). la medida de error total, El **RMSE** incluye los componentes sistemáticos y al azar, los cuales se pueden separar usando medidas para el error sistemático y el error al azar, tales como error de sesgo y desviación estándar.

$$MSE = \left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - o_i)^2 \quad (3.7)$$

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - o_i)^2} \quad (3.8)$$

Al igual que con el MAE, su alcance es de cero al infinito, con una puntuación perfecta de $MSE = 0$ o $RMSE = 0$. Debido a la segunda potencia, el MSE y RMSE son mucho más sensibles a grandes errores de pronóstico que el MAE. Esto puede ser especialmente perjudicial en la presencia de posibles valores atípicos en los conjuntos de datos y, en consecuencia, al menos con conjuntos de datos pequeños o limitados se prefiere el uso del MAE.

3.4 Pronósticos Categóricos de Predictandos Discretos (Eventos Categóricos)

Categorico, significa que el pronóstico de un enunciado consiste en afirmar que uno y solo uno de un posible conjunto de eventos va a ocurrir (ejemplo; se presentarán lluvias por encima de 60mm diarios). El pronóstico categorico no contiene expresión para la incertidumbre, lo que lo diferencia del pronóstico probabilístico. Un predictando discreto es una variable observable que toma uno y solo un valor de un conjunto finito de valores posibles. Esto lo diferencia de un predictando continuo que (al menos conceptualmente) puede tomar cualquier valor en la porción relevante de la línea real.

3.4.1 Tablas de Contingencia de 2X2

El término de tabla de contingencia fue utilizado por primera vez por el estadístico Karl Pearson en 1904 y es esencialmente un sistema de visualización utilizado para analizar y registrar la relación entre dos o más variables categóricas. Es el equivalente categorico del gráfico de dispersión utilizado para analizar la relación entre dos variables continuas. En esta verificación existe una correspondencia uno a uno, entre los valores permitidos

del pronóstico y los valores observados del predictando. En términos de la distribución conjunta del pronóstico y las observaciones, $I = J$. La situación más simple posible es el caso dicotómico $I = J = 2$, o verificación del pronóstico categórico "sí/no". Aquí hay $I = 2$ pronósticos posibles, bien que el evento ocurra ($I = 0, o y_1$) o no ocurra ($I = 1, o y_2$). Similarmente, hay $J = 2$ resultados posibles: bien que el evento ocurra (o_1) o que no ocurra (o_2).

Se necesitan estadísticas categóricas para la evaluación binaria "sí / no", los pronósticos del tipo de declaraciones de que un evento suceda o no suceda. El primer paso para verificar previsiones binarias es compilar una tabla de contingencia de 2X2 que muestra la frecuencia de los "sí" y "no" predicciones y observaciones correspondientes, como la que se muestra en la Figura 3-2.

Event forecast	Event observed		Marginal total
	Yes	No	
Yes	Hit	False alarm	Fc Yes
No	Miss	Correct rejection	Fc No
Marginal total	Obs Yes	Obs No	Sum total

Figura 3-2: Tabla de contingencia de 2X2, clasificación de eventos dicotómicos

Fuente: [36]

Hay dos casos en los que el pronóstico es correcto, ya sea un "Hit (éxito)" (evento observado y pronosticado) o un "*correctrejection* (Rechazo Correcto)" o "correcta no predicción" (evento no observado y no pronosticado) y dos casos en los que la predicción es incorrecta, ya sea una "*False alarm* (Falsa alarma)" (evento no observado y pronosticado) o un "*Miss* (evento desaparecido)" (evento observado y no pronosticado). Los llamados distribuciones marginales de los pronósticos y observaciones son los totales que se proporcionan en las columnas de la derecha y las filas inferiores de las tablas de contingencia, respectivamente. Un sistema de previsión perfecta tendría sólo "Hit (éxitos)" y "*correctrejection* (Rechazo Correcto)", con las otras celdas iguales a cero.

A menudo la tabla de contingencia se presenta en la forma como se muestra en el panel izquierdo de la Figura 3-3.

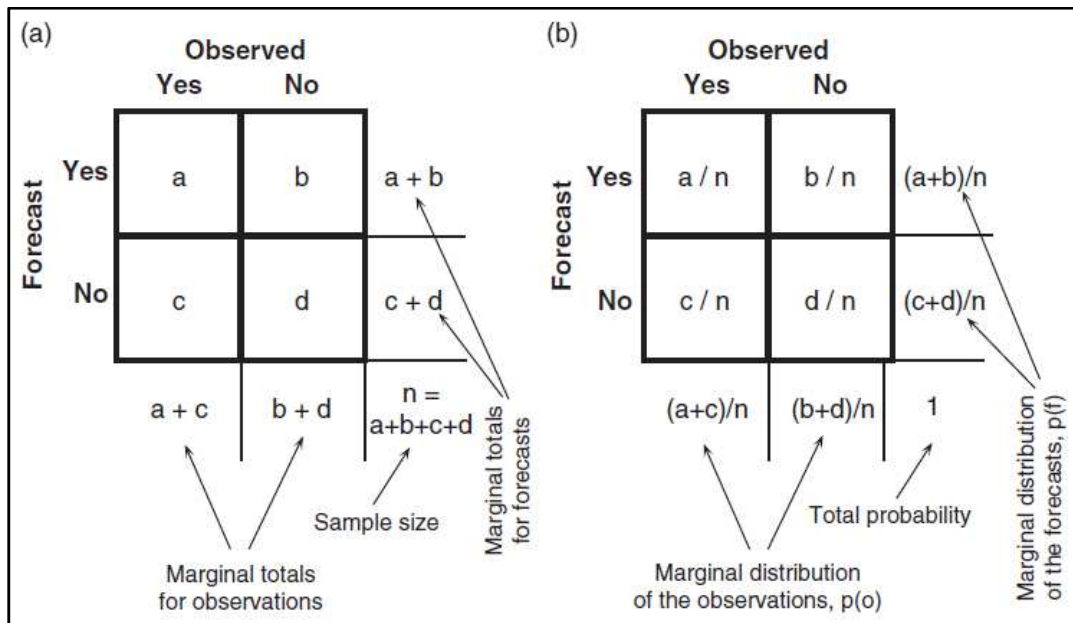


Figura 3-3: Tabla de contingencia de 2X2

Fuente: [36]

Además, si cada una de las entradas en la Figura 3-3 (panel izquierdo) se divide entre n (número de observaciones) se obtiene entonces la distribución conjunta de los pronósticos y las observaciones (Figura 3-3 panel derecho). En este panel, a/n es la frecuencia relativa o estimación de la probabilidad conjunta correspondiente $p(y_1, o_1)$, b/n corresponde a la estimación de la probabilidad conjunta $p(y_1, o_2)$, etc. El análisis en este trabajo se centra en la tabla de contingencia (Figura 3-2 o 3-3 panel izquierdo)

a = Número de veces que se pronosticó el fenómeno y éste ocurrió. Esta letra generalmente denota “Hits”

b = Número de veces que se pronosticó el evento pero éste no ocurrió. Esta llamado también “False Alarms”

c = Número de veces que no se pronosticó el evento y éste sí ocurrió. Denotado también por “misses”

d = Número de veces que no se pronosticó el evento y éste tampoco ocurrió. Denotado por “Correct non-events”

La tabla también incluye el cálculo de los “marginal totals” tanto para las observaciones como para el pronóstico. El valor en la esquina inferior derecha es el tamaño total de la muestra de verificación y debe ser igual a la suma de las cuatro cajas en la tabla. A partir de la tabla de contingencia es posible encontrar varios estadísticos para el análisis del pronóstico de precipitación.

3.4.2 Atributos Escalares de las Tablas de Contingencia de 2X2

La definición aparentemente simple de un evento binario, y la subsiguiente tabla de contingencia 2X2, esconden una complejidad bastante sorprendente. Hay una serie de medidas para abordar esta compleja cuestión y se definen destacando algunas de sus propiedades. La mayoría, si no todos, tienen un fondo histórico largo, pero que todavía se utilizan con mucha frecuencia. Hay que recordar que en ningún caso es suficiente aplicar apenas una única medida de control.

BIAS: El Bias o Bias de frecuencia compara el número de veces que un evento fue pronosticado con el número de veces que un evento fue observado. El sesgo de los pronósticos binarios compara la frecuencia de los pronósticos (Pronostico -> Sí) a la frecuencia de ocurrencias reales (Observación -> Sí) y se representa por la relación,

$$B = \frac{(a + b)}{(a + c)} [\sim \text{Fc Yes} / \text{Obs Yes}] \quad (3.9)$$

Además, puesto que (a+b) corresponde al número de pronósticos *Fc* que superan un umbral dado, y (a+c) al número total de observaciones *Obs* que supera ese mismo umbral, la fórmula anterior también se puede escribir como:

$$B = \frac{Fc}{Obs} [\sim \text{Fc Yes} / \text{Obs Yes}] \quad (3.10)$$

Esta expresión mide la relación de la frecuencia de eventos de pronóstico a la frecuencia de eventos observados. Ella indica si el sistema de predicción tiene una tendencia a subestimar (BIAS <1) o sobrestimar (BIAS > 1) eventos. No miden los puntos de la

cuadrícula pronóstico que corresponden con los puntos de la malla observados, sólo se mide frecuencias relativas en general. Puede ser difícil de interpretar cuando el número de Sí pronósticos es mucho mayor que el número de Sí observación.

PROPORTION CORRECT (PC): Es el porcentaje de pronóstico que han acertado. Son los eventos pronosticados correctamente sobre el número total de eventos.

$$PC = \frac{(a + d)}{n} [\sim (\text{Hits} + \text{Correct rejections}) / \text{Sum total}] \quad (3.10)$$

En general resuelve la pregunta, ¿qué fracción de las predicciones es correcta? El rango de PC varía desde cero (0) cuando ningún evento es pronosticado, hasta uno (1) cuando todos los eventos son pronosticados, una perfecta puntuación = 1. PC es simple e intuitivo, pero puede ser engañoso, ya que está fuertemente influenciado por la categoría más común, que son los casos que por lo general más se presentan en la meteorología, los "*CorrectRejections*".

Como se mencionó, este índice tiende a ser engañoso. Un ejemplo de los errores que puede arrojar es un ejemplo clásico en la literatura de la verificación, el caso de John Finley que fue un sargento de la Armada de los Estados Unidos, en marzo 1884 el sargento Finley inició previsiones de tornados dos veces al día sobre 18 regiones en los Estados Unidos, al este de las Montañas Rocosas. Finley afirmaba predicciones del 95,6% a 98,6% de precisión, en general para el primer período de 3 meses, dependiendo de la hora y el distrito, con algunos distritos lograba un 100% de precisión para los 3 meses.

Un crítico de los resultados indicó que el 98,2% de precisión podría obtenerse simplemente con la predicción de "no tornado", esto puede verse claramente en la Figura 3-4, en la que tienen mayor peso los casos de "*CorrectRejections*" es decir los casos en los que no se pronosticó y no se observó ningún tornado, en (a) se obtiene una precisión del 96.6% y en (b) un 98,2% de precisión. Esto ilustra claramente la necesidad de verificar los resultados con índices más significativos [36].

(a)	Tornados Observed			(b)	Tornados Observed		
		Yes	No			Yes	No
Tornados	Yes	28	72	Tornados	Yes	0	0
Forecast	No	23	2680	Forecast	No	51	2752
n = 2803				n = 2803			

Figura 3-4: Caso de Finley en la predicción de tornados en los EE.UU

Fuente: [36]

PROBABILITY OF DETECTION (POD): La probabilidad de detección mide la proporción de eventos observados que se pronosticaron correctamente, este es igual al número de eventos pronosticados correctamente sobre número total de eventos observados,

$$POD = \frac{a}{(a + c)} [\sim \text{Hits} / \text{Obs Yes}] \quad (3.11)$$

El POD responde la pregunta, ¿qué fracciones de los eventos observados fueron correctamente pronosticados? El rango del POD desde cero (0) para un mal pronóstico de eventos observados, hasta uno (1), una puntuación perfecta = 1. Sensible a los "Hits", pero hace caso omiso a las "False Alarm". Bueno para "RareEvents". Puede ser artificialmente mejorado mediante la emisión de pronósticos correctos aumentar el número de "Hit". Debe ser utilizado en conjunción con la relación de falsa alarma (FAR) [36].

FALSE ALARM RATIO (FAR): Es la fracción de eventos pronosticados y que no ocurrieron, este es igual a las falsas alarmas sobre el número de pronósticos del evento.

$$FAR = \frac{b}{(a + b)} [\sim \text{False alarms} / \text{Fc Yes}] \quad (3.12)$$

El FAR responde la pregunta, ¿qué fracción de las predicciones no ocurren? El rango de FAR varía desde cero (0) hasta uno (1), una puntuación perfecta = 0. Sensible a las "False Alarm" pero hace caso omiso a los "Miss", muy sensible a la frecuencia

climatológica del evento. Debe ser utilizado en conjunción con la probabilidad de detección (POD) [36].

CRITICAL SUCCESS INDEX (CSI): Una medida ampliamente usada para eventos raros, es *Threat Score* (TS), o Índice Crítico de Éxito (CSI), el TS combina el Hit Rate y el False Alarm Ratio en un solo estadístico para eventos que tienen poca frecuencia:

$$CSI = \frac{a}{(a + b + c)} [\sim \text{Hits} / (\text{Hits} + \text{False alarms} + \text{Misses})] \quad (3.13)$$

Este índice agrupa propiedades de POD y de FAR. El rango de CSI varía desde cero (0) hasta uno (1), una puntuación perfecta = 1, no hay habilidad para pronosticar = 0. Mide la fracción de eventos observados y/o pronosticados que se predijo correctamente. Se puede considerar como la precisión en la que los “*CorrectRejections*” han sido eliminados de la consideración. Es decir, CSI sólo se refiere a las previsiones que son importantes (suponiendo que los “*CorrectRejections*” no son importantes). Sensible a los “*Hits*”, penaliza ambos el “*Miss*” y la “*False Alarm*”. No distingue el origen del error de pronóstico. Depende de la frecuencia de los eventos climatológicos (las puntuaciones más pobres para eventos raros), ya que algunos de los “*Hits*” pueden ocurrir puramente debido al azar. Se considera una función no lineal de POD y FAR. Debe ser utilizado en combinación con otras estadísticas de la tabla de contingencia (por ejemplo, Bias, POD, FAR) [36].

EQUITABLE THREAT SCORE (ETS): El puntaje equitativo de amenaza (ETS), también conocido como *Gilbert’sSkill Score* (GSS), ajusta el número de Hits asociados con un cambio al azar, y se define como,

$$ETS = \frac{(a - ar)}{(a + b + c - ar)} \left[\frac{\sim (\text{Hits} - \text{Hits random})}{(\text{Hits} + \text{False alarms} + \text{Misses} - \text{Hits random})} \right] \quad (3.14)$$

Donde,

$$ar = \frac{(a + b) * (a + c)}{n} \quad (3.15)$$

"ar" es el número de aciertos aleatorios de los pronósticos

El ETS responde la pregunta, ¿qué tan bien hizo el pronóstico de eventos con respecto a los eventos observados (contabilizando por "Hits" que deberían ser esperados por cambios)? El rango de ETS varía desde -1/3 hasta uno (1), una puntuación perfecta = 1, no habilidad para pronosticar = 0. Mide la fracción de eventos observados y pronosticados que se predijo correctamente, ajustado por la frecuencia de los "Hits" que se esperan que ocurra simplemente por azar (por ejemplo, es más fácil de pronosticar correctamente ocurrencia de lluvia en un clima húmedo que en un clima seco). El ETS se usa a menudo en la verificación de las lluvias en los modelos de pronóstico numérico del tiempo (PNT) porque su confiabilidad permite puntajes para comparar de manera más justa a través de diferentes regímenes; sin embargo no es realmente equitativa. Sensible a los "Hits". Debido a que penaliza tanto a fallos y falsas alarmas de la misma manera, no distingue el origen del error de pronóstico. Debe ser utilizado en combinación con al menos otra estadística de tablas de contingencia (por ejemplo, Bias) [36].

HEIDKE SKILLSCORE (HSS): El HSS nos indica cuál es la exactitud del pronóstico respecto al tomado bajo cambios aleatorios,

$$HSS = \frac{(a + d) - (ec)}{n - (ec)} \left[\frac{\sim (\text{Hits} + \text{Correct Rejections}) - (\text{expected correct})_{random}}{(\text{Hits} + \text{False alarms} + \text{Misses} - \text{Hits random})} \right] \quad (3.16)$$

donde,

$$(\text{expected correct})_{random} = \frac{1}{n} [(a + c) * (a + b) + (d + c) * (d + b)]$$

El HSS intenta responder la siguiente pregunta, ¿cuál fue la exactitud del pronóstico en relación al de azar? El rango va desde $-\infty$ hasta 1, 0 indica no acierto y 1 acierto perfecto. Mide la fracción de predicciones correctas después de eliminar aquellos pronósticos que

serían correctos debido a cambios aleatorios. Esta es una forma de la puntuación de habilidad generalizada, donde la puntuación en el numerador es el número de predicciones correctas, y la previsión de referencia en este caso es el azar. En meteorología, al menos, el azar no es generalmente el mejor pronóstico para comparar, esto podría ser mejor para la climatología (a largo plazo valor medio) o persistencia (predicción = observación más reciente, es decir, sin cambios) o alguna otra norma [36].

3.5 Pronósticos de Probabilidad

Todo pronóstico involucra un nivel de incertidumbre [36]. Sin embargo, los pronósticos determinísticos (como los examinados hasta ahora) y su verificación no dan información sobre incertidumbre inherente al fenómeno meteorológico o evento bajo consideración. Dada la probabilidad de evento esperado con valores entre 0 y 1, los pronósticos probabilísticos toman en cuenta de una mejor manera la distribución conjunta subyacente entre los pronósticos y las observaciones.

El comportamiento de un sistema de pronóstico probabilístico se puede ver mediante la construcción del diagrama de confiabilidad ("*reliability diagram*") el cual es un gráfico de la frecuencia relativa observada de un evento, como función de su probabilidad de pronóstico en categorías de probabilidad definidas (por ejemplo, en intervalos de 10%). La curva resultante del diagrama de confiabilidad es entonces una indicación del ajuste entre la probabilidad media del pronóstico y la frecuencia media observada [36].

4 Técnicas Orientadas a Objetos

4.1 Conceptualización de la Técnica de Verificación Orientada a Objetos

La técnica estadística tradicional para verificación del pronóstico cuantitativo de la precipitación QPF (*"Quantitative Precipitation Forecast"*) se basa en la superposición en los puntos de grilla, y los campos de precipitación observada y pronosticada por el modelo. Sin embargo, esta técnica de grilla, también llamada la técnica orientada a mediciones (*"Measures Oriented Technique"*) [4], falla a medida que la resolución del modelo aumenta y la convección atmosférica se resuelve explícitamente [19, 4, 16, 12].

Esta falla se produce porque la técnica tradicional penaliza severamente pequeños desplazamientos de los elementos convectivos simulados con respecto a los observados, con bajos coeficientes de correlación o aún negativos, alto error cuadrático medio y valores pobres en la estadística categórica (lluvia/no lluvia). Por esta razón, métodos alternativos tales como; técnicas orientadas a distribuciones, técnicas orientadas a objetos, etc., se han venido desarrollando en los últimos decenios [7, 19, 3, 16, 48].

En estas nuevas técnicas, los objetos, los cuales son fácilmente distinguibles ante el ojo humano, son identificados simultáneamente en ambos campos pronosticado/observado [16]. Estos objetos pueden ser descritos geoméricamente, de tal manera que atributos relevantes de ambos objetos simulado y observado pueden ser comparados. Los atributos incluyen ítems tales como localización, forma, orientación, tamaño, duración, intensidad de la lluvia, etc. Usualmente, los objetos son áreas contiguas de lluvia (simulada y observada) o patrones de reflectividad del radar, etc.

4.2 Desarrollo Histórico

Uno de los precursores de la técnica orientada a objetos para verificación del pronóstico cuantitativo de la precipitación QPF ("Quantitative Precipitation Forecast") fue Ebert [19]. Estos investigadores anotaron que el modelo puede acertar la formación del campo de la precipitación pero presentar error en su posición relativa al sistema observado. Estos desplazamientos relativos son severamente castigados por la verificación estadística tradicional con coeficientes bajos o negativos de correlación, alto error cuadrático medio y valores pobres de estadística categórica (lluvia/no lluvia). Sin embargo, en estos casos, el pronóstico suministra información importante al pronosticador si se compara el comportamiento general de los sistemas de lluvias observados con los simulados. Los autores de este artículo se dieron entonces a la tarea de evaluar inicialmente este error de posición. (el comentario de la figura se hace antes de ella)

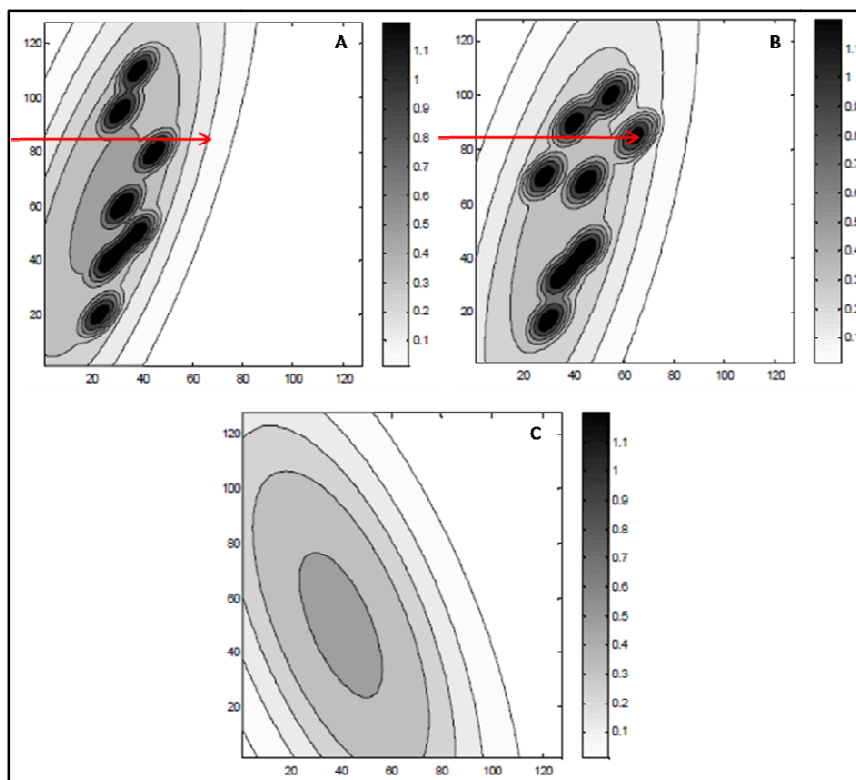


Figura 4-1: Comparación de la precipitación, (A) observada, (B) pronosticada modelo alta resolución, (C) baja resolución.

Fuente: figura adaptada por el autor [4]

El error total del pronóstico se compone de error de posición, de intensidad y error residual asociado con el patrón o forma del sistema precipitante [19]. Los autores estimaron el error total por una técnica similar al de Hoffman [28] y además introdujeron el concepto de áreas contiguas de precipitación CRA (“*Contiguous Rain Areas*”).

A medida que se tiene una alta resolución aumenta la dificultad en la técnica de verificación estadística tradicional, ilustrada en la Figura 4-1, [4]. En esta Figura se compara la precipitación observada con la simulada por un modelo de alta resolución (Figura 4-1B) y otro de baja resolución (Figura4-1C). Además, en estas figuras los campos de precipitación fueron construidos usando una función de forma elíptica usando un domino de 128 X 128 puntos de grilla [61].

Mientras que el modelo de baja resolución (Figura4-1C) presenta errores en la orientación, tamaño del campo de la precipitación, posición de su centro y no reproduce bien las elipses de menor escala, el modelo de alta resolución (Figura 4-1B) acierta de una mejor manera en la orientación, posición del centro, tamaño del campo y reproducción de las elipses de menor escala. Sin embargo, nótese que en el pronóstico de alta resolución las elipses de menor escala se encuentran desplazadas con relación a las observadas (compare por ejemplo el desplazamiento hacia la derecha de la elipse señalada por la flecha, en las (Figura 4-1B) con respecto a su posición en la (Figura 4-1A). Además, el campo de la (Figura 4-1B), en general, se encuentra un poco inclinado hacia la derecha y hacia abajo. El error en el posicionamiento de las elipses de menor escala en la Figura 4-1B con respecto a las observadas en la Figura 4-1A, conllevan a los errores que son castigados en el uso de la técnica estadística tradicional mencionados en las referencias anteriores. Para solucionar este problema, [4] sugieren el uso de una nueva técnica de verificación orientada a objetos, la cual es diferente a otras técnicas, tales como técnica orientada a medidas en puntos de grilla (ya mencionada), técnica orientada a distribuciones, etc.

4.2.1 Método de Verificación Basada en Objetos

Un fuerte avance en la técnica de verificación orientada a objetos fue dado por Davis [16]. Por ello se ilustra a continuación la técnica orientada a objetos desarrollada por

estos autores. Además, esta técnica se utiliza actualmente de manera operativa en el centro de predicciones del servicio nacional de meteorología de los Estados Unidos para verificación de pronósticos.

En este trabajo se diferencia dos etapas: el reconocimiento de los objetos y el de concordancia (“*matching*”) entre los objetos pronosticados y los observados. El reconocimiento de los objetos involucra 3 sub etapas; 1). Convolución, 2). Colocación de umbrales y 3). Filtración. Estas etapas se reproducen en la Figura 4-2, [16].

4.2.1.1 Convolución

Es un proceso de suavización (“*smoothing*”) en el cual la precipitación en un punto de grilla es substituido por el valor medio dentro de un disco de convolución con un radio de convolución dado. La Figura 4-2a muestra el ejemplo de un campo de precipitación originalmente pronosticado por el modelo WRF. La tercera dimensión en esta figura corresponde al valor de la precipitación. La Figura4-2b muestra la misma precipitación después de aplicarle el filtro de convolución (suavización) con un radio de convolución igual a la distancia entre 4 puntos de grilla. En términos gráficos, la convolución se representa por un cilindro de base definida por el radio de convolución y altura igual al valor de la precipitación media (intensidad).

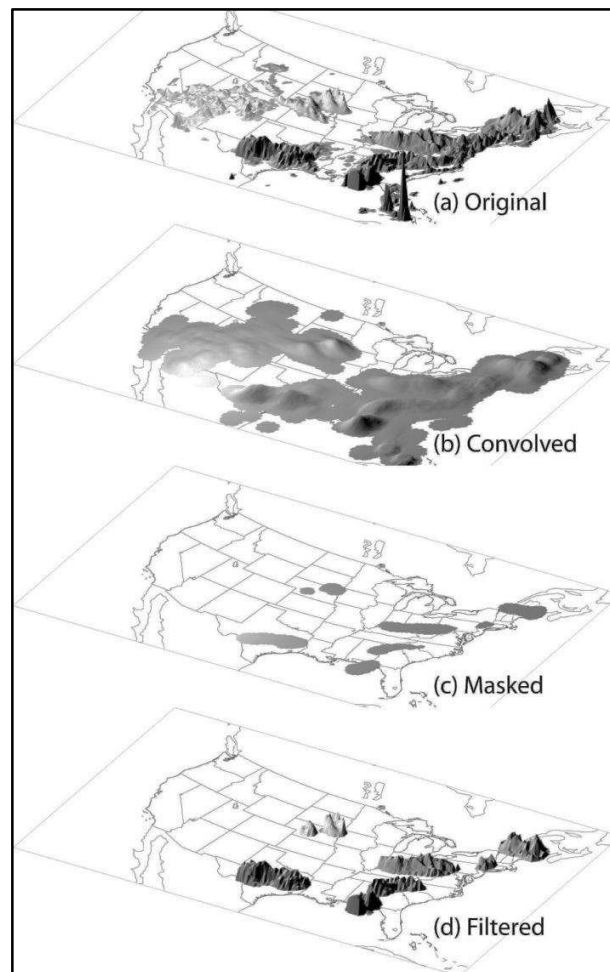


Figura 4-2: Proceso de suavizado de la precipitación

Fuente: [16]

4.2.1.2 Colocación del Umbral

Esta etapa permite detectar el contorno del objeto. Colocar umbral sin convolución no resulta en límites para el objeto, porque el campo puede contener huecos de varios tamaños y los contornos del objeto no se distinguirían bien. Los umbrales se pueden calibrar, para distinguir entre áreas de gran tamaño e intensidad, de aquellas que son más débiles y más aisladas.

El resultado de la convolución y colocar umbral, es una máscara binaria (El calificativo binario significa que el campo queda cubierto con zonas enmascaradas y zonas en

blanco). Esta máscara es colocada sobre el campo de la precipitación, dando como resultado parches contiguos de precipitación rodeadas por regiones con precipitación cero, tal como se observa en la Figura 4-2d. Note que los valores originales de precipitación son retenidos y las estadísticas en cada parche pueden ser examinadas. Osea, el proceso en el que se promedia es sólo para dibujar los objetos y no cambia para nada los valores de precipitación que se van a verificar.

4.2.1.3 Formas Geométricas

Los objetos se asocian en formas simples (elipses en muchos casos). El propósito es crear objetos con formas simples tales que el “*aspect ratio*” ángulo (orientación) y otras propiedades tengan interpretación que no sea ambigua.

4.2.1.4 Atributos o Propiedades de los Objetos

Una vez que los objetos han sido identificados a los mismos se les asigna varias propiedades tales como Intensidad, área, ángulo axial, etc. Como medida de estimación de la intensidad, en el trabajo mencionado se evaluó el percentil 25 y 95 de la distribución de la lluvia medida sobre una muestra tri-horaria durante el período julio-agosto de 2001 sobre el territorio de los EE.UU.

4.2.1.5 Concordancia (“*MATCHING*”)

La concordancia entre los objetos observado y pronosticado se basa solamente en la distancia entre los centros (o centroides) de los pares de objetos (pronosticado y observado respectivamente). Si A_f y A_0 , son las áreas de los objetos pronosticado y observado respectivamente, entonces ese par de objetos concuerdan si la distancia D que los separa es tal que,

$$D < \left(A_f^{\frac{1}{2}} + A_0^{\frac{1}{2}} \right) \quad (4.1)$$

Puesto que $A^{1/2}$ es el tamaño o dimensión lineal del objeto, de acuerdo a esta fórmula, la condición de concordancia es equivalente a decir que la separación entre centroides es menor que la suma de las dimensiones lineales de los dos objetos. Otra condición de concordancia más restrictiva es,

$$D < \frac{\left(A_f^{1/2} + A_0^{1/2}\right)}{2} \quad (4.2)$$

La condición menos restrictiva permite que dos objetos concuerden aun estando relativamente lejos el uno del otro. Con la más restrictiva los dos objetos casi se tocan.

4.2.2 Otras Aproximaciones de la Técnica Orientada a Objetos

El método brevemente descrito, a pesar de su relativa complejidad, es sin duda un gran aporte en la búsqueda de nuevos métodos de verificación. Sin embargo, en el camino de la técnica orientada a objetos, otras aproximaciones menos complejas han sido aplicadas. Entre ellas cabe mencionar [21, 10, 35, 18, 27], entre otros.

La técnica descrita por [21], se aplicó para verificar el pronóstico con el modelo MM5 y formación de la convección atmosférica durante el periodo comprendido entre el 5 abril y el 20 de Septiembre 1999, en un área sobre el lago Michigan [21]. El objeto en este caso fue un área (el tercer dominio del modelo) que cubre una extensión de 200,448 km². Los atributos del objeto en el campo simulado fueron: velocidad convectiva ascendente superior a 0.3 m/s acompañada de precipitación durante un tiempo mínimo o superior a 1h. El atributo del campo observado fue la reflectividad del radar mayor a 40dB por un mínimo de 1h, en cualquier parte del tercer dominio. Los atributos mencionados sirvieron de condición de concordancia entre el pronóstico y las observaciones.

La técnica orientada a objetos [10], se usó para verificar el pasaje del frente de brisa marina sobre un dominio específico ubicado sobre la costa oriental de la Florida, fenómeno que sucede en el correr de las horas diurnas. El paso del frente de brisa es

caracterizado por un cambio de vientos del continente hacia el mar ("*offshore*") a vientos de dirección contraria, es decir del mar hacia el continente ("*onshore*"). Vientos sostenidos entre 0° y 180° se consideran del este y entre 180° y 360° del oeste. Para la aplicación de la técnica orientada a objetos, en este trabajo el objeto es el área del dominio. En calidad del atributo del objeto, se considera la fracción de área del dominio con transición de frente de brisa (viento del este) f_o en la observación, y f_R en el pronóstico respectivamente.

De acuerdo a esta metodología, si $f_o = 0$ (todo el dominio se encuentra cubierto con viento de oeste) y $f_R = 0$ (todo el dominio cubierto también con viento del oeste) se dice que, en ese día específico, no se presentó transición de frente de brisa y se considera como una concordancia "*hit*".

En el trabajo de [27], la técnica orientada a objetos se aplica no para verificar pronósticos sino para identificar y caracterizar eventos de lluvia extrema que tienen lugar en períodos cortos de tiempo (1h) [27]. En este trabajo, los objetos son áreas con precipitación extrema (mayor a 55.4 mm) que varían desde cientos hasta un poco más de 10,000 km² (100x100km). Los atributos son la intensidad de la lluvia y condiciones ambientales tales como CAPE ("*Convective Available Potential Energy*"), agua precipitable, punto de rocío y velocidad y dirección del viento medio por capas de nubes.

4.3 Técnica Orientada a Objetos Utilizada en el Presente Trabajo de Tesis

En el presente trabajo de tesis, se pretende usar una técnica similar a [27, 10, 21] para verificar pronósticos de precipitación diaria que supere un umbral dado. En este caso, el dominio total es dividido en áreas iguales de 50x50km, cada una de las cuales representan el objeto. El atributo del objeto es el total diario de precipitación en cada una de estas áreas tanto en el dominio observado como en el pronosticado. La condición de concordancia es que el total diario supere el umbral (60mm) al menos en algún punto de grilla en objetos correspondientes en ambos dominios. Para alcanzar esto; 1). Un algoritmo especial busca la precipitación máxima en cada uno de los objetos y 2). Se

analiza la condición de concordancia (para cada una de estas áreas de 50x50km en ambos dominios) a través de tablas de contingencia.

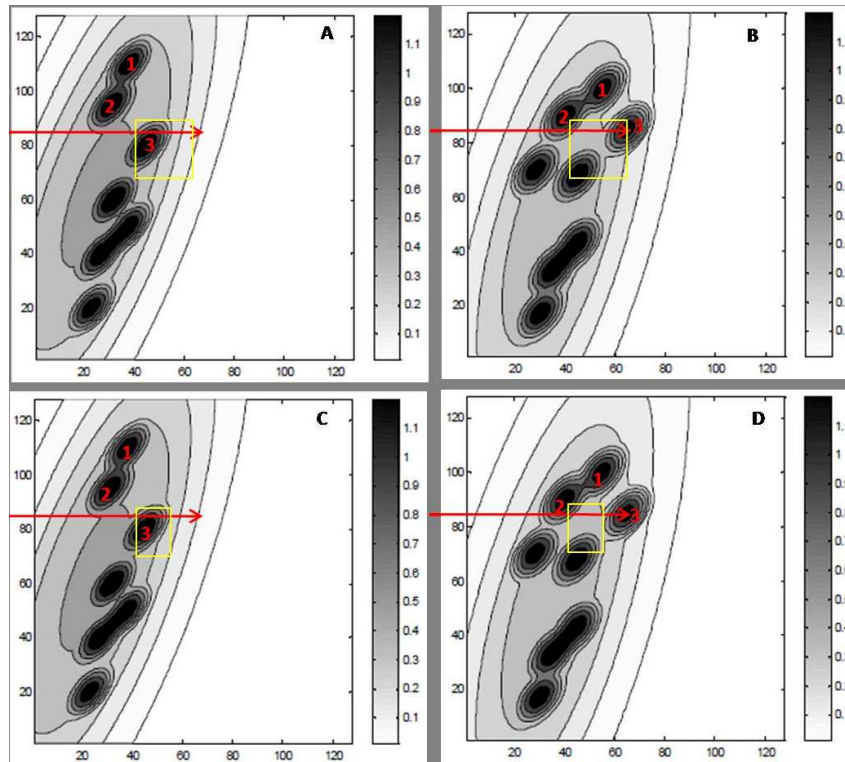


Figura 4-3: Se resalta la reducción del área del objeto, (A) observado área mayor, (B) modelado área mayor, (C) observado área menor y (D) observado área menor.

Fuente: figura adaptada por el autor [4]

La diferencia del tamaño de las áreas (objetos) es constante [27]. Se espera que, a medida que el tamaño del objeto disminuya en el método aquí propuesto, el nivel de acierto también disminuya, ya que se reduce el número de objetos que concuerdan, tal como se muestra en la Figura 4-3. En esta figura se ha resaltado con marco amarillo la reducción del área del objeto y también se ha numerado las elipses de menor escala (objetos) en la parte superior de la figura. Como se puede apreciar, al reducir el tamaño del objeto (panel inferior), se pierde la concordancia en el objeto No. 3 de los campos observado y pronosticado (Figura4-3C y 4-3D respectivamente).

En el límite cuando el tamaño del área tiende a cero, el método propuesto aquí se reduce al método de verificación de la estadística tradicional en puntos de grilla.

5 Metodología

5.1 Área de Estudio

El área de estudio comprende el dominio No. 3 configurado en el modelo WRF instalado en la Fuerza Aérea Colombiana (FAC), delimitados por las coordenadas (9°9'28"N 77°8'35"W), (9°9'28"N 70°44'19"W), (0°52'17"N 70°44'19"W) y (0°52'17"N 77°8'35"W), ver Tabla 5-1 y comprende los departamentos de Norte de Santander, Santander, Boyacá, sur de Bolívar, sur de Sucre, Córdoba, Antioquia, Choco, Caldas, Risaralda, Quindío, Cundinamarca, Tolima, Huila, Valle del Cauca, Oriente del Cauca, Norte de Caquetá y Putumayo, Guaviare, Meta, Casanare y occidente de Arauca, como se puede observar en la Figura 5-1.



Figura 5-1: Dominio del modelo WRF de la FAC, con resolución espacial de 5 Km.

Fuente: [19]

Tabla 5-1: Coordenadas del dominio No. 3 del WRF de la FAC

DOMINIO	LATITUD	LONGITUD
No. 3	0°52'17"N - 9°9'28"N	77°8'35"W - 70°44'19"W

5.1.1 Descripción Físico-Geográfica

A continuación se realiza una descripción física y geográfica del área de estudio, haciendo espacial énfasis en la diversidad de las condiciones topográficas, hidrográficas y del clima.

Esta área, principalmente está concentrada sobre la Región Andina, cubre una extensión de 653.872 kilómetros cuadrados, que representan el 57% de la extensión de todo el país y comprende los departamentos de Norte de Santander, Santander, Boyacá, sur de Bolívar, sur de Sucre, Córdoba, Antioquia, Choco, Caldas, Risaralda, Quindío, Cundinamarca, Tolima, Huila, Valle del Cauca, Oriente del Cauca, Norte de Caquetá y Putumayo, Guaviare, Meta, Casanare y occidente de Arauca. Su relieve, típico andino, es tan abrupto como variado. Comprende tanto el Macizo Colombiano, donde los Andes se desprenden en las tres estribaciones conocidas en como cordilleras Oriental, Central y Occidental, y entre las cuales se desarrollan superficies medias y bajas, como los valles de diversos ríos y la planicie aluvial del Magdalena. De la misma manera, está rodeada por las regiones de la Amazonía, al sur, la Orinoquía, al oriente, la Pacífica al occidente, la Andina Occidental al occidente y la Caribe al norte, y compartiendo una frontera terrestre internacional con Venezuela, al nororiente.

5.1.2 Topografía

El área de estudio abarca un territorio extenso pero sobre todo muy heterogéneo, partiendo desde el suroccidente, se divide en el Macizo colombiano en tres cordilleras, la cordillera Occidental, la Cordillera Central y la Cordillera Oriental. La cordillera Central está separada de la Occidental una distancia promedio de 400 m por una falla geológica ocupada por el río Patía al sur y por el río Cauca al norte, sobre ella encontramos picos de más de 5.000 msnm cubiertos de nieves permanentes, entre ellos el Nevado del Ruiz, Nevado del Quindío y Nevado del Tolima, y ubicado en el extremo norte encontramos la Serranía de San Lucas con una extensión de 16.000 km². La cordillera Oriental se separa gradualmente hacia el este, creando la cuenca del río más importante de Colombia, el Magdalena. Esta cordillera se extiende hacia el noroeste y alcanza su mayor altitud en 5.000 m, formando la Sierra Nevada del Cocuy antes de entrar al

territorio venezolano. Hacia el sur, sobre el piedemonte andino encontramos la serranía de la macarena que es el punto de encuentro de los ecosistemas andinos, de la amazonia y la Orinoquia. [34]

Las tres cordilleras tienen picos principalmente de formación volcánica de más de 4.000 msnm. Muchos de estos volcanes son activos y han causado destrucción y muertes en el pasado debido a las explosiones de gas y ceniza como también a las avalanchas de hielo y lodo. El occidente del país está sujeto a una mayor actividad telúrica, lo que demuestra la inestabilidad de su naturaleza geológica. Al noroccidente de la cordillera Occidental aparece un sistema montañoso llamado serranía del Baudó, que continúa por el Darién girando al oeste hacia Panamá, al norte encontramos el Nudo de Paramillo y tres ramificaciones de la cordillera, conocidas con los nombre de serranía de Abibel, San Jerónimo y Ayapel [34].

Al suroriente del área de estudio encontramos la Serranía de Chiribiquete con su pico más sobresaliente de 910 m, llamado el Cerro Otaré y la Serranía del Tunahi. Entre la cordillera oriental y central el valle del Magdalena medio, al norte, la llanura del Caribe, al occidente, la llanura del Pacífico, al oriente, la llanura del Orinoco y al sur y suroriente la llanura amazónica. En la Figura 5-2 se puede observar el relieve topográfico del área de estudio.

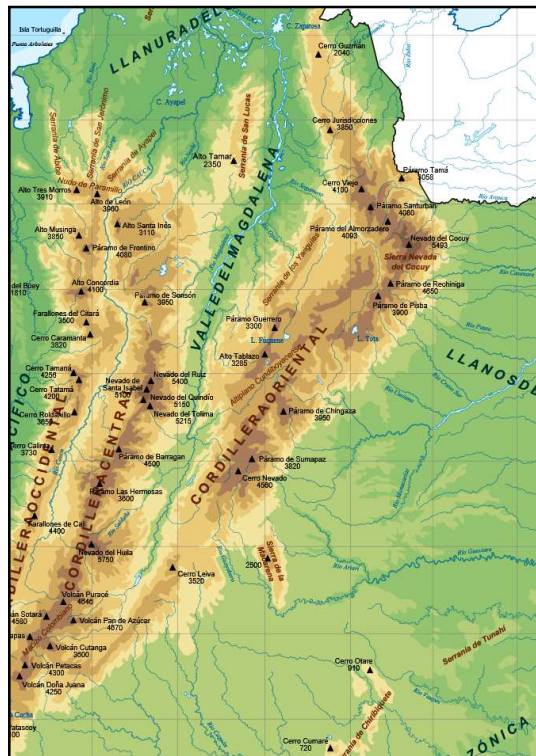


Figura 5-2: Topografía del área de estudio

Fuente: [31]

5.1.3 Hidrografía

La hidrografía en el área de estudio podría ser considerada de las más ricas del mundo, ya que dentro del dominio tenemos influencia de los principales ríos del Colombia, entre ellos el Atrato, el Sinú, el Cauca, el Magdalena, el Meta y el Guaviare. En esta área encontramos algunos de los principales centros hidrográficos del país (donde provienen las principales fuentes de agua dulce), entre ellos, ubicado al suroccidente, el Macizo Colombiano, considerado por la UNESCO Reserva Mundial de la Biosfera y posee más de 60 lagunas, y es el principal conductor de agua dulce de Colombia, donde nacen los ríos Magdalena, Cauca, Patía y Caquetá. Hacia el occidente, el Cerro Caramanta, ubicado a una altura de 3.900 msnm, donde nacen los ríos San Juan, San Juan Antioqueño y Risaralda. Hacia el norte sobre la cordillera Occidental, el Nudo de Paramillo, formando los ríos Sinú, León, San Jorge y Riosucio. De norte a sur, sobre la cordillera Oriental, encontramos el Nudo de Santurbán, donde nace la Sierra de Mérida y los ríos Zulía, Lebrija y otros afluentes del río Catatumbo y Arauca, el Páramo de

5.1.4 Clima

El clima dentro del área de estudio es muy variado debido a las fluctuaciones en altura que ocurren por ser una región de montañas, es así como el régimen de temperatura se caracteriza de acuerdo a los pisos térmicos, que consisten en la disminución de la temperatura a medida que aumenta la altura sobre el nivel del mar. La disminución ocasiona cambios en las condiciones ambientales a medida que se asciende y por consiguiente cambios en los ecosistemas y las formas de vida. Sin embargo el clima se puede clasificar en súper-húmedo en la región del Chocó; húmedo sobre el piedemonte llanero y hacia el norte, y en el departamento de Antioquia; semi-húmedo sobre la cordillera oriental, y árido en la región del valle interandino del Magdalena, en los departamentos de Huila y Tolima.

El principal factor determinante de las precipitaciones es la Zona de Confluencia Intertropical (ZCIT), su fluctuación latitudinal marca las épocas de invierno y verano. Sin embargo, la compleja topografía andina influye también en la cantidad y distribución de las lluvias en la medida en que la dinámica del aire se ve afectada por el relieve, al encontrarse con las cordilleras, las corrientes ascienden, condensándose y generando las llamadas lluvias orográficas. Así mismo, a lo largo del Valle del Magdalena y en cañones de ríos como el Chicamocha, el relieve montañoso circundante sirve como escudo para las corrientes de aire secas que tienden a ascender, generando climas secos y cálidos en las partes bajas, mientras que húmedos y lluviosos en las altas.

El área de estudio es especialmente variable en la distribución espacial y temporal de la precipitación. Algunas zonas, como el Catatumbo, presentan un promedio de 4.000 mm anuales con distribución bimodal, así como también son lugares lluviosos el Magdalena Medio y algunas partes de las Cordilleras Central, Occidental y el Macizo Colombiano; la región de mayor precipitación la encontramos hacia el occidente, sobre la región del Chocó, cuyas precipitaciones superan los 7.000 mm anuales y siguen una distribución mono-modal. Así mismo, se pueden encontrar sectores como el cañón del Chicamocha, el altiplano cundiboyacense, la región de Cúcuta, la Tatacoa en Huila y Candelaria en Boyacá, cuyas precipitaciones no sobrepasan los 1.500 mm anuales. En la figura 5-4 se puede observar mejor la distribución espacial y temporal de precipitación en el área de estudio.

Las precipitaciones máximas en 24 horas están mayormente concentradas en la región del Chocó, con un promedio anual superior a 200mm, de la misma manera la región del piedemonte llanero y amazónico, región del Magdalena Medio y región del Catatumbo. Sobre la cordillera occidental y central un promedio anual superior a 75 mm. Y en algunas regiones de Cundinamarca y Boyacá el promedio anual es apenas superior a 50mm. En la figura 5-4 se puede observar mejor la distribución del promedio anual de las precipitaciones máximas en 24 horas en el área de estudio.

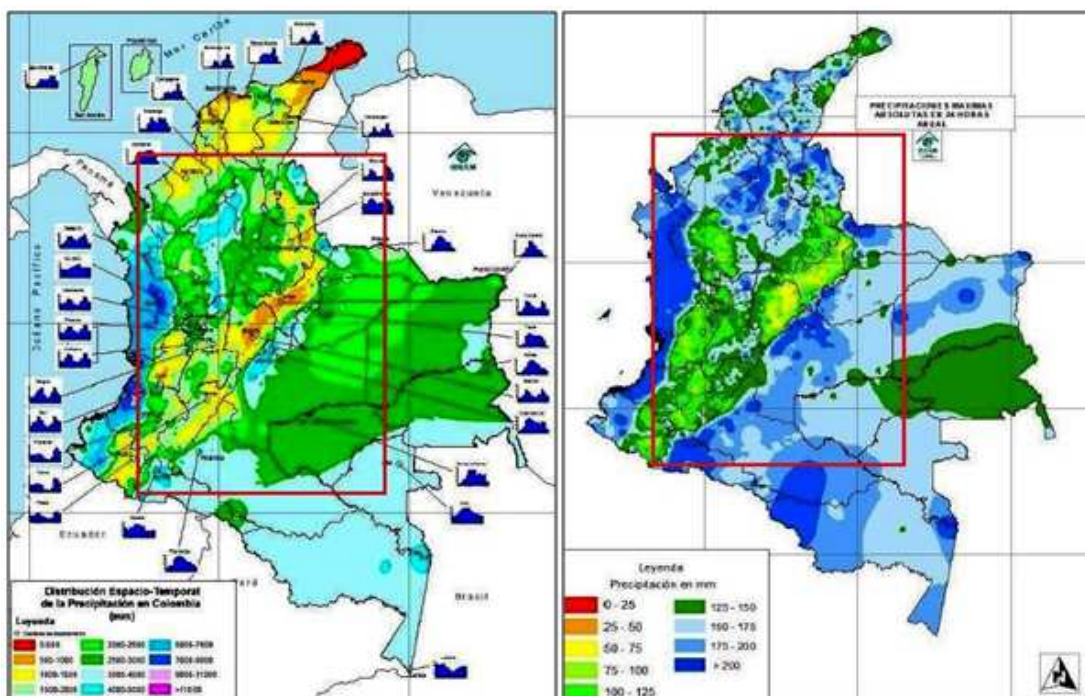


Figura 5-4: Promedio anual de la precipitación (izquierda), promedio anual precipitaciones máximas en 24 horas (derecha)

Fuente: [30]

5.2 Análisis y Control de Calidad de los Datos

Observados

Antes de proceder a cualquier análisis, es necesario garantizar que la información prevista por instrumentos hidrometeorológicos posee un adecuado control de calidad tal como recomienda la Organización Meteorológica Mundial (OMM) [38]. El objetivo es garantizar que las series de tiempo de las distintas estaciones están en realidad mostrando señales asociadas a los sistemas meteorológicos que se presentaron, y no señales que tengan involucrados errores que ocurren durante la adquisición, manipulación, formato, transmisión y archivo de los datos [1].

En la literatura existen infinidad de formas de realizar un control de calidad, y para ello también existe infinidad de software que permiten realizar esta labor, pero particularmente para esta investigación se siguen algunos lineamientos que no se encontraron en los sistemas de software que se encuentran disponibles, por ello se tomó la decisión de diseñar un script en NCAR *CommandLanguage* (NCL), que permitiera realizar un control de calidad de acuerdo a los lineamientos requeridos para esta investigación.

5.2.1 Datos Pluviométricos

Se usaron datos diarios de precipitación, medidos en (mm), de 1.490 estaciones meteorológicas del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) distribuidas dentro del dominio en el que simula el modelo WRF de la FAC, ver Figura 5-5.

Para el desarrollo de este trabajo se usaron datos pluviométricos de estaciones distribuidas en 28 de departamentos, 167 estaciones en Antioquia, 6 en Arauca, 60 en Bogotá, 31 en Bolívar, 126 en Boyacá, 29 en Caldas, 14 en Caquetá, 30 en Casanare, 89 en Cauca, 21 en Cesar, 33 en Chocó, 53 en Córdoba, 171 en Cundinamarca, 3 en Guaviare, 112 en el Huila, 3 en Magdalena, 63 en Meta, 6 en Nariño, 73 en Norte de

Santander, 17 en Putumayo, 3 en Quindío, 23 en Risaralda, 109 en Santander, 15 en Sucre, 111 en Tolima, 119 en Valle del Cauca, 1 en Vaupés y 2 en Vichada, como se puede ver en la Tabla No. 5.2, en donde además, se clasifican las estaciones de acuerdo al tipo de estación, automática (AM), climatológica ordinaria (CO), climatológica principal (CP), meteorológica (ME), pluviográfica (PG), pluviométrica (PM), sinóptica principal (SP) y sinóptica secundaria (SS).

Tabla 5-2: Estaciones meteorológicas usadas en el estudio

TIPO DE ESTACIÓN	ANTIOQUIA	ARAUCA	BOGOTÁ D.C.	BOLIVAR	BOYACA	CALDAS	CAQUETA	CASANARE	CAUCA	CESAR	CHOCO	CORDOBA	CUNDINAMARCA	GUAVIARE	HUILA	MAGDALENA	META	NARIÑO	NORTE DE SANTANDER	PUTUMAYO	QUINDIO	RISARALDA	SANTANDER	SUCRE	TOLIMA	VALLE DEL CAUCA	VAUPES	VICHADA	TOTAL
AM	6	0	0	0	4	0	1	0	0	0	0	2	1	0	2	0	1	0	1	1	0	0	0	0	1	3	0	0	23
CO	28	2	17	4	21	5	7	4	13	2	4	10	25	2	22	0	11	1	12	4	0	6	20	1	21	10	0	0	252
CP	9	0	3	0	10	1	1	4	2	1	0	1	5	0	4	0	3	0	8	0	0	1	6	1	6	2	0	1	69
ME	2	0	4	0	2	0	0	0	0	0	0	0	12	0	0	0	0	0	0	0	0	0	0	0	2	6	0	0	28
PG	14	0	13	0	18	3	0	9	7	0	0	1	28	0	9	0	6	0	2	3	0	0	9	0	12	4	0	0	138
PM	103	4	22	27	71	19	5	13	66	18	29	38	99	1	74	2	40	5	49	9	2	15	72	13	67	90	1	1	955
SP	4	0	1	0	0	1	0	0	1	0	0	0	1	0	0	1	1	0	1	0	1	1	2	0	0	3	0	0	18
SS	1	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	0	0	0	0	2	1	0	0	7
Tot	167	6	60	31	126	29	14	30	89	21	33	53	171	3	112	3	63	6	73	17	3	23	109	15	111	119	1	2	1490

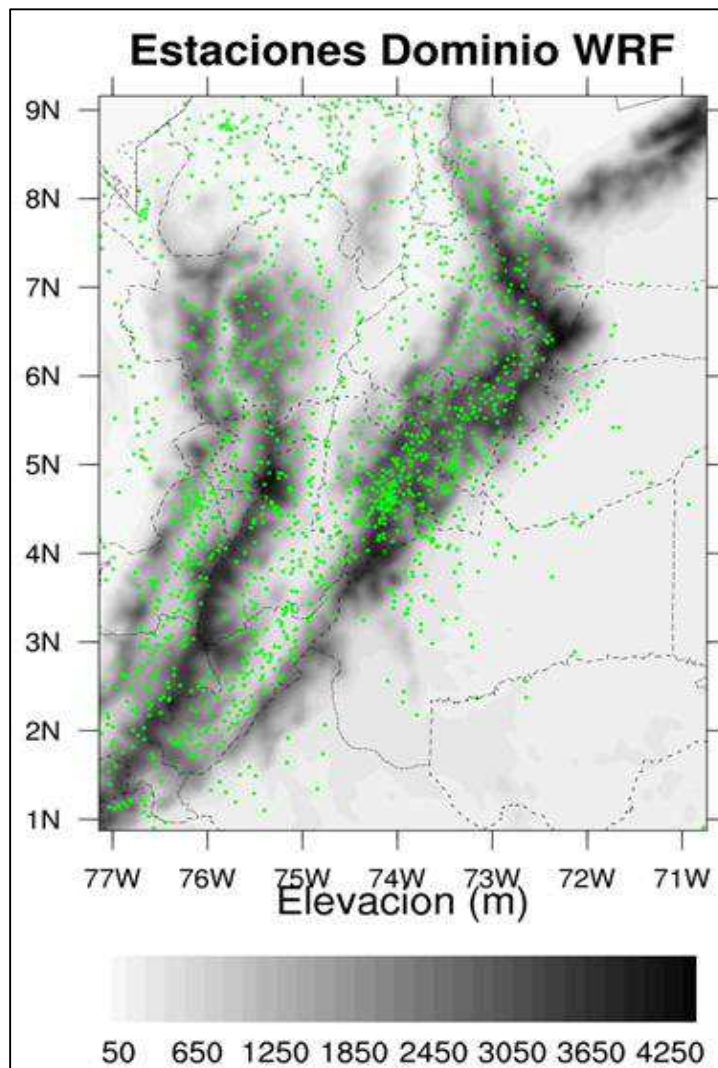


Figura 5-5: Distribución espacial de la estaciones en el dominio del WRF.

5.2.2 Análisis y Control de Calidad

La información de precipitación suministrada por la Red Pluviométrica Nacional del IDEAM fue sometida a un test de calidad en el que las 1.490 estaciones dentro del dominio fueron evaluadas bajo los siguientes lineamientos:

- Se verifica que la información de cada estación no contenga datos de precipitación inferiores a cero, "la serie no puede contener precipitaciones negativas"

- Se verifica que la información no contenga datos registrados con fechas iguales, *“duplicidad en las fechas”*
- Se verifica que la información no contenga datos con precipitaciones iguales durante cinco días consecutivos, esta verificación se realiza debido a que en muchos casos el observador omite la lectura pluviométrica y realiza el mismo registro del día anterior. *“duplicidad en los datos”*
- Se verifica que la información no contenga datos superiores ± 5 desviaciones estándar *“identificación de datos atípicos”*

Al realizar el test se obtuvo que, 917 estaciones presentaron problemas con el test de calidad, 0 estaciones presentaron datos negativos, 0 estaciones presentaron datos con fechas duplicadas, siete (7) estaciones presentaron datos duplicados en cinco días consecutivos, 916 estaciones presentaron datos atípicos superiores a ± 5 desviaciones estándar, como se muestra en la Figura 5-6

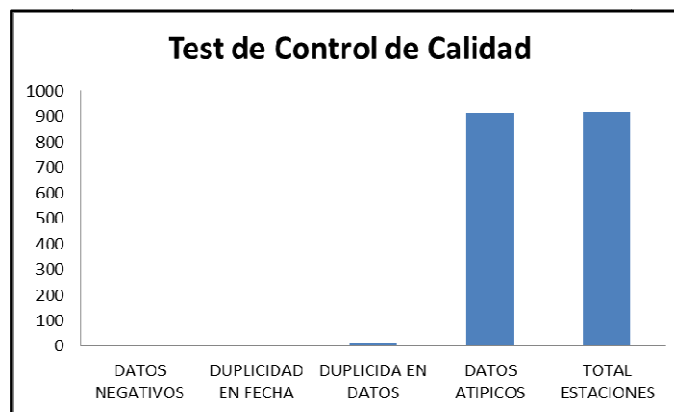


Figura 5-6: Estaciones con que no cumplen el test de calidad

Una vez realizado el test de calidad, se realiza una verificación individual de cada estación problema, mediante dos procedimientos de verificación visual:

En el primero se realiza un análisis estadístico descriptivo de las estaciones que presentaron duplicidad en los cinco días consecutivos, para ello se realizó un script en NCL llamado *“Analis_Datos”* que permitió visualizar la serie temporal de los datos con su

media, la variación de los datos respecto a la media y al Oceanic Niño Index (ONI), la distribución de los datos en un histograma de frecuencias, la distribución de los datos en un Boxplot, un resumen de la estadística descriptiva de los datos y además una gráfica de la serie temporal de las precipitaciones presentadas en las estaciones vecinas en un rango de 10 km.

Una de las estaciones que presentó duplicidad fue la estación con ID. "26150150", el resultado del análisis se puede ver en la Figura 5-7 y 5-8. En la Figura 5-8, se señala en círculos rojos la poca correlación que existe con las estaciones vecinas en las fechas 23 al 29 de abril y 22 al 26 de mayo (fechas con duplicidad en el dato registrado), los resultados de las 7 estaciones con duplicidad se pueden encontrar en el Anexo A.

Después del análisis individual se determinó que en 3 de las 7 estaciones es necesario modificar la serie de precipitación y asignar valores como no observados en donde presentaron problema, ya que los datos no tienen ningún tipo de correlación con las estaciones vecinas.

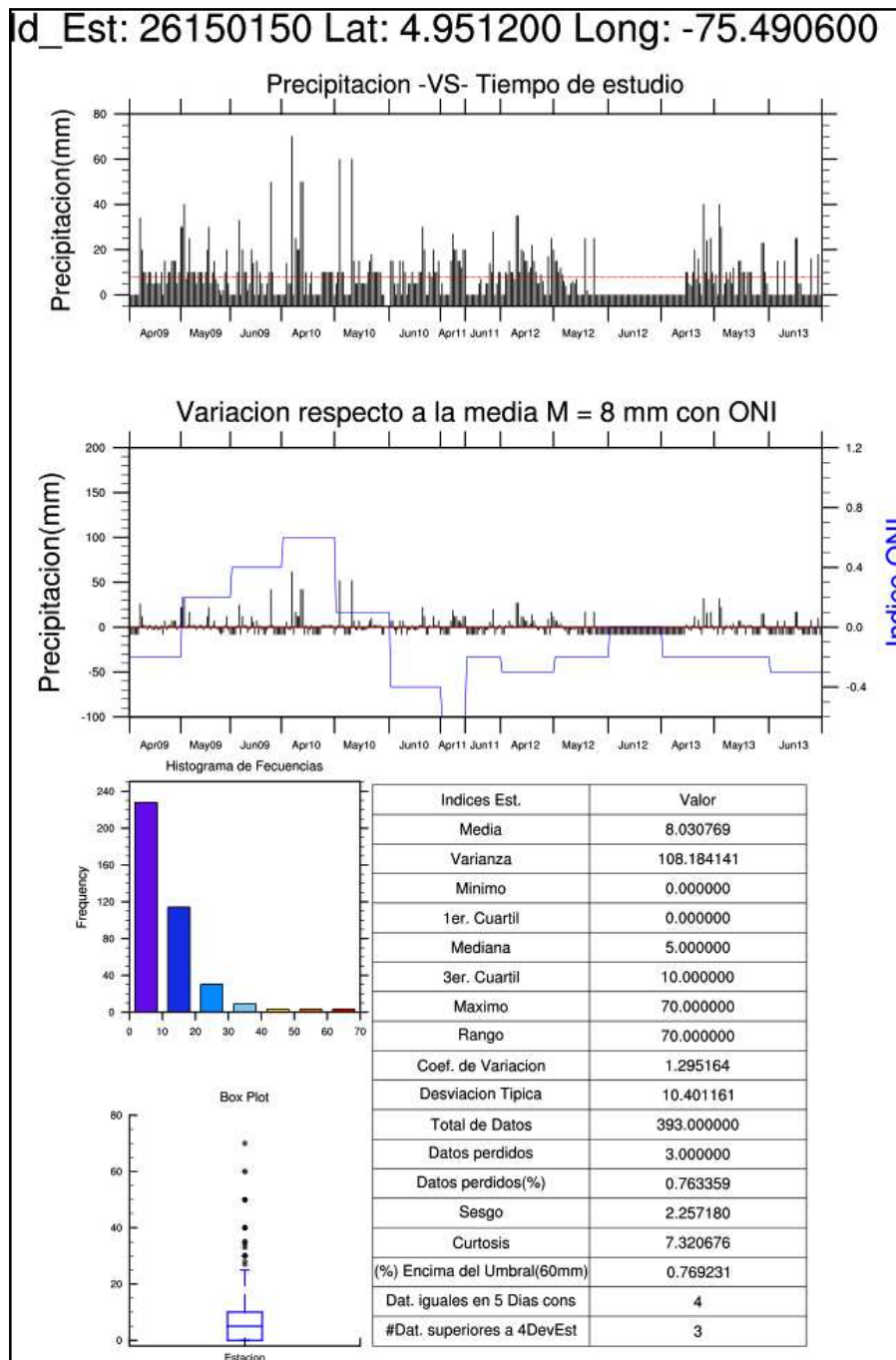


Figura 5-7: Análisis de datos

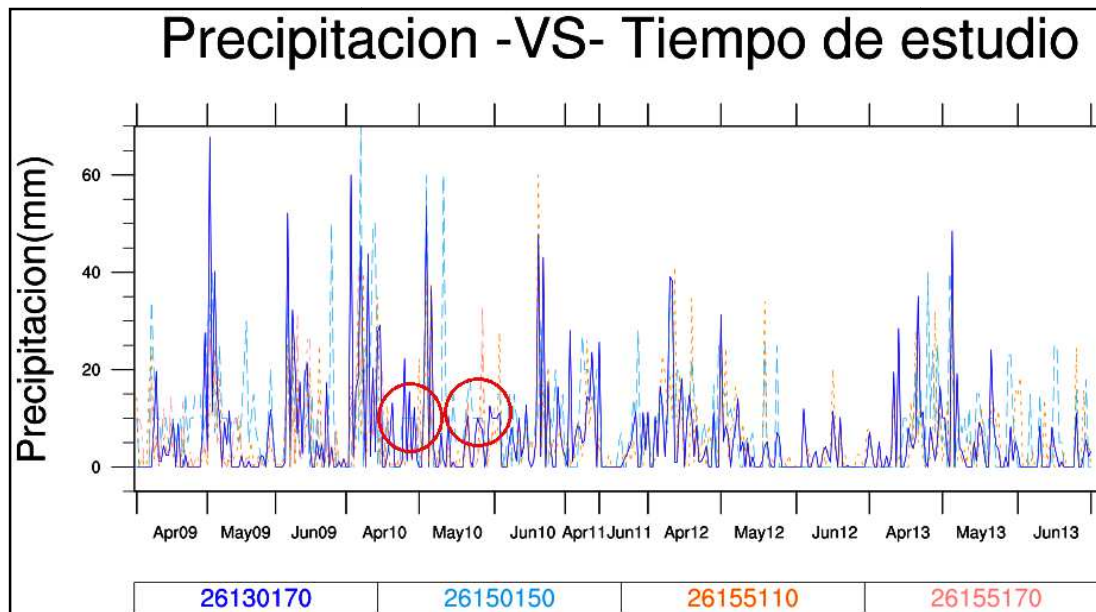


Figura 5-8: Análisis de datos, en círculo rojo los días en los que existe duplicidad de datos en 5 días consecutivos.

En el segundo procedimiento de verificación, se realiza un script en NCL llamado (*"Estaciones_ConSd"*), este script permitió modificar automáticamente la serie de precipitación de las estaciones que presentaron datos atípicos, es decir datos que se encuentren por encima de ± 5 desviaciones estándar, en el script se identifican los datos atípicos y se modifica ese dato siempre y cuando, por lo menos una de las estaciones vecinas presenten al menos el 50% de precipitación del dato atípico.

En la Figura 5-9 se muestra un ejemplo en el que se puede observar gráficamente los valores de precipitación durante el día 14 de abril de 2010. En él se presentaron 6 datos atípicos, mostrados en color rojo dentro de los círculos rojos. Una vez identificados, se verifica que por lo menos una de las estaciones cercanas al dato en un rango de 10 km, presente el 50% de la precipitación del dato atípico, si no se cumple el criterio es modificado ese dato como un dato faltante, de lo contrario continua con su dato de precipitación.

Este control de calidad permitió contar con datos reales en las series temporales de precipitación. Lo que a su vez mejora la interpolación de los datos observados, haciéndolos más consistentes y aproximados a la realidad durante cada día de estudio.

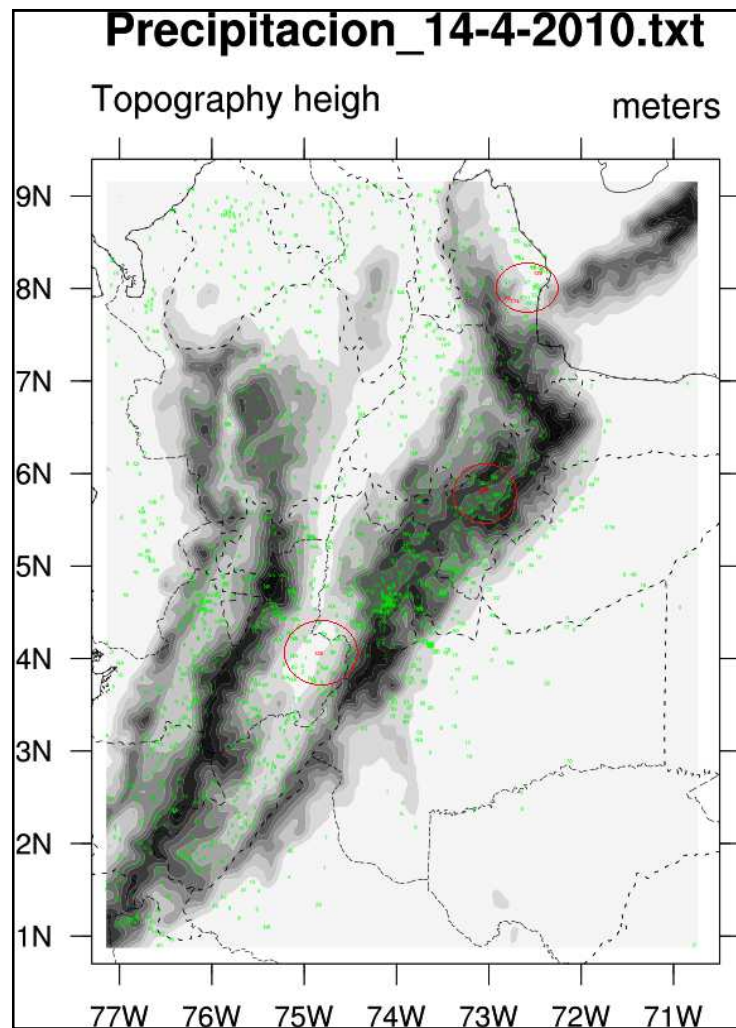


Figura 5-9: Datos de precipitación atípicos (rojo) y datos no atípicos (verde).

5.3 Verificación Método de Interpolación

Para poder realizar una comparación del dato observado contra el simulado en cada punto de grilla en el que el modelo WRF simula, es necesario llevar los datos de precipitación de las estaciones a cada punto de grilla del modelo, esto se realiza mediante un método de interpolación. Para efectos de esta investigación se ha

seleccionado la metodología de interpolación de Cressman por su simplicidad y los buenos resultados obtenidos.

La metodología de Cressman [23] involucra una aproximación iterativa para la elaboración del análisis objetivo, cuyo propósito es la disminución del error en cada iteración. Esto se logra empleando secuencialmente radios de influencia cada vez menores, como se puede ver en la Figura 5-10. En cada paso, un nuevo valor se calcula para cada punto de la malla basado en un factor de corrección que se determina en relación a cada estación dentro del radio de influencia. Para cada estación el error se define como la diferencia entre el valor de la estación y el valor actual del análisis en ese punto (una interpolación desde la malla a las coordenadas de la estación). El esquema original de Cressman emplea un factor de peso para la corrección que es una función radial cuadrática, sin embargo modificaciones posteriores involucran variedad de funciones de peso, especialmente Gaussianas. Estos factores de corrección son aplicados a cada celda de la malla antes de pasar a la siguiente iteración. Si una celda no posee estaciones dentro del tercer radio especificado, se asume dato faltante para dicha celda.

Para verificar la calidad de la interpolación se realiza un script NCL llamado "Verifi_Interpolacion_PorDia", el cual realiza la interpolación de los datos de precipitación de las estaciones mediante el método de Cressman, además, permite obtener una visualización en líneas de contorno de los datos interpolado y ubica espacialmente el valor de precipitación de cada estación. Como se puede ver en la Figura 5-11 existe una relación directa entre las líneas de contorno y los datos pluviométricos de las estaciones, lo que garantiza que el método es consistente y arroja buenos resultados en la precipitación interpolada. En la Figura 5-11 se observa en color magenta, las líneas de contorno de la precipitación interpolada mediante el método Cressman y en color verde el valor de precipitación de la estación meteorológica.

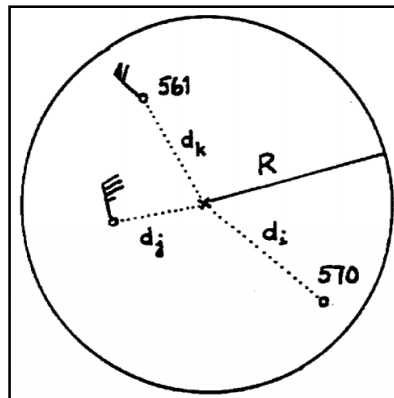


Figura 5-10: Análisis iterativo del método de Cressman

Fuente: [23]

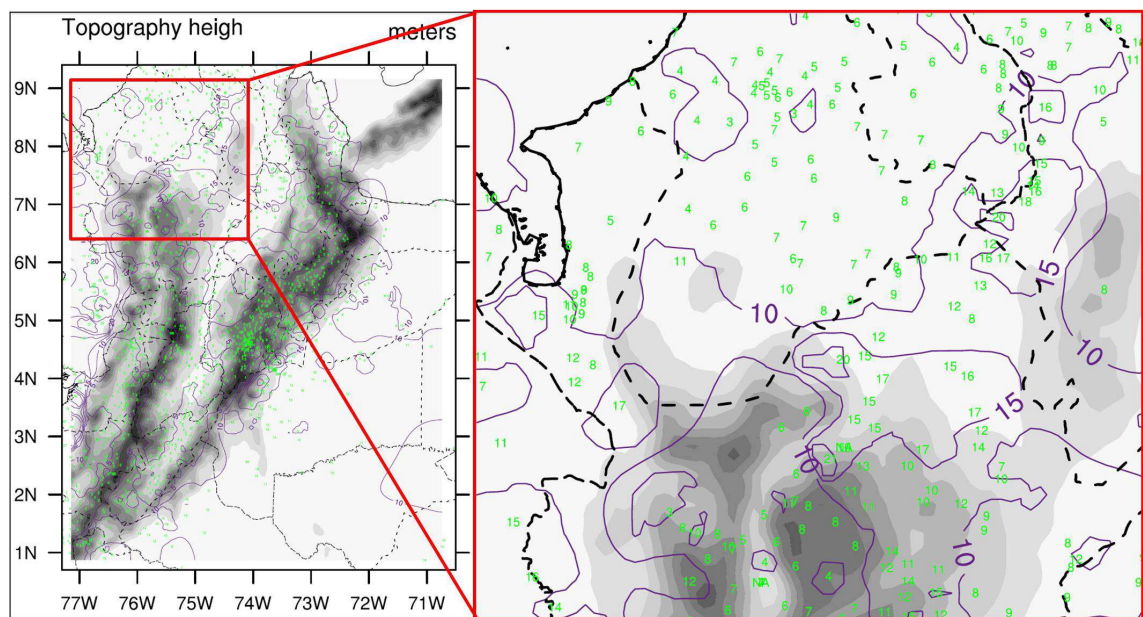


Figura 5-11: Interpolación de Cressman para el día 30 de junio de 2013.

Fuente: El autor

5.4 Simulaciones del Modelo WRF

Las simulaciones que se evalúan, fueron generadas con el modelo de pronóstico numérico Weather Research and Forecasting (WRF), implementado en la Fuerza Aérea Colombiana (FAC) en su versión 2.1, operativo desde el año 2008. La versión del núcleo dinámico (ecuaciones de modelo) es el Advanced Research WRF (ARW).

El modelo consta de tres módulos principales:

- **WRFSI:** Inicializa la grilla del modelo con las especificaciones del usuario y datos iniciales y de frontera provenientes de algún modelo global de circulación atmosférica además de datos de topografía, uso de suelo y suelos.
- **3DVAR:** Modulo para la alimentación del modelo datos de medición meteorológica. De esta forma se ajustan las condiciones iniciales del modelo global a las condiciones observadas en situ.
- **WRF:** El modulo con el mismo nombre del sistema total, realiza los cálculos numéricos para el pronóstico mediante un gran número de esquemas que describen los múltiples sistemas atmosféricos.

El modelo consta de la posibilidad de simular paralelamente hasta 8 grillas (Dominios), una insertada en la otra. La configuración usada para las simulaciones del WRF en esta investigación consta de tres dominios a diferentes resoluciones.

5.4.1 Condiciones Marco

Condiciones que debe respetar la configuración del modelo:

- El modelo debe alimentar las condiciones iniciales con datos de observaciones, por lo tanto las simulaciones de prueba deben incluir el módulo 3DVAR.
- El modelo debe simular mínimo tres dominios insertados.

- La distancia horizontal entre los puntos de grilla debe ser 45km, 15km y 5km para cada uno de los tres dominios respectivamente.
- Cada grilla del modelo debe contener 27 niveles en la vertical.
- La extensión de los tres dominios debe ser:
 1. Dominio: Desde -45° a -100° longitud
 2. Dominio: Todo el territorio nacional incluyendo la isla de San Andrés y Providencia
 3. Dominio: El área andina
- El tiempo de cómputo para la simulación de 24 horas incluyendo la captura de los datos de entrada para este periodo no debe sobrepasar 4 horas.

5.4.2 Ajustes del Modelo WRF

▪ Parámetros del "namelist.input"

Los parámetros del *namelist* más importantes son mencionados a continuación junto con su descripción, si se desea visualizar todos los parámetros ver Anexo B.

`intervalseconds = 10800`

Intervalo en segundos correspondiente al intervalo de tiempo entre los archivos de los datos de entrada proveniente del modelo global GFS. El intervalo es de tres horas.

El parámetro "true" significa que los datos de frontera para el primer dominio provienen del modelo GFS. Los parámetros "false" tienen como efecto que el segundo y el tercer dominio recibe las condiciones de frontera del primer y segundo dominio respectivamente.

historyinterval = 180, 180, 180,

Especifica el intervalo entre las horas para el cual los archivos de salida del modelo deben presentar datos.

frames per outfile = 8, 8, 8,

Indica el número de intervalos de tres horas que un solo archivo de salida debe contener. La duración del intervalo se fija con el parámetro historyinterval. De esta forma se logra separar los datos de salida en diferentes archivos. En el presente caso cada archivo contiene ocho intervalos de tres horas (historyinterval=180) para así obtener archivos de salida correspondientes a intervalos sucesivos de 24 horas de los cuales cada archivo contiene un solo intervalo de 24 horas.

ioformhistory = 2

Los datos de salida se escriben en formato "netcdf".

timestep = 240,

Especifica el paso de tiempo del primer dominio en segundos. Este parámetro es decisivo sobre la estabilidad numérica del modelo.

svert = 1, 1, 1,

evert = 27, 27, 27,

El modelo consta de 27 niveles sigma. Con este valor bajo las extensiones y resoluciones de los dominios dados se logra terminar una simulación de 72 horas en aproximadamente 5:20 horas.

feedback = 1

Habilita la retroalimentación de un dominio de alta resolución al dominio de menor resolución ("Dominio Madre") en el cual se encuentra insertado. Datos de la grilla de mayor resolución reemplazan los datos en el dominio madre donde los puntos de grilla del dominio madre están dentro del área del dominio insertado.

mpphysics = 0, 3, 6,

Especifica el esquema de microfísica aplicado en cada dominio. Esta variable se especifica durante las fases del ajuste cualitativo y cuantitativo mediante los resultados de simulaciones de pruebas del modelo. Los esquemas de microfísica se aplican solamente en el segundo y tercer dominio donde la resolución espacial horizontal es de 15 km y 5 km respectivamente. La resolución de 45 km del primer dominio no justifica la aplicación de un esquema de microfísica. El paso de tiempo relacionado con esta resolución de 240 segundos es más grande que el tiempo de vida de la gran parte de los procesos que simulan los esquemas microfísicas. De esa forma no se pueden esperar mayores mejoras aplicando microfísica en el primer dominio pero sí aumentos en el tiempo de cómputo ya que estas rutinas son muy dispendiosas en sus cálculos.

cuphysics = 2, 1, 0,

Determina el esquema de convección de cúmulos. Estos esquemas se aplican solamente para el primer y segundo dominio. Estos esquemas asumen resoluciones de 25 km o menores por lo cual no se aplican en el tercer dominio con una resolución de 5 kilómetros. Existen cuatro esquemas a la elección: Kain-Fritsch-ETA, Betts-Miller-Janjic, Grell-Devenyi y el antiguo esquema de Kain-Fritsch como se usa en el MM5. El esquema de Grell-Devenyi se basa en el esquema de Arakawa considerando un conjunto de nubes. Este esquema se diseñó para escalas sinópticas y globales (mayor a 100km) y no se aplica en la configuración del modelo. En lugar del esquema Kain Fritsch antiguo se usa el Kain-Fritsch-ETA el cual trae varios ajustes.

5.4.3 Uso del Modelo WRF en la investigación

Para la presente investigación se utilizaron las corridas generadas por el modelo numérico de predicción meteorológica WRF "*WeatherResearch and Forecasting*" instalado en la Fuerza Aérea Colombiana (FAC) en su dominio No. 3. El cual está configurado con límites espaciales en las coordenadas (9° 9'28"N 77° 8'35"W), (9° 9'28"N 70°44'19"W), (0°52'17"N 70°44'19"W) y (0°52'17"N 77° 8'35"W), como se puede observar en la Figura 5.1, con una distancia de puntos de grilla de 5 X 5 km. Esta cobertura geográfica presenta una topografía compleja que incluye tres ramificaciones (occidental, central y oriental) de la Cordillera de los Andes y el Valle del Magdalena.

El modelo implementado en la FAC realiza dos corridas al día, una a las 00:00 UTC y la otra a las 12:00 UTC, en esta investigación solo se tomó las corridas de la 00:00 UTC. La configuración del Dominio No. 3, usa para la modelación de la precipitación la parametrización en microfísica No. 6, correspondiente a “The WRF single-moment 6-class microphysics scheme (WSM6)” el cual contempla el hidrometeoro “Graupel” en la formación de la precipitación.

El pronóstico de precipitación de lluvias, que se desea validar en esta investigación, es generada por el modelo hasta 72horas adelante, de las cuales solo serán utilizadas las primeras 36. Las primeras 12 horas de pronóstico son desechadas, debido a que el modelo emplea este tiempo para calibración y ajuste de las condiciones iniciales y de frontera, este procedimiento se denomina “Spin-Up”. Las siguientes 24horas, es decir las comprendidas entre la hora 12 y 36 de pronóstico son utilizadas para cuantificar la precipitación acumulada en 24 horas (de las 12:00 a las 12:00 UTC del siguiente día).

Cada salida del modelo genera un directorio con la fecha del día de simulación, dentro de este directorio se genera tres archivos, nombrados con el día correspondiente de pronóstico, es decir si el modelo corrió el día 2010-04-12, crea un directorio con ese nombre y dentro él, genera tres archivos con los nombres 2010-04-12, 2010-04-13 y 2010-04-14. Cada archivo tiene un peso de 450 MB y contiene las simulaciones de 24 horas. Como ya se mencionó, solo se usó para esta investigación la hora 12 y 36 del pronóstico, si se aplica al ejemplo anterior, solo se usan los archivos 2010-04-12 y 2010-04-13. Los archivos de salida se encuentran en formato netCDF, sin extensión, por lo que fue necesario agregar a cada archivo de salida una extensión “.nc” para que los sistemas de posprocesamiento permitieran su manipulación.

El principal objetivo de la investigación es verificar la manera en la que el modelo simula las precipitaciones, por ello, se eligió para el análisis la época de lluvias de los meses de Abril, Mayo y Junio que corresponde a la primera época de lluvias del país, además se eligieron los años 2009, 2010, 2011, 2012 y 2013, años en los que se cuenta con simulaciones del modelo y que además fueron años en los que el país estuvo bajo influencia del fenómeno La Niña, lo que generó fuertes precipitaciones e inundaciones.

La información de las simulaciones del modelo WRF de la FAC reposa en cintas magnéticas, siendo necesario subir cada cinta a un servidor que permitirá acceder y extraer la información de las simulaciones en las fechas mencionadas. Durante los cinco años de análisis, el modelo WRF de la FAC presentó algunas fallas técnicas que impidieron la simulación de algunos días, estos días se relacionan en la Tabla 5-3.

Tabla 5-3: Fechas en las que no se cuenta con información de simulaciones del WRF

MES	2009	2010	2011	2012	2013
ABRIL	-	-	15-30	-	-
MAYO	7, 8 y 11	-	1-31	-	-
JUNIO	21	16	1-10	-	-

El análisis de las simulaciones se realizó durante 393 días con los que se cuenta con información, correspondientes a los meses de abril, mayo y junio, de los años 2009, 2010, 2011, 2012 y 2013

5.5 Metodología de Validación

Como se describió en el capítulo 3 y 4, existen diferentes métodos para verificar pronósticos meteorológicos, para efectos de esta investigación se busca comparar las Técnicas de Estadística Tradicional con La Técnica Orientada a Objetos e identificar la ganancia en exactitud de una contra otra, por ello se ha dividido la metodología de verificación en tres partes, la primera consiste en una verificación mediante técnicas de estadística tradicional con variable continua, la segunda en técnicas de estadística tradicional con variable discreta o tablas de contingencia de 2 x 2 y la tercera en técnicas orientadas a objetos con variable discreta o tablas de contingencia 2 x 2.

La herramienta de visualización empleada para los cálculos de cada metodología de verificación fue el *NcarCommandLanguage* "NCL", ya que es una herramienta amigable, cuenta con infinidad de recursos para la visualización, permite una fácil y eficiente manipulación de grandes volúmenes de información y es especialmente usada para investigaciones científicas alrededor del mundo. En ella, fueron escritos los scripts que realizan las tres metodologías mencionadas.

Para ello, se tomaron datos observados y simulados por el WRF. Los datos observados, fueron tomados después someterlos al proceso de control de calidad mencionado en la sección 5.2, y se ajustaron en un formato “.txt” que facilitara su manipulación. Los datos simulados, fueron organizados en diferentes directorios, clasificándolos por años y fecha de simulación, para facilitar el acceso a los archivos de simulación en formato “.nc”.

Contando con la información simulada y observada, se implementa cada metodología de verificación. La manera en la que se realizó cada una de estas verificaciones, se describe a continuación mediante diagramas de flujo para facilitar su comprensión.

5.5.1 Técnicas de Estadística Tradicional con Variable Continua

Como se describió en la sección 3.2, las variables continuas nos permiten verificar cuánto difieren las simulaciones de las observaciones. Para ello se emplean algunos índices de verificación, que nos muestran mediante diferentes aproximaciones estas diferencias. Los índices usados son el Mean Error “BIAS” (ecuación 3.4), *Mean Absolute Error* “MAE” (ecuación 3.5), *Mean Squared Error* “MSE” (ecuación 3.6) y Raíz del MSE “RMSE” (ecuación 3.7).

Como se puede observar en la Figura 5-12, el procediendo de verificación se inicia con el ingreso de los datos observado y simulados. Se establece el número de días que se desea verificar, para nuestro periodo de análisis 393 días.

Para cada día de análisis se realiza una interpolación de los datos observados mediante el método de Cressmas, llevándolos a cada punto de grilla del modelo. Con los datos simulados, se resta la hora 36 a la hora 12 de simulación (precipitación acumulada en 24 horas) como se mencionó en la sección 5.4. Teniendo los datos observados y simulados en cada punto de grilla, se realiza una iteración día a día en la que se compara observado y simulado, hasta completar los 393 días de análisis, y finalmente se realiza los cálculos de los índices BIAS, MAE, MSE y RMSE. Al finar el ciclo, los resultados de los análisis son graficados junto con la topografía del terreno, arrojando los resultados que se muestran capítulo 6.

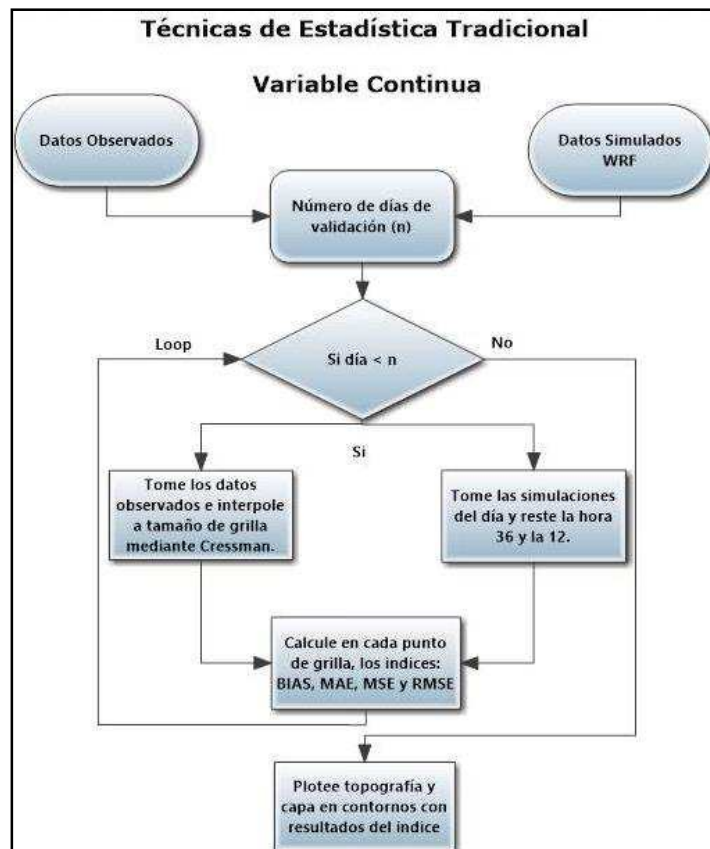


Figura 5-12: Diagrama de flujo para la técnica de estadística tradicional variable continua.

5.5.2 Técnicas de Estadística Tradicional con Variable Discreta

La metodología de verificación mediante técnicas de estadística tradicional con variable discreta, o también llamado, mediante eventos categóricos o tablas de contingencia 2x2; consiste en un enunciado en el cual se afirma que uno y solo uno de los conjuntos de eventos posibles va a ocurrir, para este caso se tomarán eventos de precipitaciones superiores a 60mm diarios, con el fin de identificar precipitaciones intensas. Estos eventos son clasificados de acuerdo a lo mencionado en la sección 3.3.

Como se puede observar en la Figura 5-13, el procediendo de verificación se inicia con el ingreso de los datos observados y simulados. Se establece el número de días de análisis, para este caso 393 días (periodo de análisis). Para cada día de análisis se realiza una interpolación de los datos observados mediante el método de Cressman, llevándolos a cada punto de grilla del modelo. Con los datos simulados, se resta la hora

36 a la hora 12 de simulación (precipitación acumulada en 24 horas) como se mencionó en la sección 5.4.

Con los datos observados y simulados en cada punto de grilla, se realiza una verificación en la que únicamente serán consideradas las precipitaciones superiores a 60mm. Las precipitaciones inferiores a este umbral serán rechazadas, y de esta manera clasificar en cada punto de grilla la ocurrencia de "un evento" o un "no evento" (evento categórico o dicotómico), se realiza esta identificación en ambos campos observado y simulado de manera iterativa para cada día de análisis.

Los resultados almacenados son comparados y de acuerdo a su resultado será contabilizado como un "*Hit*", "*False Alarm*", "*Miss*" o "*CorrectRejection*", llevándolos de esta manera a tablas de contingencia de 2x2. Una vez obtenidos los resultados son calculados los índices BIAS, POD, FAR, CSI y ETS, con las ecuaciones mencionadas en la sección 3.4.3. Los resultados de los análisis son graficados junto con la topografía del terreno, arrojando los resultados que se muestran en el capítulo 6.

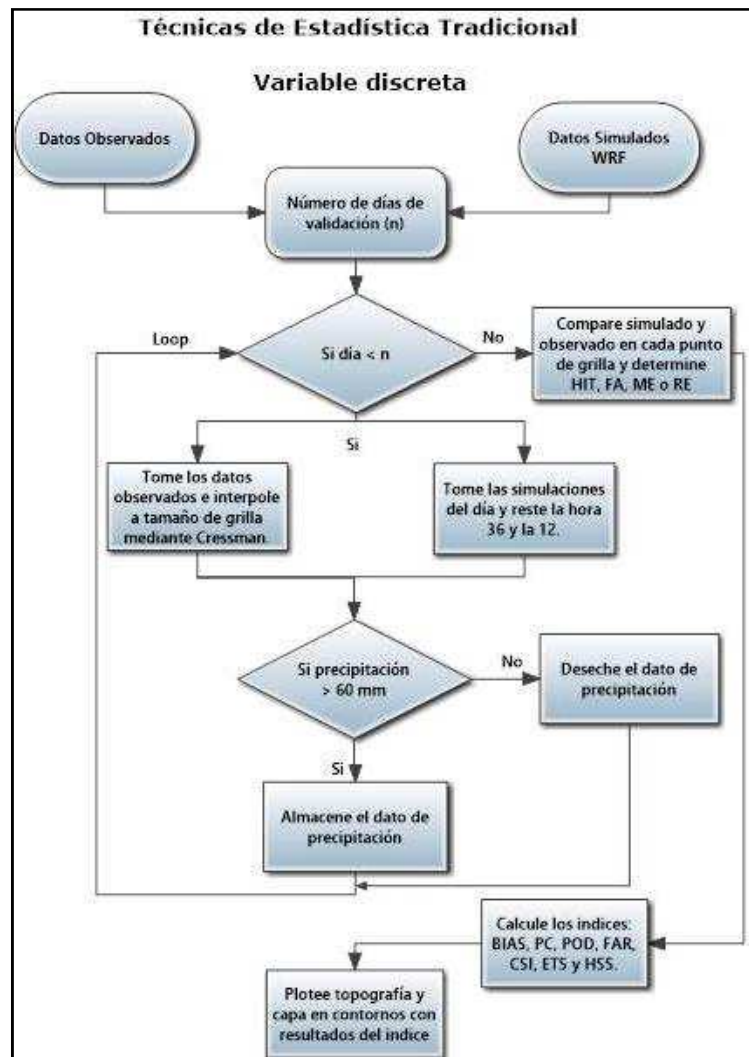


Figura 5-13: Diagrama de flujo para la técnica de estadística tradicional variable discreta.

5.5.3 Técnicas Orientadas a Objetos con Variable Discreta

La metodología de verificación mediante técnicas orientadas a objetos con variable discreta, o también llamado, mediante eventos categóricos o tablas de contingencia 2x2; consiste en la identificación de objetos los cuales son fácilmente distinguibles ante el ojo humano, para nuestro caso de estudio los objetos serán las precipitaciones más intensas que se presenten en áreas de 50x50 km, una vez identificados en ambos campos, observado y simulado, son llevados a tablas de contingencia de 2x2.

Como se puede observar en la Figura 5-14, el procedimiento de verificación se inicia con el ingreso de los datos observado y simulados. Se establece el número de días de análisis, para este caso 393 días (periodo de análisis). Para cada día de análisis se realiza una interpolación de los datos observados mediante el método de Cressman, llevándolos a cada punto de grilla del modelo. Con los datos simulados, se resta a la hora 36 la hora 12 de simulación (precipitación acumulada en 24 horas) como se mencionó en la sección 5.4.

Con los datos observados y simulados en cada punto de grilla, se realiza una verificación que consiste en la identificación de los objetos dentro de cada área de 50 X 50 km, es decir, se busca dentro de esa área las precipitaciones que sean superiores a un umbral de 60 mm y que además sea la mayor precipitación presentada dentro del área (precipitación más intensa representativa del área), una vez identificados en ambos campos tanto simulado como observado, se verifica si cumple los criterios establecidos (>60mm y mayor precipitación en el área), si es así, los resultados son almacenados, de lo contrario son rechazados. De esta manera son clasificados como "un evento" o un "no evento" (eventos categorizados o dicotómicos), y se repite el procedimiento día a día durante el periodo de análisis.

Los resultados almacenados son comparados y de acuerdo a su resultado será contabilizado como un "*Hit*", "*False Alarm*", "*Miss*" o "*CorrectRejection*" dentro del área, llevándolos de esta manera a tablas de contingencia de 2x2. Una vez obtenidos los resultados son calculados los índices BIAS, POD, FAR, CSI y ETS, con las ecuaciones mencionadas en la sección 3.4.3. Los resultados de los análisis se grafica junto con la topografía del terreno, arrojando los resultados que se muestran en el siguiente capítulo.

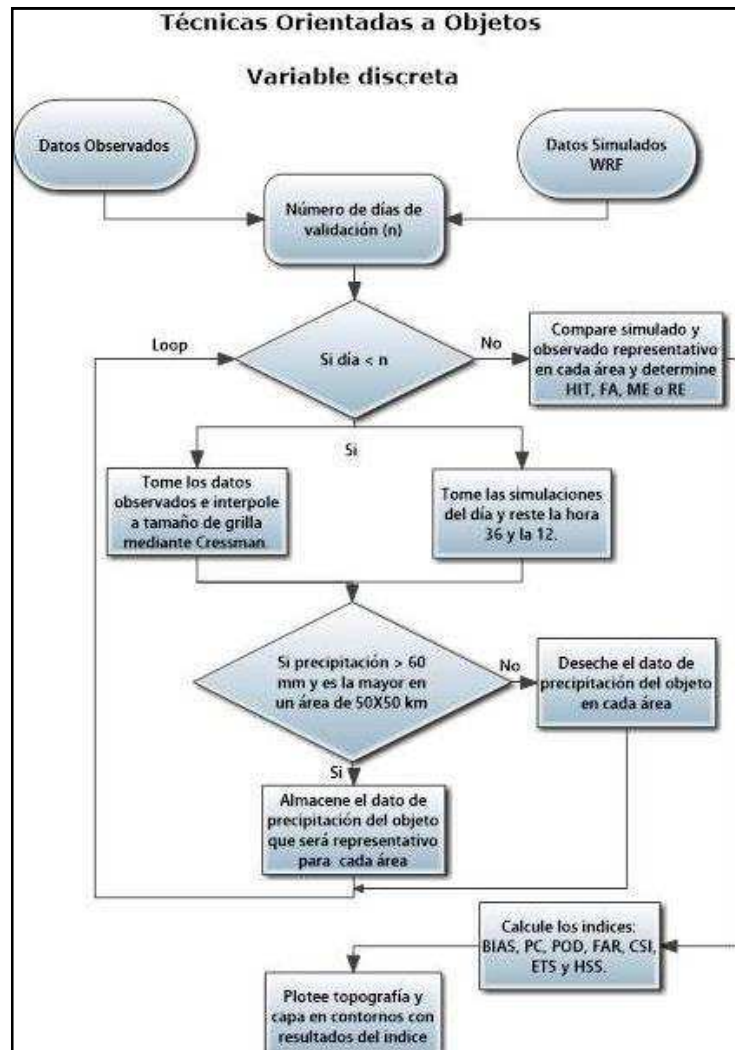


Figura 5-14: Diagrama de flujo para la técnicas orientadas a objetos variable discreta.

6 Análisis de los Resultados

Se presentan los resultados de 393 corridas del modelo WRF y se comparan con los resultados observados correspondientes, el análisis se divide en resultados de las validaciones mediante Técnicas de Estadística Tradicional de variable continua, Técnicas de Estadística Tradicional por tablas de contingencia o análisis categórico (dicotómico) y Técnicas Orientadas a Objetos por tablas de contingencia o análisis categórico (dicotómico), para ver gráficamente todos los índices calculados ver Anexo C. En las graficas que se presentan a continuación fueron graficados algunos de los principales aeropuertos del país con el fin de facilitar un análisis cuantitativo de los resultados arrojados por el índice en cada uno de ellos.

Antes de entrar al análisis de cada metodología de validación se observa gráficamente la ubicación de las estaciones dentro del dominio y la diferencia del promedio de la precipitación simulada por el WRF y la observada por las estaciones. La Figura 6-1 muestra con puntos verdes la ubicación de las estaciones dentro del dominio y se demarca en círculos rojos las regiones en las que existe falencia de datos. La Figura 6-2 muestra la diferencia del promedio de la precipitación modelada y observada, en color magenta se observan valores negativos que denotan subestimación, pues fue pronosticada menos precipitación de la que se observó, en color verde se observan valores positivos que denotan sobrestimación, pues se pronostico más de lo que se observó, en círculos rojos se demarcan las regiones con falencia de datos que casualmente coincide con las regiones en las que existe mayor sobrestimación de la precipitación.

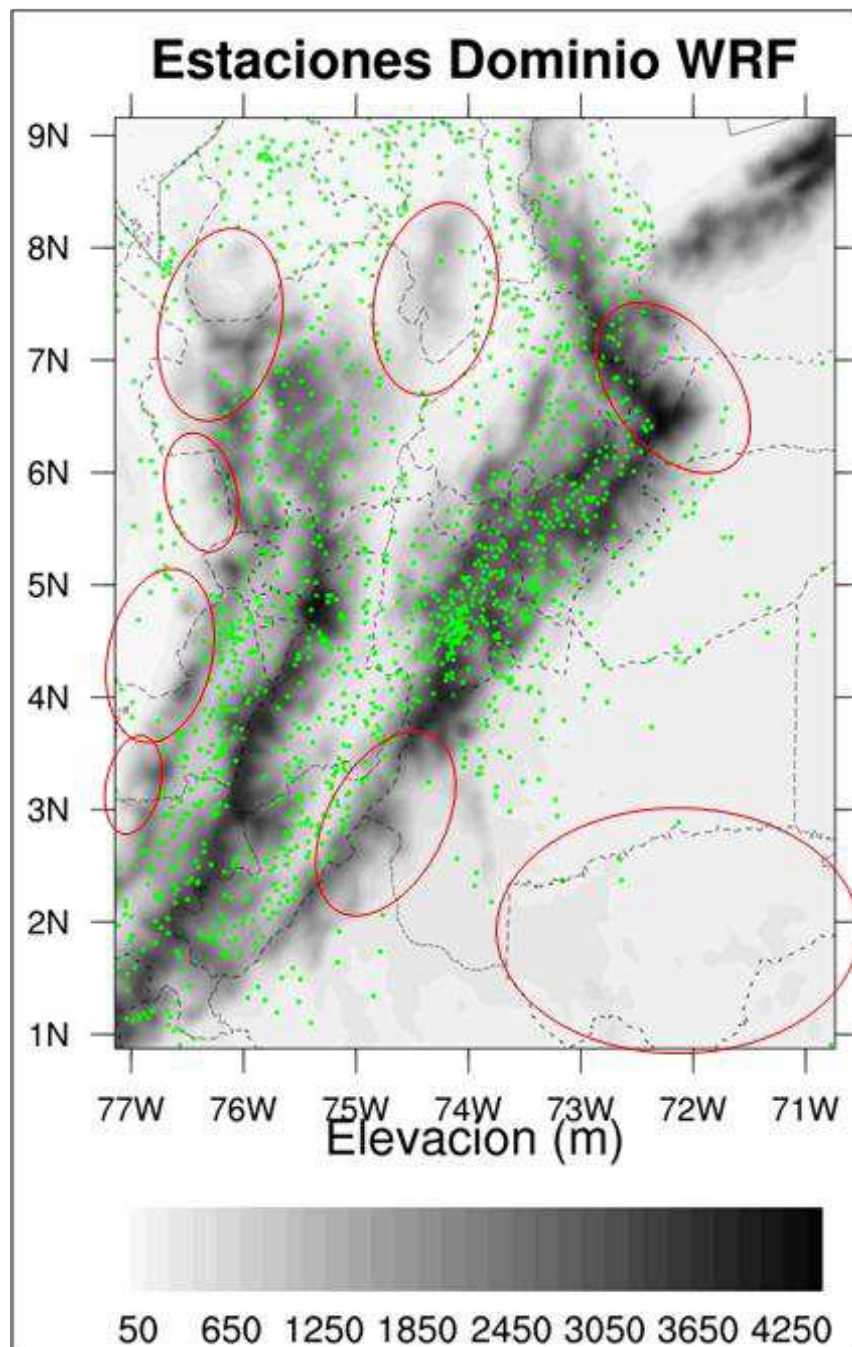


Figura 6-1: Estaciones dentro del dominio (puntos verdes) y regiones con falencia de datos (círculos rojos).

métodos de validación, se encontrarán resultados que se ven afectados por lo mencionado.

6.1 Técnicas de Estadística Tradicional con Variable Continua

En las Técnica de Estadística tradicional (TET) con Variable continua, los resultados de los índices calculados varían en sus resultados de un índice a otro, es por ello que los rangos en las escalas de colores (ubicados a las derecha de cada imagen) son diferentes; debido a esto, fue necesario colocar un valor de referencia que represente el área de análisis, denotando como representativo el punto de grilla ubicado en el centro de regiones de 50 X 50 km.

▪ BIAS y MAE

El BIAS de la estadística tradicional, calculado por la fórmula de la ecuación 3.5, indicando que hay una sobrestimación sobre las tres cordilleras, sin embargo es recomendado analizar los resultados del BIAS en conjunto con el MAE, ya que las discrepancias entre el campo observado y modelado son maximizadas con este último, permitiendo identificar más fácilmente las regiones con sobrestimación. Lo anterior se debe a que en la fórmula del MAE mostrada en la ecuación 3.6, aplica el valor absoluto de la diferencia entre simulado - observado, eliminando valores negativos.

Al analizar en conjunto los resultados de los dos índices (Figura 6-3) se observa que esta sobrestimación es más intensa al oeste de la cordillera occidental; en el noroeste y oeste del departamento de Antioquia, sur de Córdoba, este y centro del departamento del Chocó, oeste de Risaralda, oeste del Valle del Cauca, noroeste del Cauca y norte de Nariño.

También se nota sobrestimación al norte de la Cordillera Central; centro y norte del departamento de Antioquia, norte de Caldas, norte y centro del Tolima.

Sobre la cordillera oriental se nota sobrestimación al noroeste y noreste del departamento de Boyacá, noroeste y sureste de Cundinamarca, departamento de Santander, oeste de Norte de Santander, límites entre el departamento del Casanare y Boyacá, noroeste del Meta y noroeste de Caquetá.

Además se observa sobrestimación sobre el departamento del Guaviare y sureste del Meta, cabe anotar que en esta última apreciación los resultados no son demasiado confiables debido a la falencia de datos observados en la región.

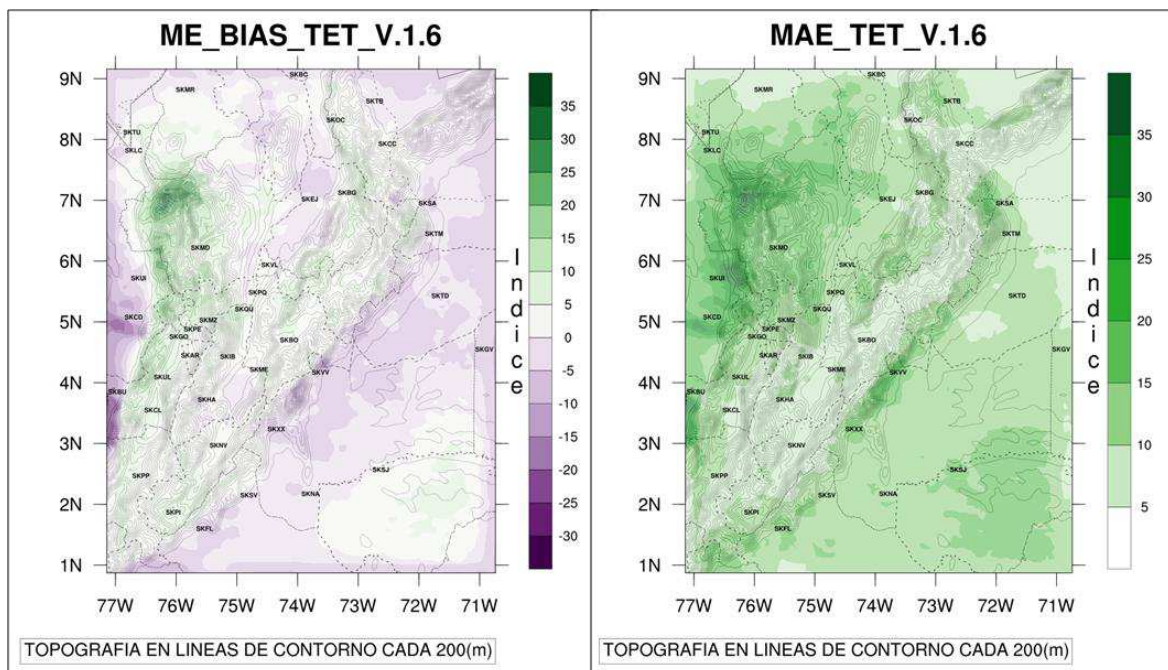


Figura 6-3: Resultado BIAS (izq.) y MAE (der.), TET con variable continua

▪ MSE y RMSE

En la Figura 6-4 (panel izquierdo) se observa el resultado arrojado por el MSE calculado la fórmula de la ecuación 3.7, mediante su análisis se puede observar una sobrestimación mucho más marcada, especialmente en las regiones que presentan datos atípicos (con discrepancia entre simulado - observado), ya que en su ecuación las diferencias son elevadas al cuadrado.

Es así que para este caso de análisis este índice arroja resultados de sobrestimación sobre gran parte de la cordillera occidental; norte, noroeste y oeste de Antioquia, este y centro del Chocó, oeste de Risaralda, oeste del Valle del Cauca y norte de Nariño.

Sobre la cordillera central se observa sobrestimación al norte y este del departamento de Antioquia, sur y suroeste de Bolívar, este de Caldas, y límites entre los departamentos de Santander, Bolívar y Antioquia.

Sobre la cordillera oriental; sobrestimación sobre el noreste del departamento de Boyacá, límites entre Boyacá y Casanare, noroeste de Guaviare, y noroeste y oeste de Caquetá. Además de la sobrestimación sobre el departamento de Guaviare y sureste del Meta.

En la Figura 6-4 (panel derecho) se observan los resultados arrojados por el RMSE calculados mediante la fórmula de la ecuación 3.8. En esta fórmula se aplica la raíz cuadrada a los resultados obtenidos del MAE, el RMSE muestra mayores valores sobre la cordillera occidental y norte de la central, en general sobre el departamento de Antioquia, además de la serranía de San Lucas, piedemonte llanero y amazónico y el departamento de Guaviare.

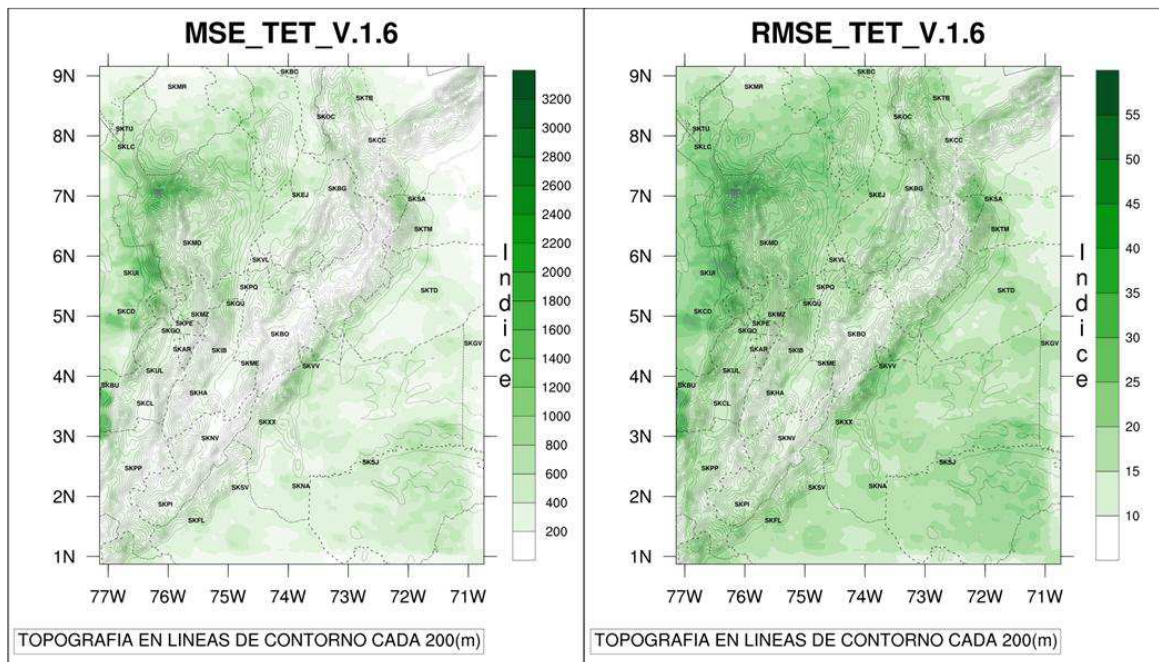


Figura 6-4: Resultado MSE (izq.) y RMSE (der.), TET con variable continua

6.2 Técnicas de Estadística Tradicional con Variable Discreta

En las Técnica de Estadística tradicional (TET) con Variable Discreta, es clasificado en cada punto de grilla la ocurrencia de un evento o no evento (precipitación mayor a 60 mm) en los campos observado y simulado. Una vez clasificados, son llevados a tablas de contingencia de 2 X 2 y son calculados los índices correspondientes cumpliendo los criterios mencionado en la sección 3.4. Los resultados de los índices calculados varían en sus resultados de un índice a otro, y es por ello que los rangos en las escalas de colores (ubicados a la derecha de cada imagen) son diferentes. Debido a esto, fue necesario colocar un valor de referencia que represente el área de análisis y que además facilite la comparación con los resultados de la Técnica Orientada a Objetos de la siguiente sección, denotando como representativo el punto de grilla ubicado en el centro de regiones de 50 X 50 km.

▪ BIAS

El BIAS mediante eventos categóricos se calcula con la fórmula de la ecuación 3.9, sus resultados confirman lo calculado en variable continua, con altos BIAS (sobrestimación) al oeste de la cordillera occidental; puntualmente al oeste de Antioquia, en la región que limita el Chocó con el suroeste antioqueño y con el norte del Valle del Cauca.

También se presenta sobrestimación entre la cordilleras occidental y central; en el departamento del Cauca, al noreste de Nariño, al norte del departamento de Caldas y oeste antioqueño, al norte de Antioquia, entre las serranías de Abibe, San Jerónimo y Ayapel, en límites con Córdoba y al sur del departamento de Córdoba.

Entre las cordilleras central y oriental; al este de la cordillera central, al norte del Tolima, caldas y este antioqueño, norte de Antioquia en límites con Bolívar y sur y centro de este mismo departamento. Al oeste de la cordillera oriental, al noroeste de Cundinamarca, oeste de Boyacá, en la región que limita a Santander con Boyacá y noreste de Santander.

Sobre y al este de la cordillera oriental; en la región que limita Norte de Santander con Cesar y Santander. Al este de la cordillera oriental; en la región que limita Boyacá con Arauca y Casanare, al este de Cundinamarca, al oeste del Meta y en el límite entre Meta y Caquetá, en el límite entre Caquetá y Huila, al noroeste de Guaviare, al este de Caquetá y en la región límite entre Meta y Vichada.

Región norte y centro de la misma, serranía de San Lucas, lado oeste de la cordillera oriental, entre los departamentos de Boyacá y Santander, también es notorio la subestimación sobre el valle interandino del Magdalena. Ver Figura 6-5.

De las regiones marcadas, además del efecto orográfico, la sobreestimación también se debe a la ausencia de observaciones, como lo son: el noroeste de Antioquia, oeste y noreste de de Guaviare y oeste del municipio de Saravena. Por lo tanto se puede concluir que el BIAS no solo está asociado a la topografía, sino también a la falta de datos.

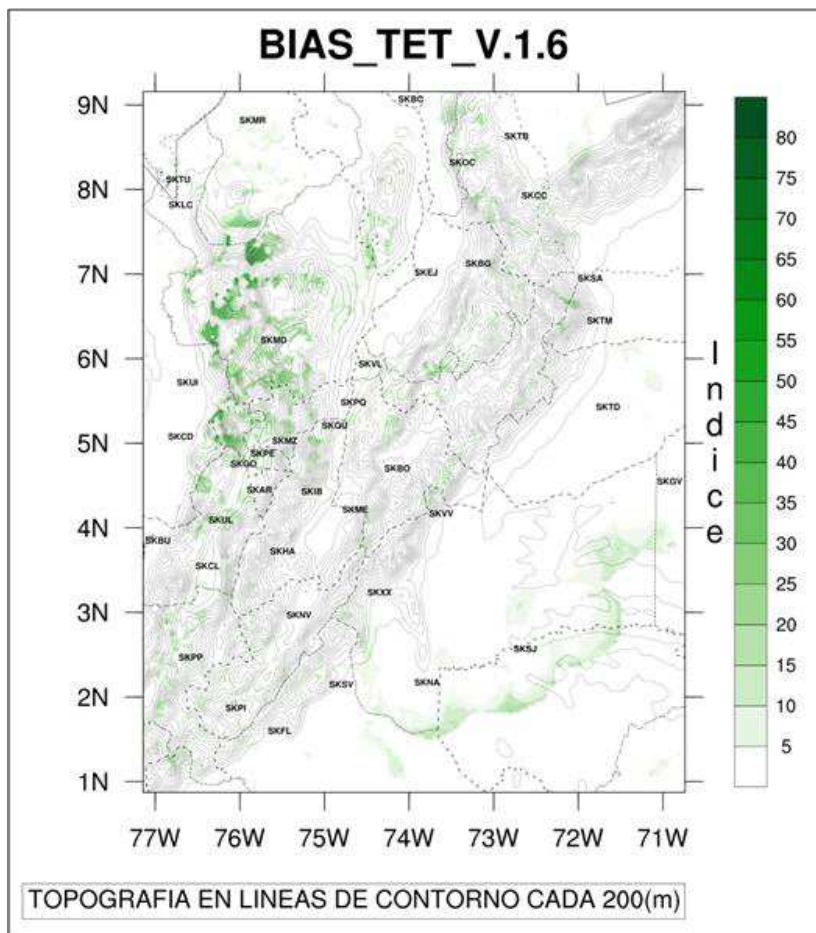


Figura 6-5: Resultados BIAS, TET con variable discreta

▪ **POD y FAR**

El POD, el cual se recomienda ser analizado junto con el FAR, indica que fracciones de los eventos pronosticados fueron correctamente observados, como se puede ver en la Figura 6-6 (panel izquierdo), se encuentra que las regiones con mayor acierto están ubicadas al norte de la cordillera oriental y central sobre el departamento de Antioquia, norte de Valle del Cauca, Risaralda, Caldas, sur de la serranía de San Lucas, oeste de Saravena, norte de Guaviare y este de Meta. Por otra parte el FAR, indica la fracción de las predicciones que no ocurren, arroja resultados contrarios a los que muestra el POD, como se puede ver en la Figura 6-6 (panel derecho).

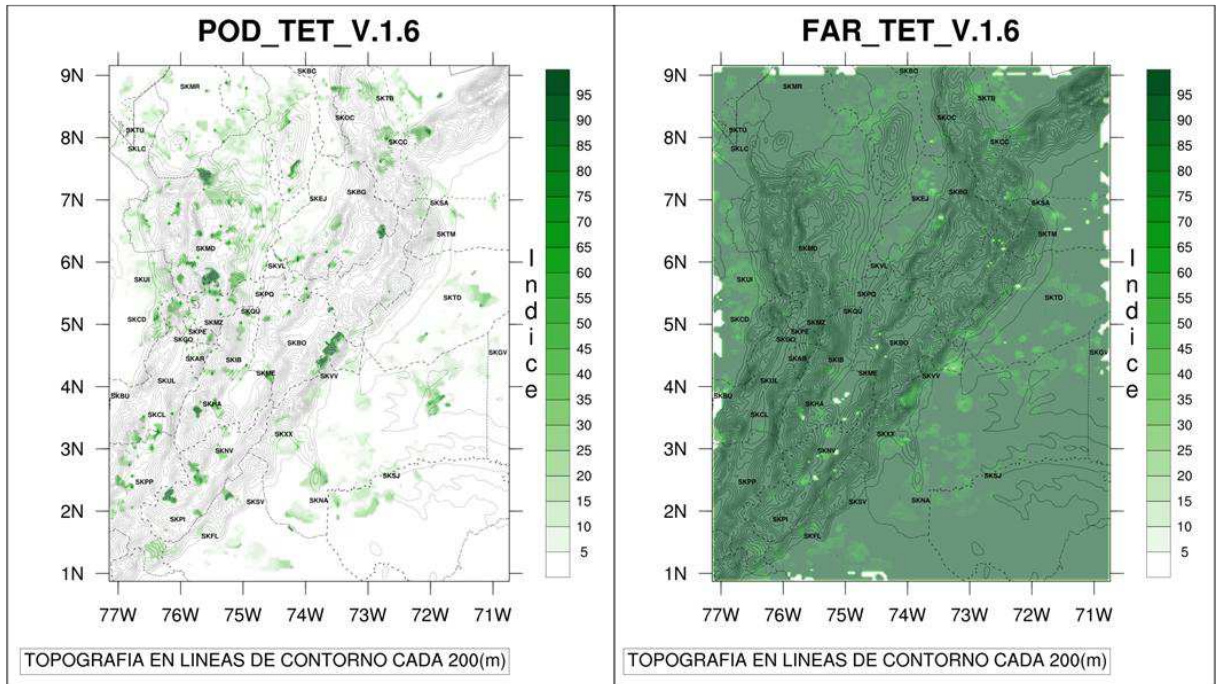


Figura 6-6: Resultados POD y FAR, TET con variable discreta

▪ **CSI**

El CSI marca qué tan bien corresponden los eventos pronosticados con los observados. En la Figura 6-7, se puede ver que las regiones con mayor correspondencia están sobre algunas regiones en los costados este y oeste de la cordillera oriental, sobre el valle interandino del Magdalena y los departamentos de Meta y Guaviare. Los mayores valores de correspondencia se presentaron al oeste del departamento de Tolima.

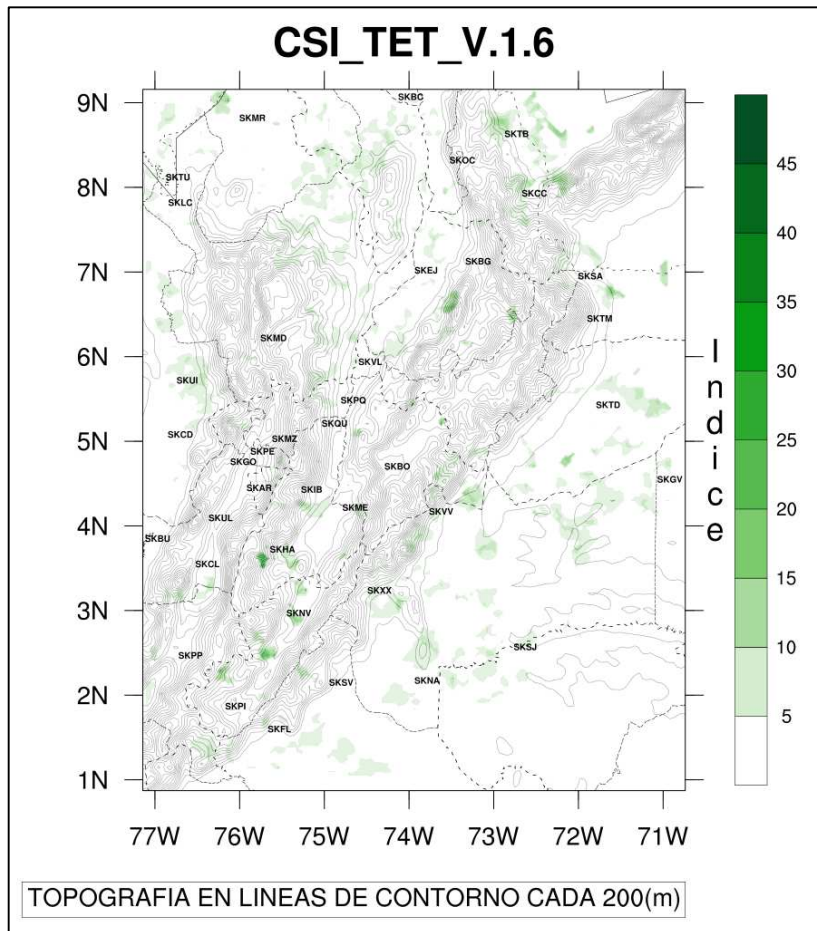


Figura 6-7: Resultado del CSI, TET con variable discreta

▪ ETS

El ETS muestra qué tan bien el modelo hizo el pronóstico de eventos con respecto a los eventos observados, en la Figura 6-8 se puede ver que el modelo expone habilidad para pronosticar sobre toda la cordillera oriental, la central y parte de la occidental, al igual que sobre el departamento de Guaviare y norte de la cordillera Central y Occidental. Mostrando un pico de máxima habilidad al oeste del departamento de Tolima.

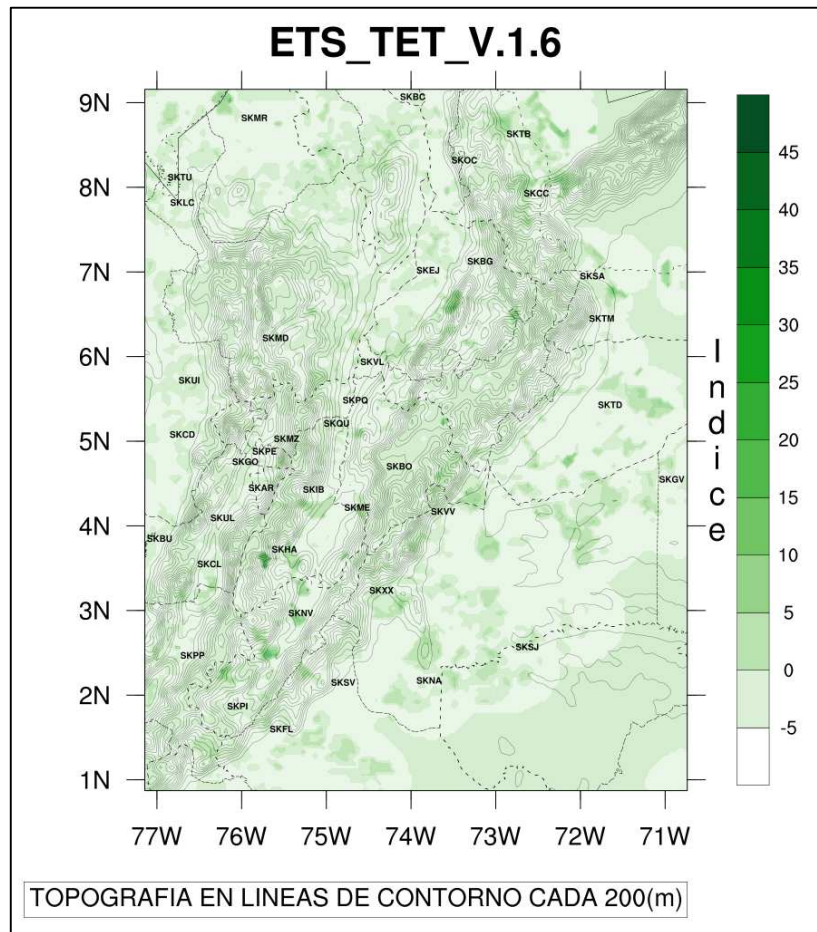


Figura 6-8: Resultado del ETS, TET con variable discreta.

6.3 Técnicas Orientadas a Objetos con Variable Discreta

Para tener una primera impresión sobre los resultados obtenidos en cada índice con las técnicas orientadas a objetos (TOO), inicialmente se realiza un análisis de los resultados arrojados por las tablas de contingencia de manera gráfica, número de “Hits”, “False Alarm”, “Miss o Miss Event” y “CorectRejection o RareEvent”, como se indica en la Figura 6-9. Este análisis permite además verificar si el método se está aplicando correctamente. La tabla de contingencia, por ejemplo, en el punto de Cundinamarca ubicado más al oriente y al sur, arroja los siguientes resultado: Hits=23, False Alarm=37, Miss Event=82, y CorectRejection=251, lo que representa un total de 393 eventos, o sea, los 393 días considerados.

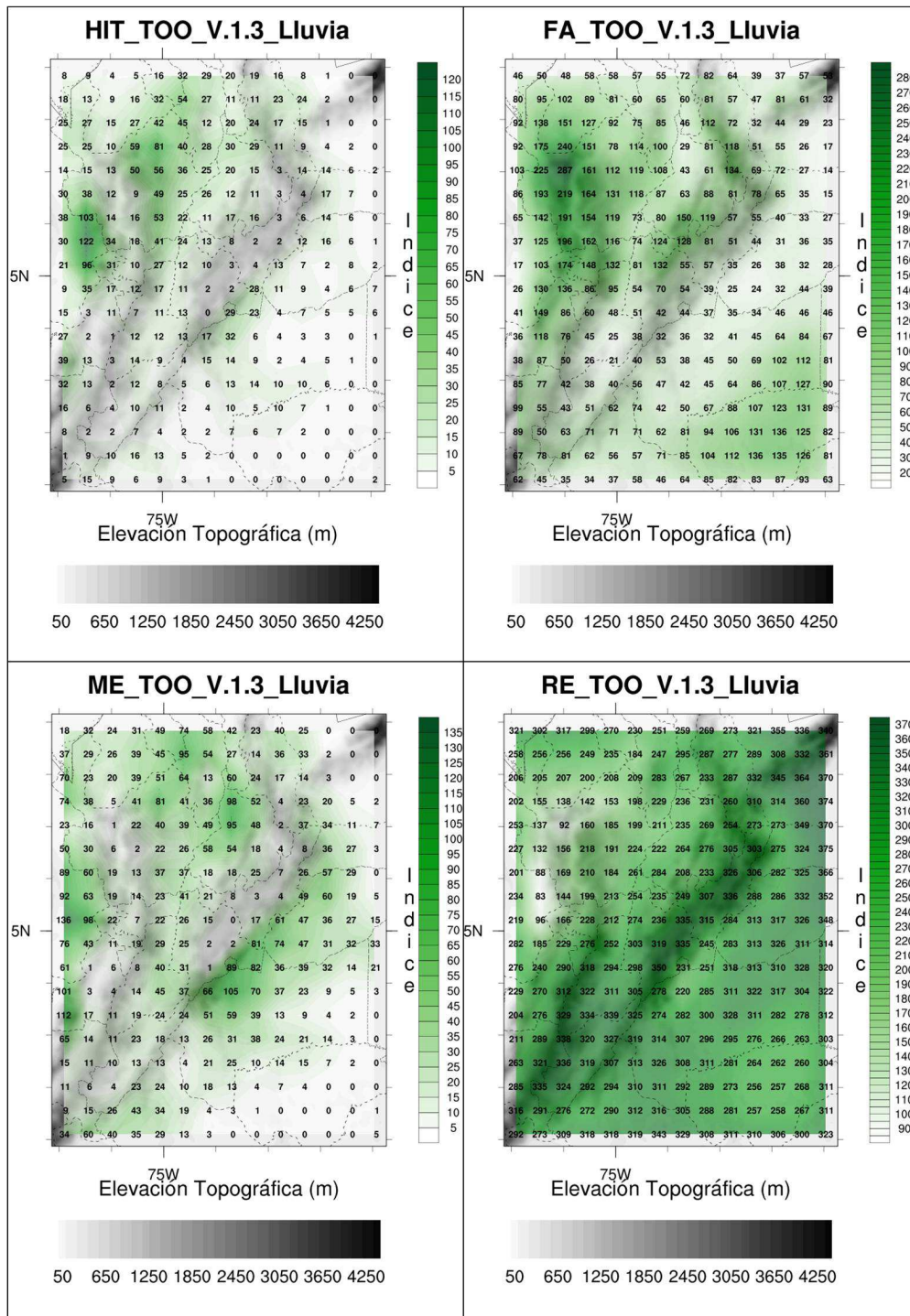


Figura 6-9: Resultados de las tablas de contingencia con técnicas orientadas a objetos

Las TOO muestran una concentración de los Hit sobre el costado oeste de la cordillera occidental, al igual que al norte de la cordillera occidental y central, sobre el valle interandino del Magdalena Medio, límites del departamento del Meta y Cundinamarca, piedemonte amazónico, este de la serranía de la Macarena y al oeste de Saravena. Las False Alarm, se encuentran distribuidas sobre las partes más altas de la cordillera occidental, central y oriental, al igual que sobre el departamento de Guaviare, debido a la falta de datos en esa región. Los Miss Event están concentrados sobre el piedemonte llanero y amazónico, entre el valle interandino del Magdalena, entre la cordillera central y la serranía de San Lucas, al norte de la cordillera central y occidental y al oeste de la cordillera occidental. Los CorectRejection están concentrados sobre casi toda la topografía del dominio, las regiones con menos CorectRejection están ubicadas al norte y oeste de la cordillera occidental y norte de la central, al igual que en el piedemonte llanero y amazónico, y el valle interandino del Magdalena.

▪ **BIAS**

EL BIAS calculado con las técnicas orientadas a objetos, arroja una de las gráficas más importantes en los resultados de este trabajo, los máximos están asociados con la topografía y con puntos en los que hay pocas observaciones, las regiones en las que se presentan estos máximos se encuentran bien marcadas y son fáciles de distinguir en la Figura 6-10. En ella se observan las más altas sobrestimaciones sobre el oeste y noreste departamento de Guaviare, en los límites entre el Guaviare y norte del Vichada, límites entre el departamento de Santander y Norte de Santander, sur de Cundinamarca, este del departamento de Valle del Cauca, en límites con Huila y sur de Huila, y el norte de la cordillera occidental y central.

Comparando la Figuras 6-5 y 6-10 se nota una significativa mejoría del BIAS en la TOO en comparación con el BIAS de la TET para variable discreta. Entre las regiones para las cuales esta ostensible mejoría es evidente se puede mencionar: Al oeste de la cordillera occidental; al oeste de Antioquia, en la región que limita el Chocó con el suroeste antioqueño y con el norte del Valle del Cauca, en el valle del Cauca, a lo largo del lado este de la cordillera central y en Córdoba, en el valle interandino del Magdalena exceptuando la vertiente oeste de la cordillera oriental en Cundinamarca oeste de Boyacá, sur y noroeste de Santander. También se nota en el Norte de Santander

exceptuando el límite al sur con Santander y a lo largo del lado este de la cordillera oriental excluyendo el piedemonte amazónico en los límites entre Meta y Caquetá.

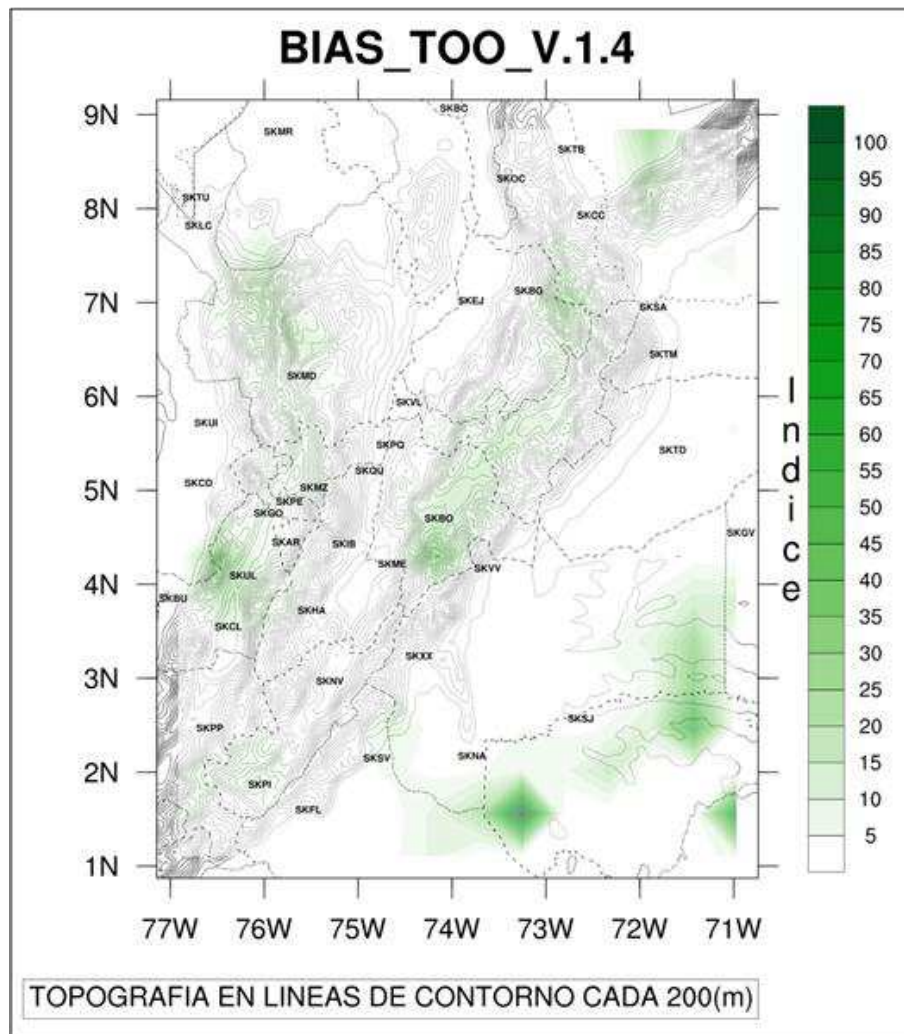


Figura 6-10: Resultado de BIAS, TOO con variable discreta.

▪ POD y FAR

Como se muestra en la Figura 6-11 (panel izquierdo), el POD muestra regiones con mayores aciertos sobre la región norte de la cordillera occidental y central, oeste de la occidental, límites entre los departamentos de Cundinamarca y Boyacá, límites entre los departamentos de Santander y Norte de Santander, y sobre el norte de Guaviare y sur

del Meta. Por otro lado, en la Figura 6-11 (panel derecho), el FAR muestra las regiones en las que la que la fracción de las predicciones no ocurre, concentrándose los resultados de este índice principalmente sobre las partes altas de las cordilleras.

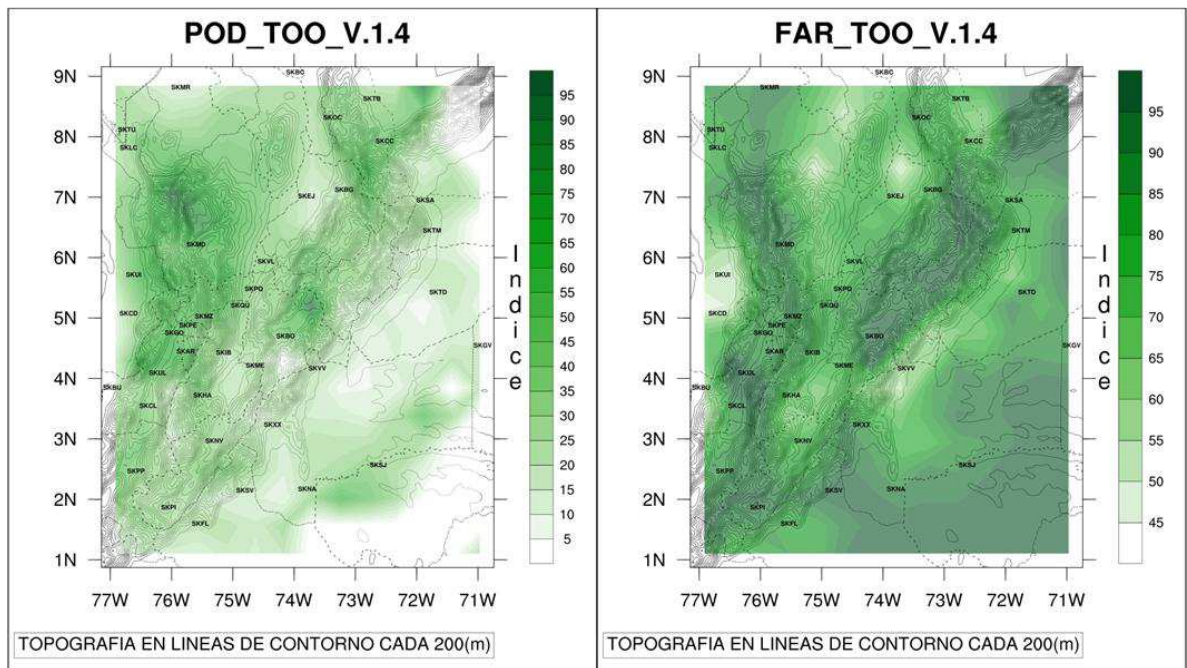


Figura 6-11: Resultados POD (izq.) y FAR (der.), TOO con variable discreta.

Es de notar que en los puntos donde se presenta alto POD, por lo regular, también se presenta alto FAR. Esto limita la capacidad del índice POD para calificar, por sí solo, la calidad del pronóstico

▪ **CSI**

El índice CSI, como se mencionó anteriormente, castiga las False Alarm y los Miss Event como se puede ver en la fórmula de la ecuación 3.13. Por esta razón el CSI se considera un índice más robusto que el POD. Los resultados del CSI se muestran en la Figura 6-12. En ella se observa que las regiones donde mayormente corresponden los eventos pronosticados con los observados se encuentra concentradas sobre el piedemonte llanero, norte y este de la serranía de la Macarena, este de Caquetá, entre el valle interandino del Magdalena, noreste de la cordillera central y serranía de San Lucas, oeste de la cordillera occidental.

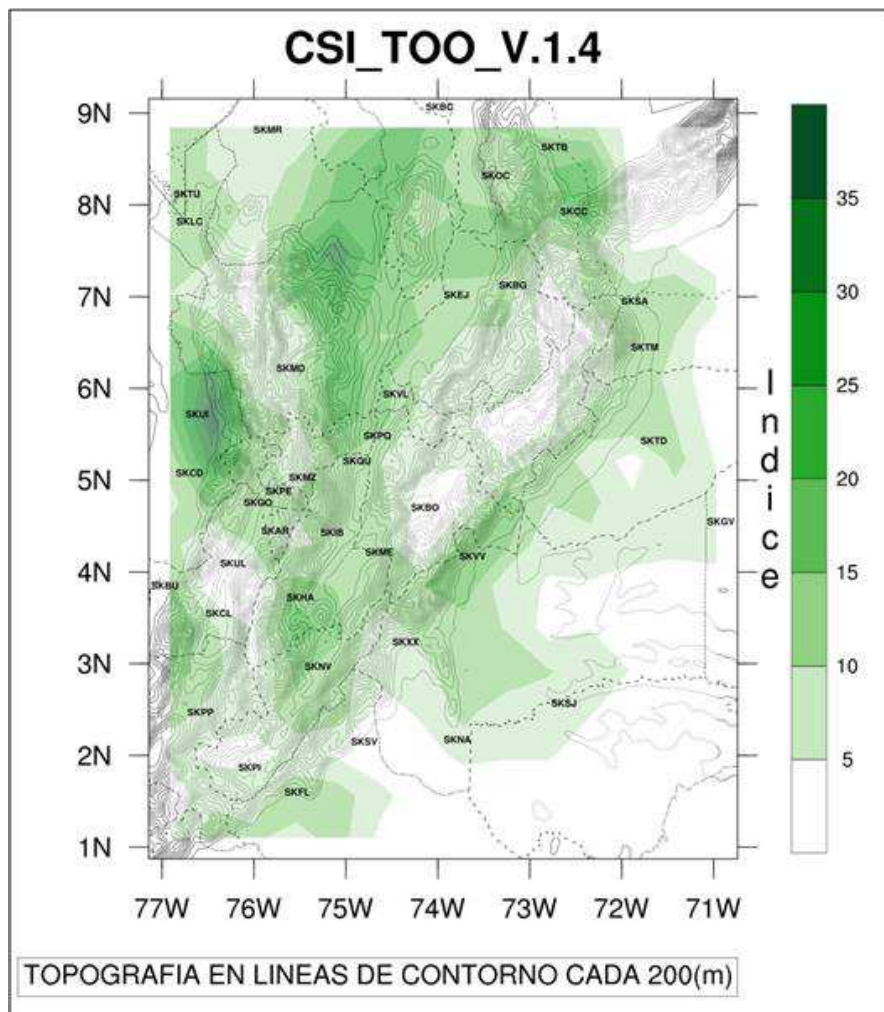


Figura 6-12: Resultados CSI, TOO con variable discreta.

Comprando las figuras 6.10 y 6.12, se puede observar que los puntos con alto BIAS y alto CSI son excluyentes, lo que es entendible físicamente y representa una fortaleza de este último índice.

▪ ETS

El ETS al igual que el CSI penaliza las False Alarm y los Miss Event, pero a diferencia de éste, también penaliza el número de aciertos aleatorios de los pronósticos. En la Figura 6-13 se muestra los resultados del ETS, en ella se indica que tan bien se hizo el pronóstico con respecto a los eventos observados. La regiones en las que se realizó

mejor el pronóstico están concentradas sobre el norte y noreste de la cordillera central, norte del departamento de Huila y sur de Tolima, norte de la cordillera oriental, piedemonte llanero en límites entre el departamento de Cundinamarca y Meta, norte de Arauca y centro de Casanare.

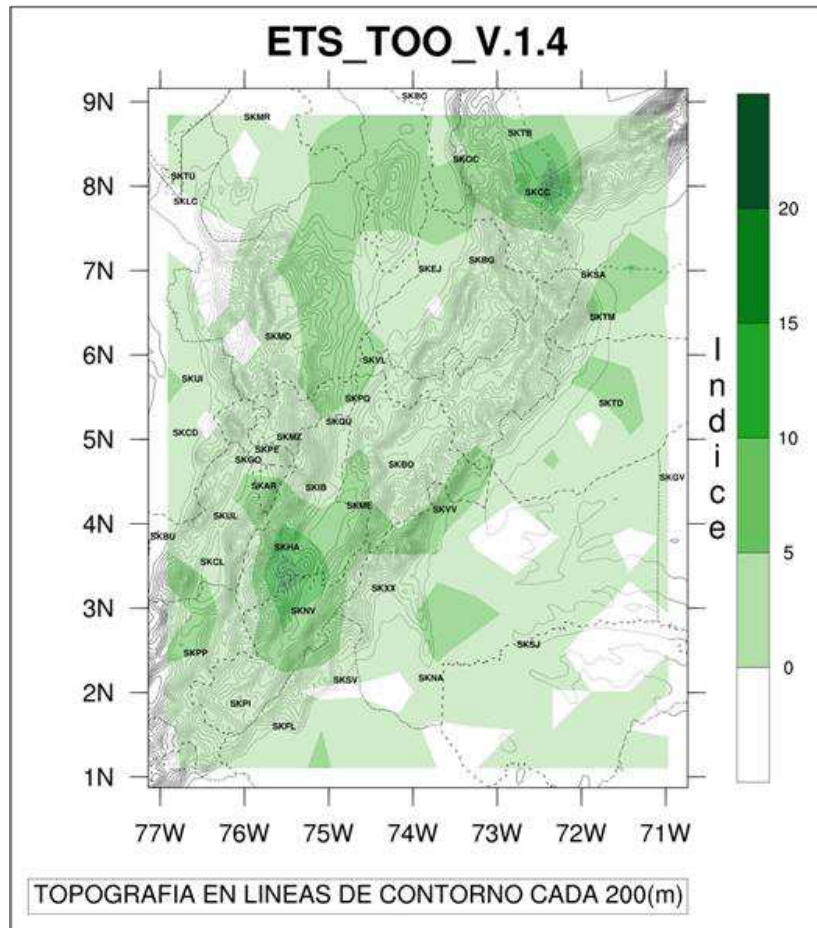


Figura 6-13: Resultados ETS, TOO con variable discreta.

- **Ganancia en exactitud de la técnica orientada a objetos contra la técnica de estadística tradicional**

Con el propósito de medir la ganancia en exactitud al implementar la Técnica de Estadística Tradicional contra la Técnica Orientada a Objetos, se calculó el promedio de los resultados arrojados por cada uno de los índices calculados, obteniendo una

ganancia muy significativa al realizar la validación del pronóstico implementando las TOO, tal como se relaciona en la Tabla 6-1.

Tabla 6-1: Ganancia en exactitud de la TET contra la TOO

	TET	TOO	GANANCIA EN EXACTITUD (%)
HIT	0,34	13,48	97,46
FALSE ALARM	13,37	76,69	82,57
MISS	6,25	26,81	76,68
CORRECT REJECTION	373,04	276,02	26,01
BIAS	2,84	5,41	47,42
POD	0,04	0,30	85,98
FAR	0,93	0,86	8,49
CSI	0,01	0,10	86,26
ETS	0,01	0,03	80,17

7 Conclusiones

Los objetivos de este trabajo de investigación están enmarcados en la eficiente implementación y análisis de la Técnica Orientada a Objetos y su comparación con la Estadística Tradicional en puntos de grilla, determinando la ganancia en exactitud del uno contra el otro e identificando las regiones en las que el modelo WRF sobrestima y subestima la precipitación. Sobre estos aspectos, el análisis del resultado de esta investigación nos permite realizar las siguientes conclusiones:

1. Se implementó la técnica de estadística tradicional en puntos de grilla, realizando dos análisis, uno mediante variable continua, en la que se calcularon los índices BIAS, MAE, MSE, RMSE, y el otro mediante eventos categóricos o tablas de contingencia de 2x2, en la que se calcularon los índices BIAS, POD, FAR, CSI y ETS. Las dos técnicas permitieron corroborar los resultados obtenidos.

Tras el análisis de los resultados, se logró concluir el efecto que tiene la carencia de datos en la validación y los pronósticos. Las regiones en las que el modelo sobrestima las precipitaciones generalmente están asociadas con la mayor complejidad en la topografía, pero ésta no es la única razón para que el modelo sobrestime. Mediante el análisis del BIAS y la ubicación espacial de las estaciones meteorológicas se pudo observar que en las regiones en las que existía carencia de estaciones, los resultados arrojaban valores muy altos de sobrestimación de la precipitación independiente de si estaba o no ubicado sobre terreno complejo.

2. Se implementó la técnica orientada a objetos con tablas de contingencia de 2x2, los índices calculados fueron BIAS, POD, FAR, CSI y ETS.

Tras el análisis, se logró determinar que la mejor manera de ver los resultados de los pronósticos es en el índice CSI, que nos indica qué tan bien corresponden los eventos pronosticados con los observados, las regiones de mayor porcentaje son el

pedemonte llanero, oeste de Caquetá, valle interandino del Magdalena, norte y noreste de la cordillera Central, este y oeste de la serranía de San Lucas, este de Choco, oeste del Valle del Cauca y norte del Cauca.

3. Como se pudo demostrar, la Técnica de Estadística Tradicional en variable categórica castiga los pequeños desplazamientos de los sistemas, por el contrario la Técnica Orientada a Objetos no penaliza este desplazamiento y arroja mayor exactitud en los resultados, además la técnica orientada a objetos muestra áreas representativas que facilitan la ubicación de regiones con mejores aciertos en las predicciones.
4. El modelo WRF implementado en la FAC, con parametrización de la microfísica de Groupel, sobrestima los valores de la precipitación en regiones con topografía compleja, presentando mayor sobrestimación sobre la cordillera oriental al oeste de Villavicencio, al este y oeste del departamento del Valle del Cauca, límites entre los departamento de Norte de Santander y Santander, y norte, este y oeste del departamento de Guaviare. Además, presenta subestimación en el centro del departamento del Tolima, este de la serranía de San Lucas, centro y sur de Choco, oeste del Valle del Cauca, y norte y nororiente de la serranía de la Macarena.

Las regiones en las que el modelo WRF de la FAC genera pronósticos confiables son el piedemonte llanero, departamentos de Casanare, oeste de Arauca, oeste de Meta, límite entre Huila y Tolima, valle interandino del Magdalena, sur del Cesar, norte y noreste de Norte de Santander, este de Choco y oeste de la serranía de San Lucas.

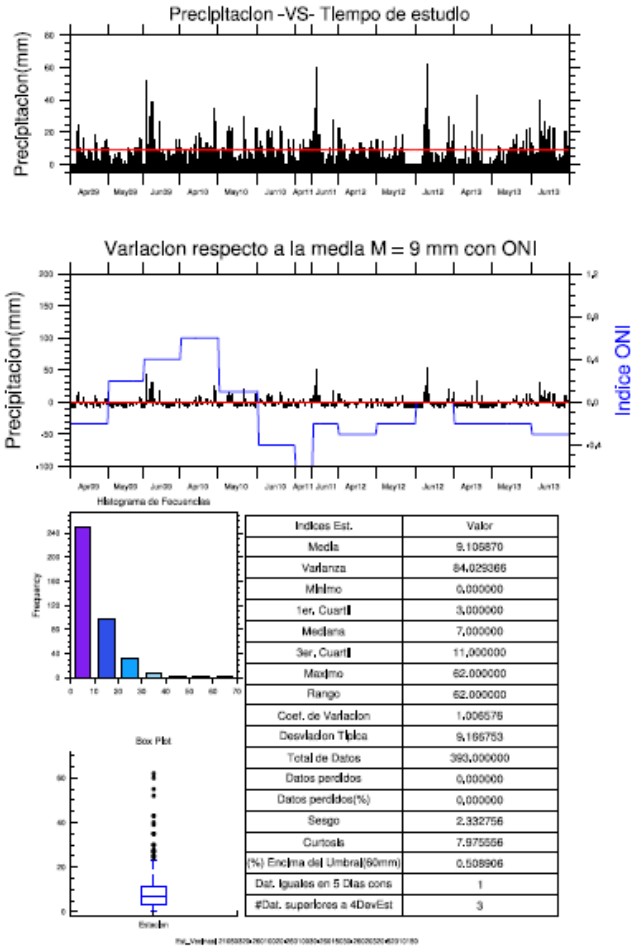
En resumen, los dos resultados más novedosos de esta investigación son: 1). La técnica de orientada a objetos mejora ostensiblemente la manera de verificar los pronósticos de precipitación extrema realizados con modelos de alta resolución como el WRF y 2). La calidad de los pronósticos en Colombia está altamente asociada a 2 factores: La topografía y la carencia de observaciones en algunas regiones.

Este trabajo de investigación, permitió identificar las regiones en las que el modelo WRF de la FAC realiza mejores predicciones meteorológicas, lo cual permite a los pronosticadores del tiempo en la FAC emitir pronósticos con un mayor grado de

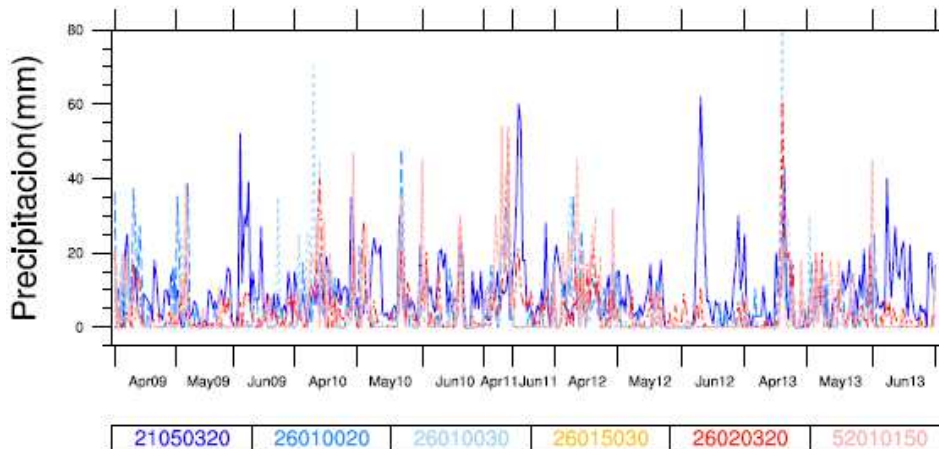
probabilidad de acierto, lo que a su vez mejora las predicciones y el adecuado suministro de información meteorológica para el desarrollo de operaciones militares y aéreas en el país.

A. Anexo: Estación con datos de precipitación iguales en 5 días consecutivos

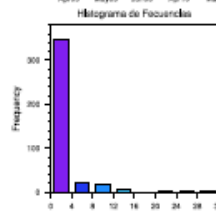
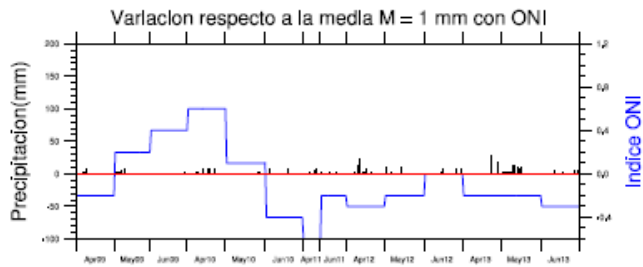
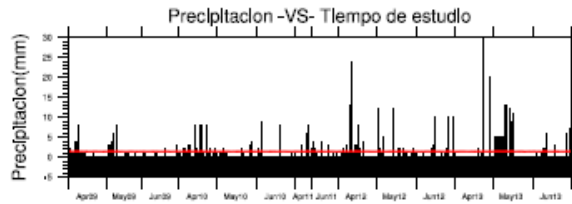
Id_Est: 21050320 Lat: 2.351800 Long: -76.496400



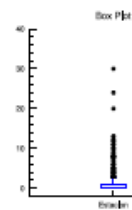
Precipitación -VS- Tiempo de estudio



Id_Est: 21140010 Lat: 3.291600 Long: -74.904000

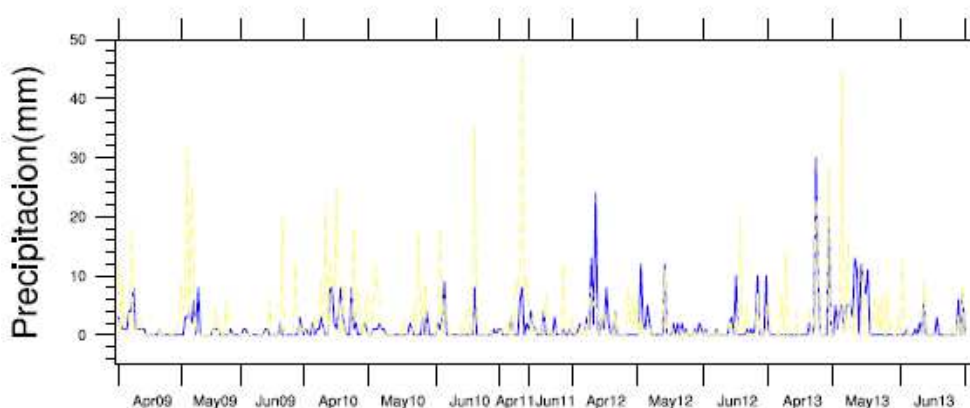


Índice Est.	Valor
Medio	1,400000
Varianza	10,331785
Mínimo	0,000000
1er. Cuartil	0,000000
Mediana	0,000000
3er. Cuartil	1,000000
Máximo	30,000000
Rango	30,000000
Coeff. de Variación	2,318049
Desviación Típica	3,245268
Total de Datos	390,000000
Datos perdidos	1,000000
Datos perdidos(%)	0,254453
Sesgo	4,231244
Curtosis	24,500012
(%) Encima del Umbral(50mm)	0
Dat. iguales en 5 Días cons	1
#Dat. superiores a 4DevEst	3



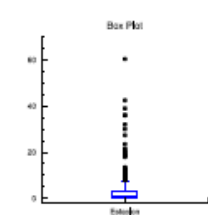
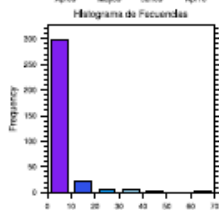
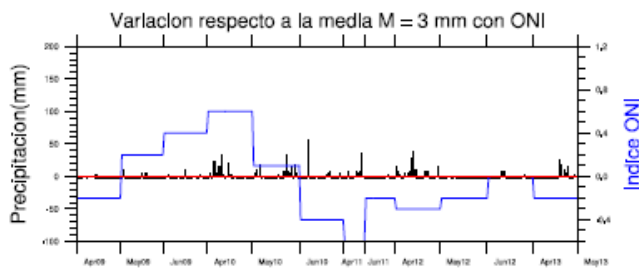
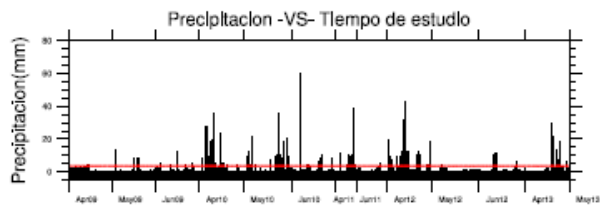
File: D:\Estad\21140010\140229\21140010

Precipitación -VS- Tiempo de estudio



21140010	21140030	21145060
----------	----------	----------

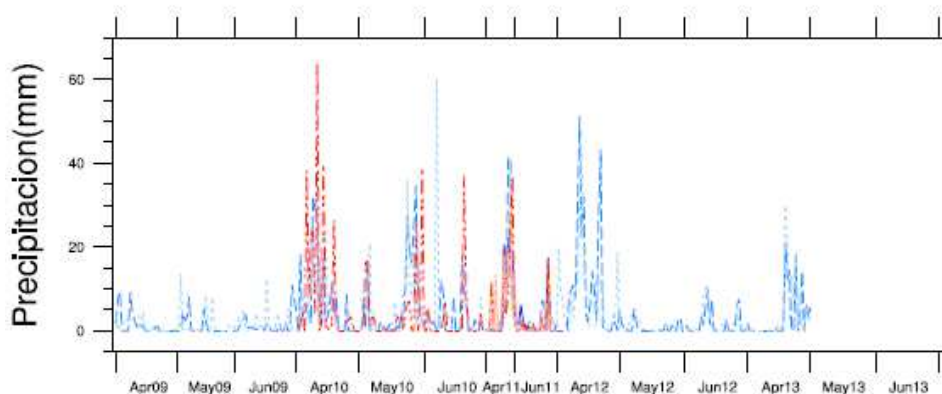
Id_Est: 21201190 Lat: 5.214500 Long: -73.907300



Índice Est.	Valor
Media	3,528067
Varianza	56,503057
Modo	0,000000
1er. Cuartil	0,200000
Mediana	0,800000
3er. Cuartil	3,200000
Máximo	60,400000
Range	60,400000
Coef. de Varianción	2,014273
Desviación Típica	7,195550
Total de Datos	393,000000
Datos perdidos	62,000000
Datos perdidos(%)	15,778081
Sesgo	3,888007
Curtois	15,423735
(%) Encima del Umbral(50mm)	0,302115
Dat. iguales en 5 días consec.	3
#Dat. superiores a 4DevEst	5

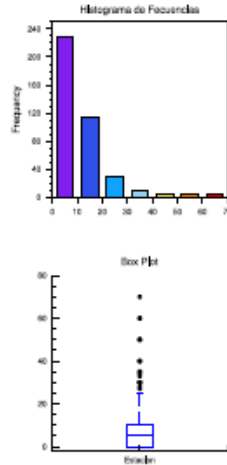
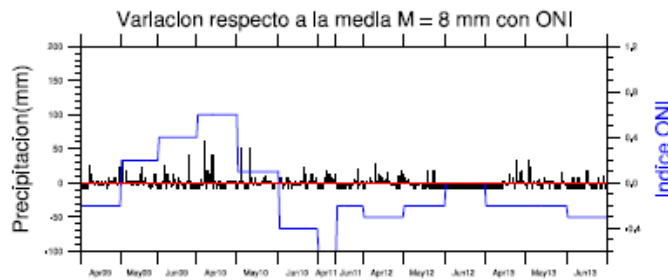
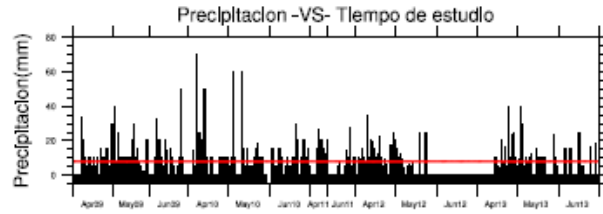
file_name=21200950-21201180-21201190-21201380-21205400-24010270

Precipitación -VS- Tiempo de estudio



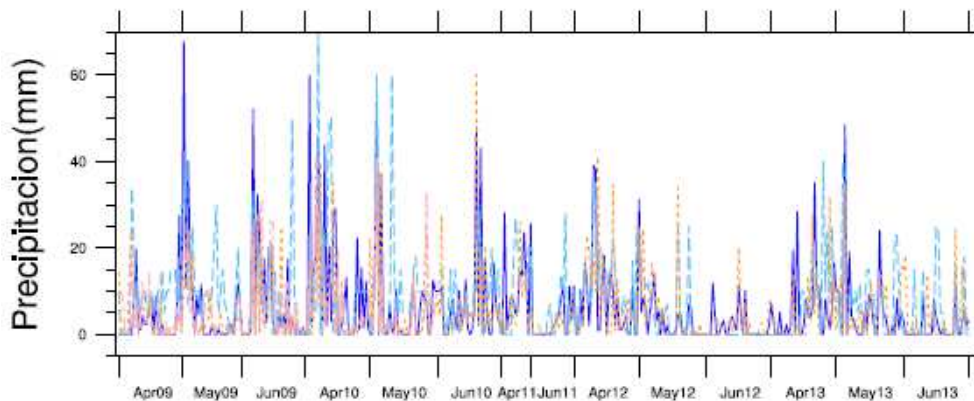
21200950 21201180 21201190 21201380 21205400 24010270

Id_Est: 26150150 Lat: 4.951200 Long: -75.490600



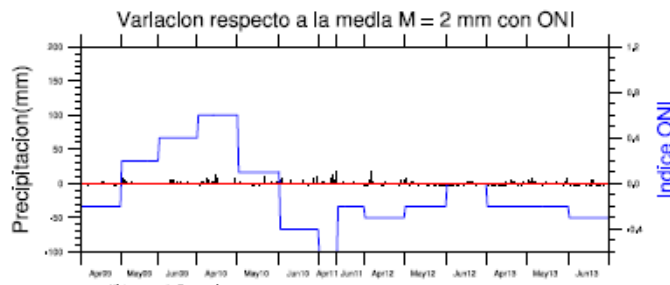
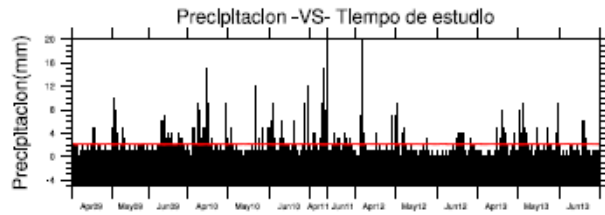
Indice Est.	Valor
Moda	8,030769
Varianza	108,184141
Mínimo	0,000000
ter. Cuartil	0,000000
Mediana	5,000000
Sei. Cuartil	10,000000
Máximo	70,000000
Rango	70,000000
Coef. de Variación	1,295164
Desviación Típica	10,401161
Total de Datos	393,000000
Datos perdidos	3,000000
Datos perdidos(%)	0,783359
Sesgo	2,257180
Curtosis	7,320676
(%) Encima del Umbral(50mm)	0,759231
Dat. iguales en 5 Dias cons	4
#Dat. superiores a 4DevEst	3

Precipitacion -VS- Tiempo de estudio

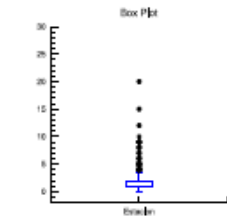
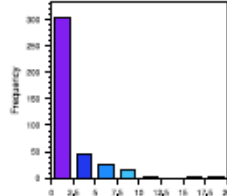


26130170	26150150	26155110	26155170
----------	----------	----------	----------

Id_Est: 53110030 Lat: 3.633300 Long: -76.715200

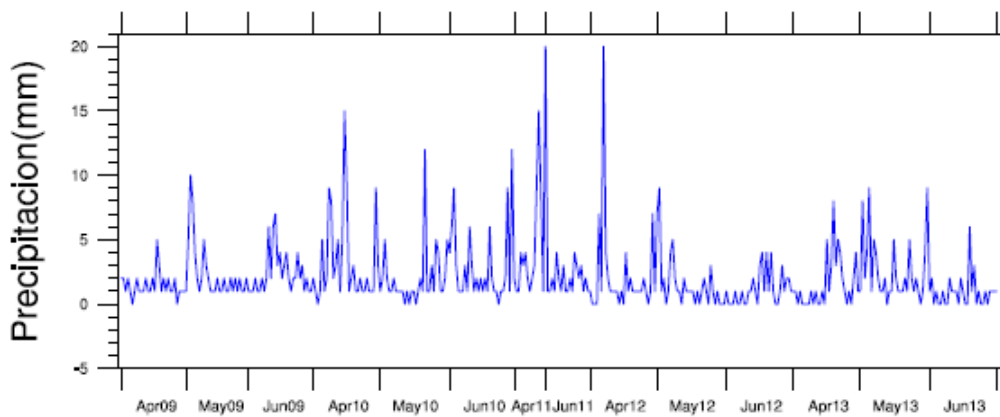


Histograma de Frecuencias



Índice Est.	Valor
Moda	2,129771
Varianza	6,893831
Mínimo	0,000000
1er. Cuartil	1,000000
Mediana	1,000000
3er. Cuartil	2,000000
Máximo	20,000000
Rango	20,000000
Coef. de Variación	1,232814
Desviación Típica	2,626111
Total de Datos	393,000000
Datos perdidos	0,000000
Datos perdidos(%)	0,000000
Sesgo	3,168069
Curtosis	13,810885
(%) Encima del Umbral(50mm)	0
Dat. iguales en 5 Días cons	1
#Dat. superiores a 4DevEst	4

Precipitación -VS- Tiempo de estudio



53110030

B. Anexo:NameList de la configuración del modelo WRF de la FAC

"namelist.input domains"

```
&domains
time_step           = 240,
time_step_fract_num = 0,
time_step_fract_den = 1,
max_dom             = 3,
s_we                = 1,  1,  1,
e_we                = 137, 121, 142,
s_sn                = 1,  1,  1,
e_sn                = 104, 142, 184,
s_vert              = 1,  1,  1,
e_vert              = 27,  27,  27,
dx                  = 45000, 15000, 5000,
dy                  = 45000, 15000, 5000,
grid_id             = 1,  2,  3,
parent_id           = 1,  1,  2,
i_parent_start      = 1,  47,  33,
j_parent_start      = 1,  14,  45,
parent_grid_ratio    = 1,  3,  3,
parent_time_step_ratio = 1,  3,  3,
feedback            = 0,
smooth_option       = 0
```

"namelist.input physics"

```
&physics
mp_physics          = 0,  3,  6,
ra_lw_physics       = 1,  1,  1,
ra_sw_physics       = 2,  2,  2,
radt                = 10, 10,  5,
sf_sfclay_physics  = 1,  1,  1,
sf_surface_physics  = 3,  3,  3,
bl_pbl_physics      = 1,  1,  1,
bldt                = 0,  0,  0,
cu_physics          = 2,  1,  0,
cudt                = 2,  2,  2,
isfflx              = 1,
ifsnow              = 0,
icloud              = 1,
surface_input_source = 1,
num_soil_layers     = 6,
maxiens             = 1,
maxens              = 3,
maxens2             = 3,
maxens3             = 16,
ensdim              = 144,
```

"namelist.input dynamics"

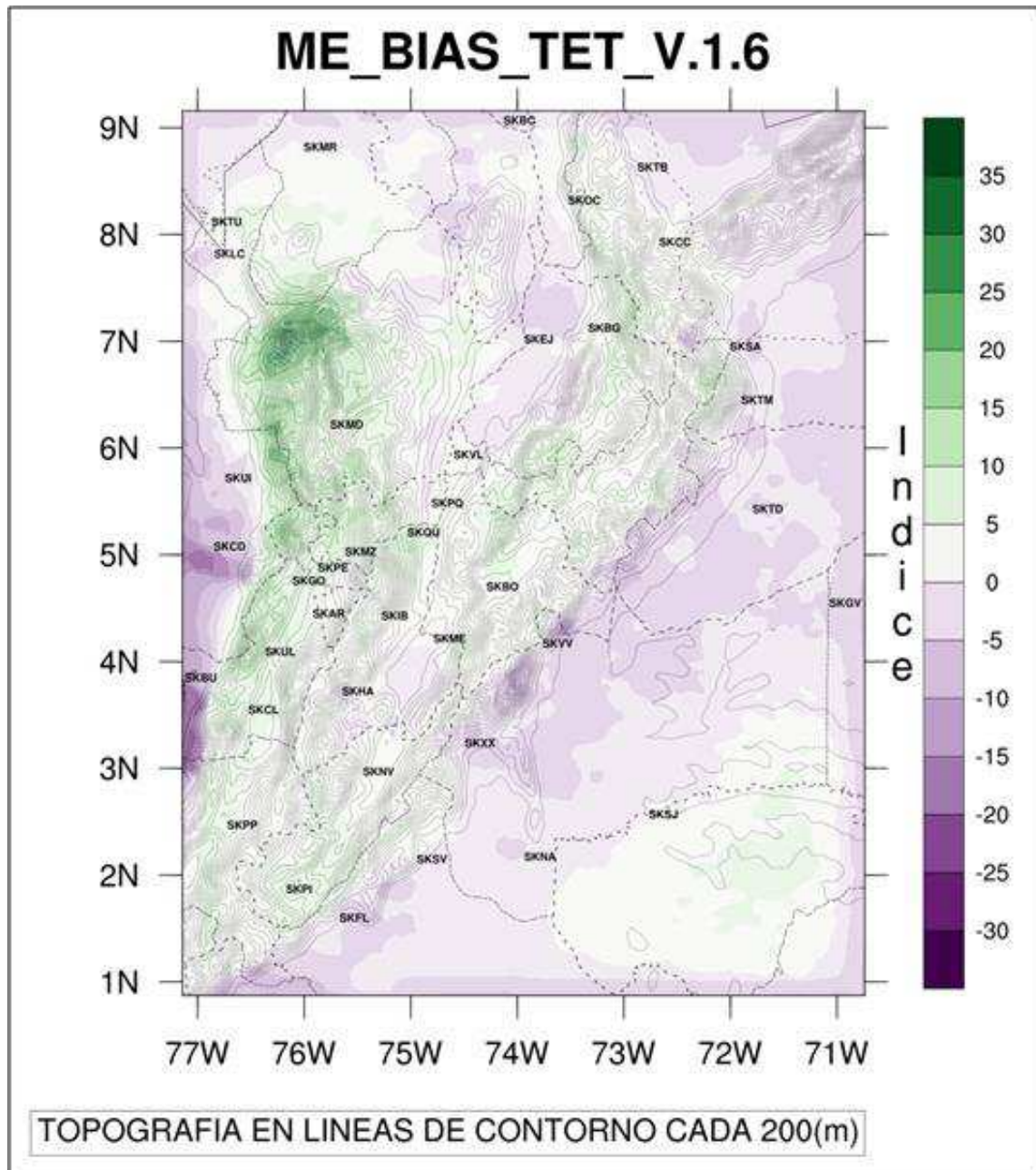
```
&dynamics
dyn_opt           = 2,
rk_ord            = 3,
w_damping         = 1,
diff_opt          = 1,
km_opt            = 4,
damp_opt          = 1,
base_temp         = 290.
zdamp             = 6000., 6000., 6000.,
dampcoef          = 0.01, 0.01, 0.01
khdif             = 0, 0, 0,
kvdif             = 0, 0, 0,
smdiv             = 0.1, 0.1, 0.1,
emdiv             = 0.01, 0.01, 0.01,
epssm            = 0.1, 0.1, 0.1
non_hydrostatic   = .true., .true., .true.,
h_mom_adv_order   = 3, 3, 3,
v_mom_adv_order   = 3, 3, 3,
time_step_sound   = 4, 4, 4,
h_sca_adv_order   = 3, 3, 3,
v_sca_adv_order   = 3, 3, 3,
```

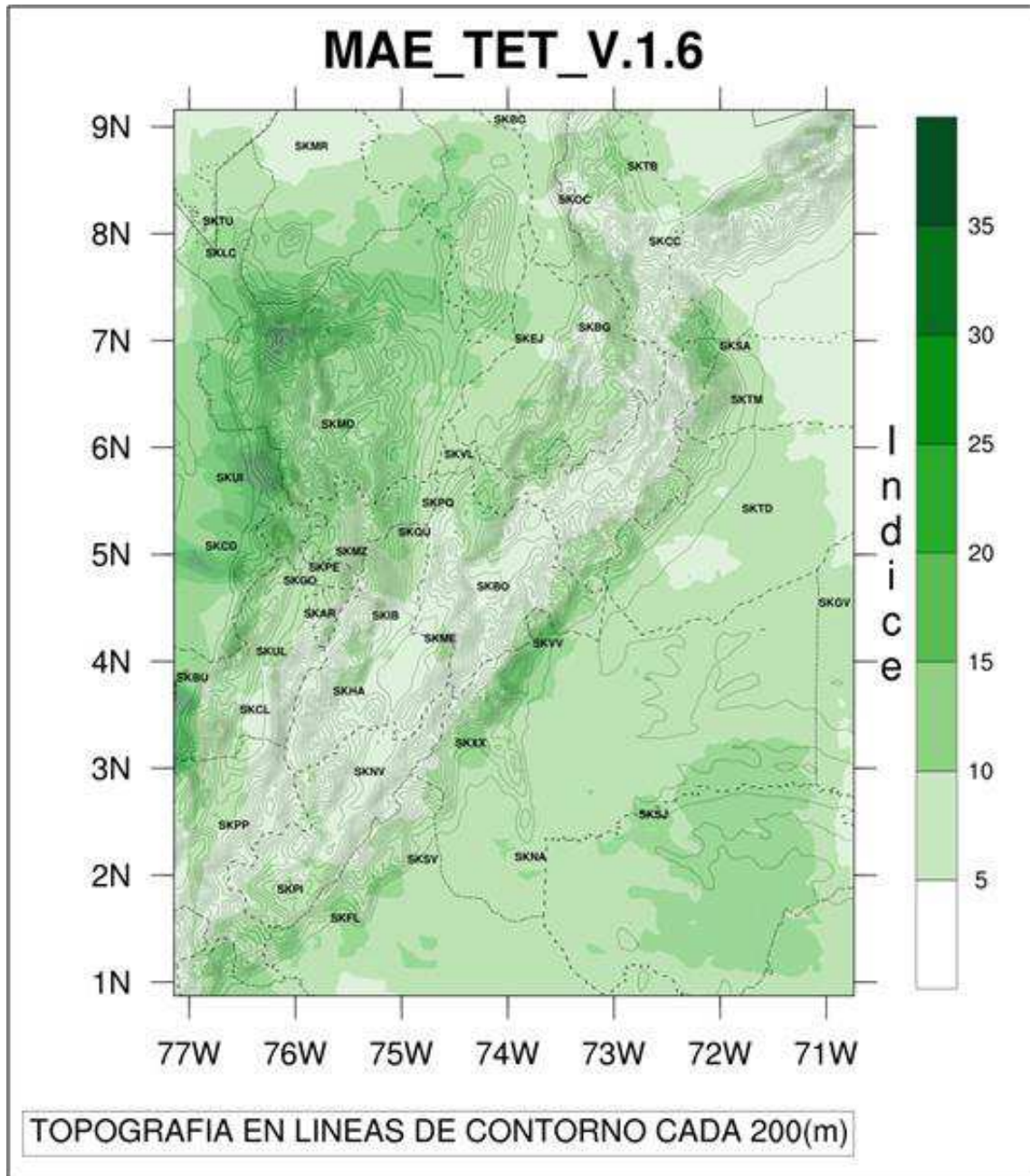
"*namelist.inputbdy_control*"

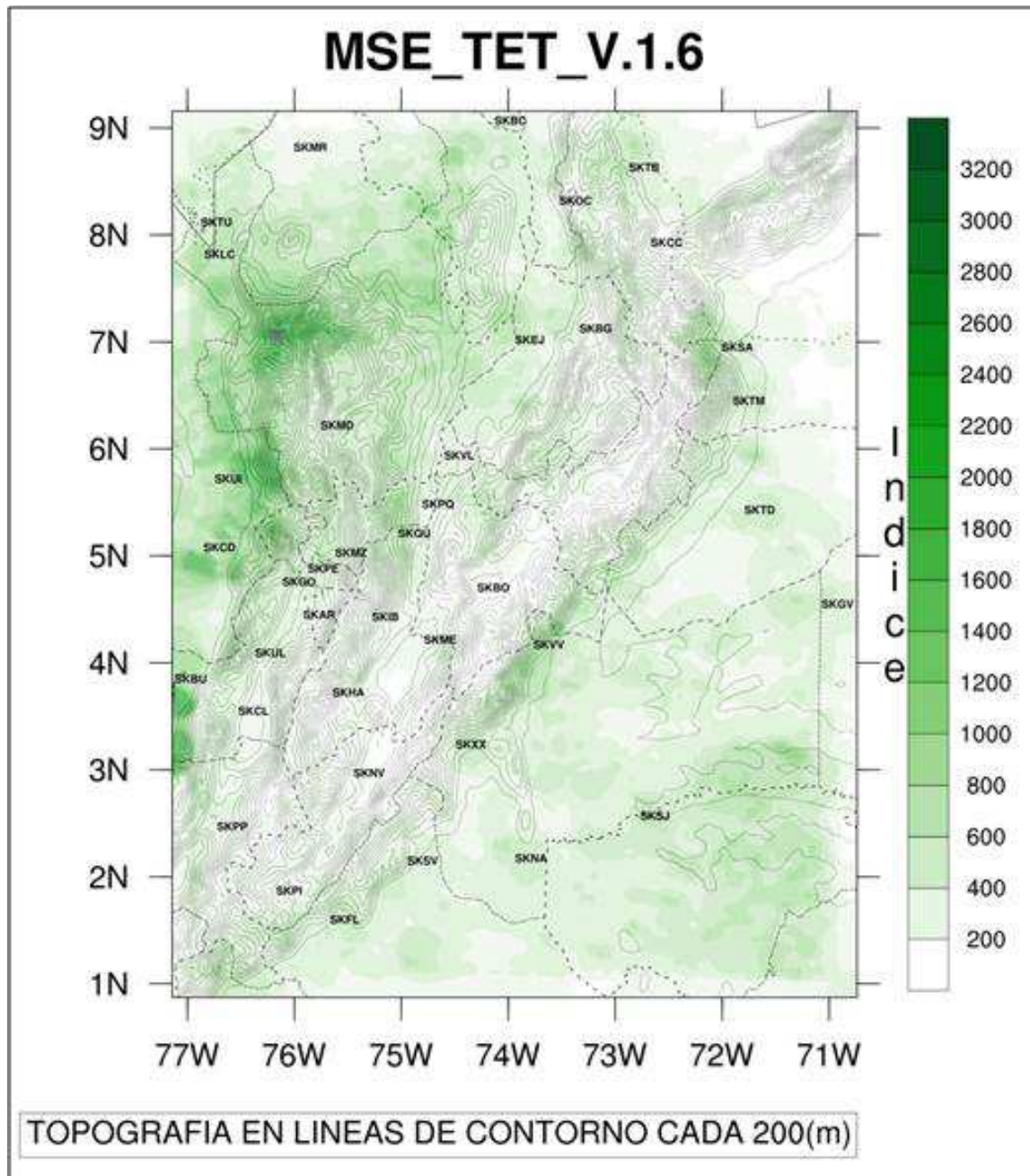
```
&bdy_control
spec_bdy_width = 5,
spec_zone      = 1,
relax_zone     = 4,
specified      = .true., .false.,.false.,
periodic_x     = .false.,.false.,.false.,
symmetric_xs   = .false.,.false.,.false.,
symmetric_xe   = .false.,.false.,.false.,
open_xs       = .false.,.false.,.false.,
open_xe       = .false.,.false.,.false.,
periodic_y     = .false.,.false.,.false.,
symmetric_ys   = .false.,.false.,.false.,
symmetric_ye   = .false.,.false.,.false.,
open_ys       = .false.,.false.,.false.,
open_ye       = .false.,.false.,.false.,
nested        = .false., .true., .true.,
```

C. Anexo: Resultados de las técnicas de validación implementadas

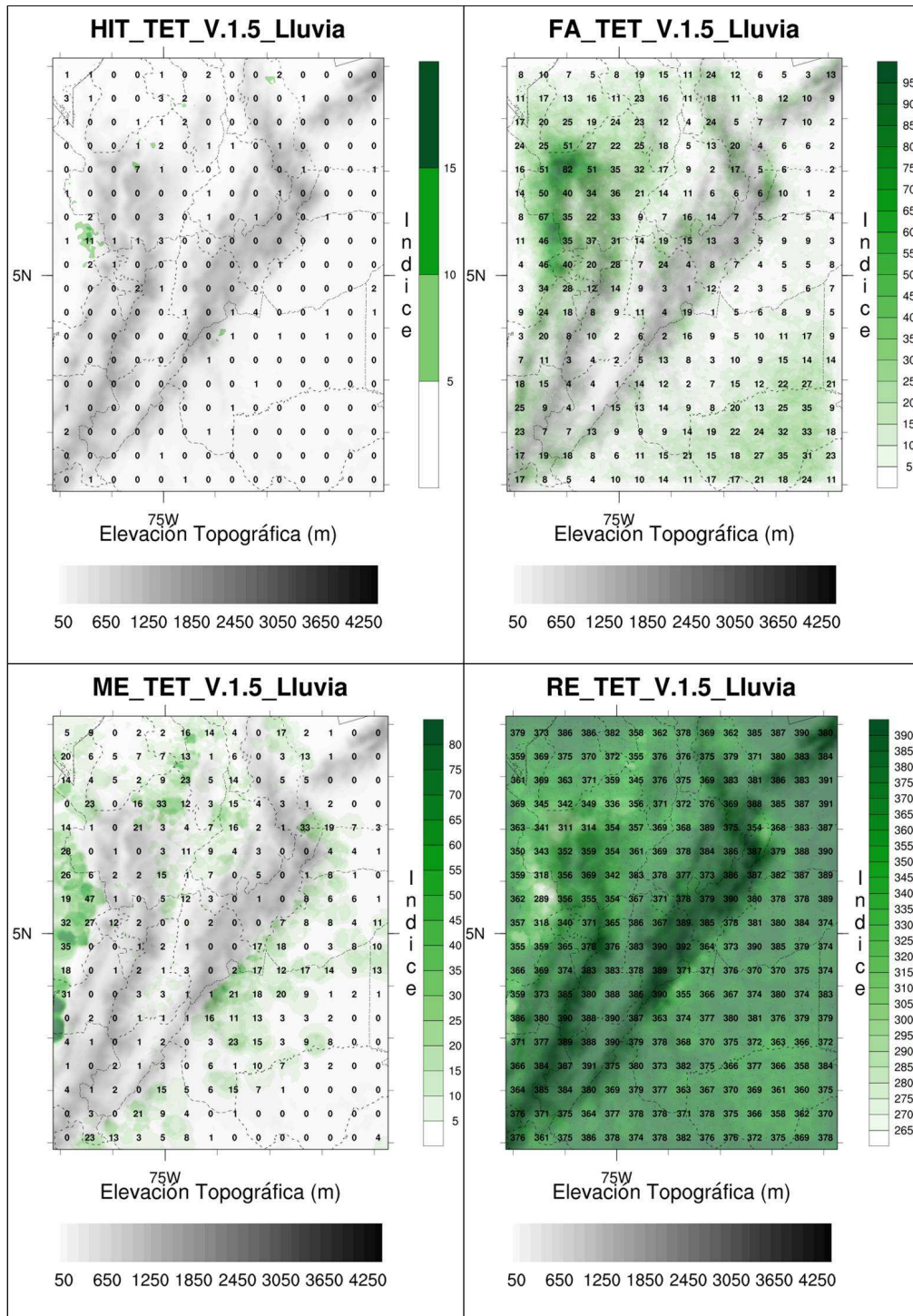
Técnicas de estadística tradicional con variable continua

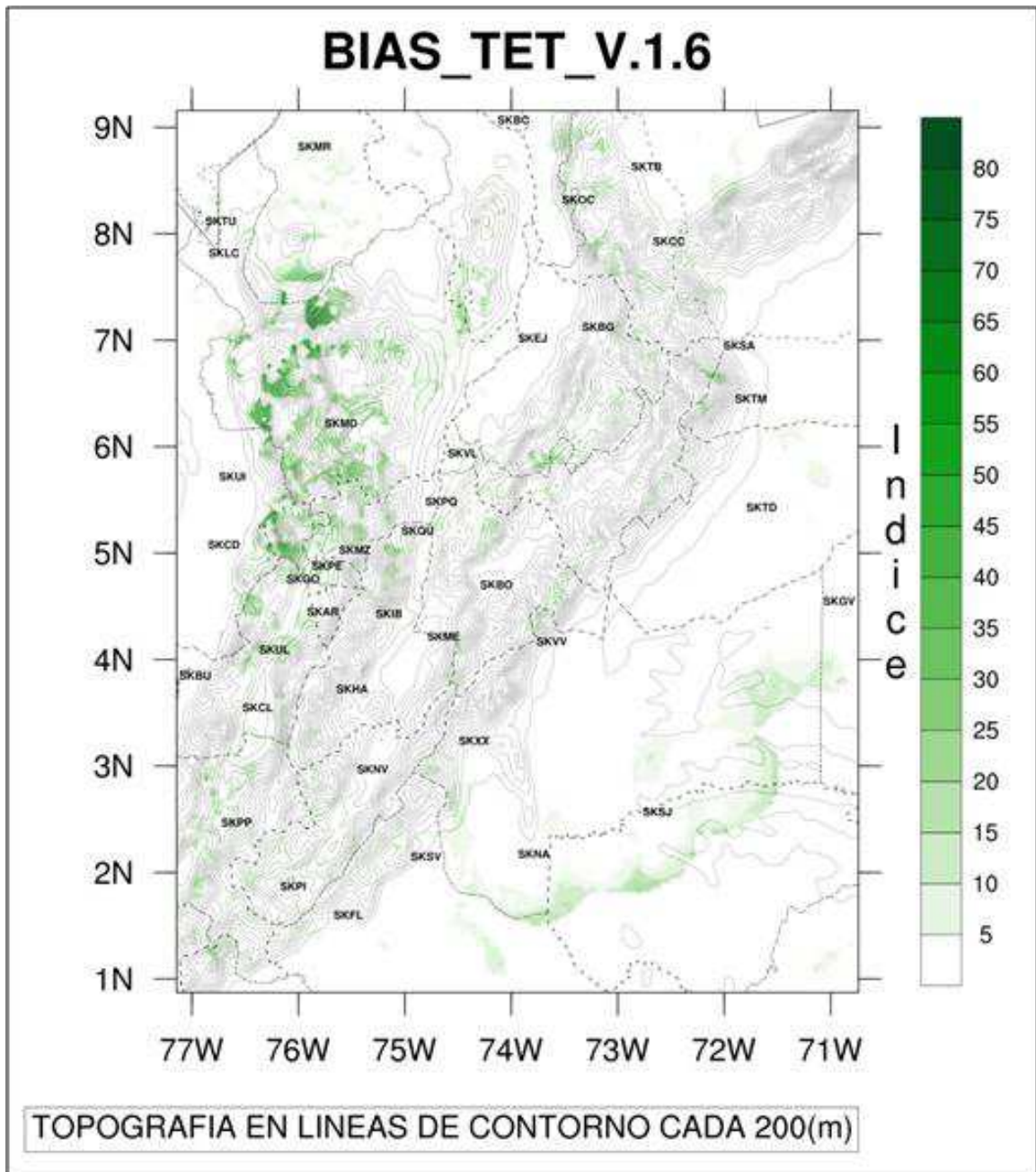


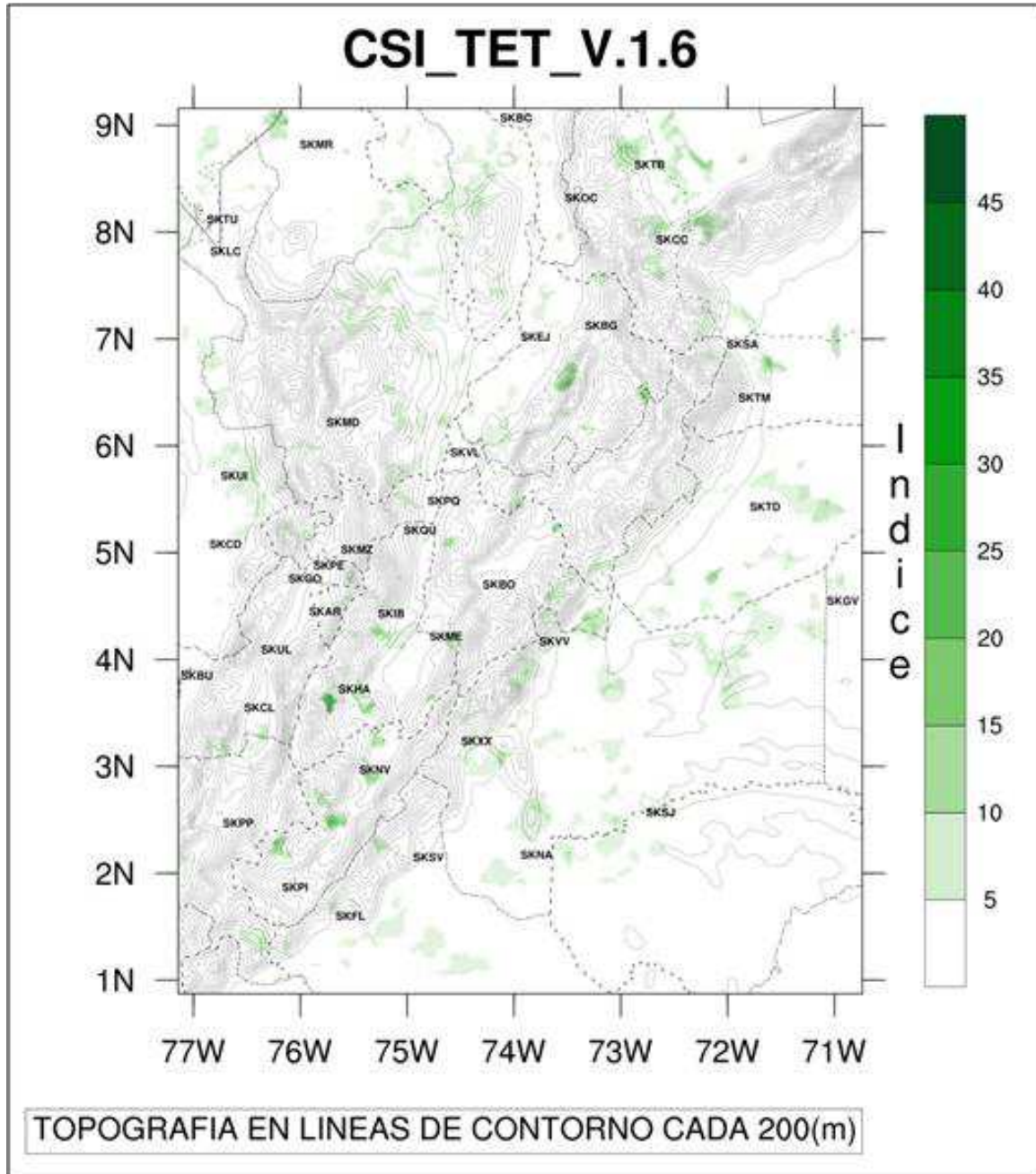


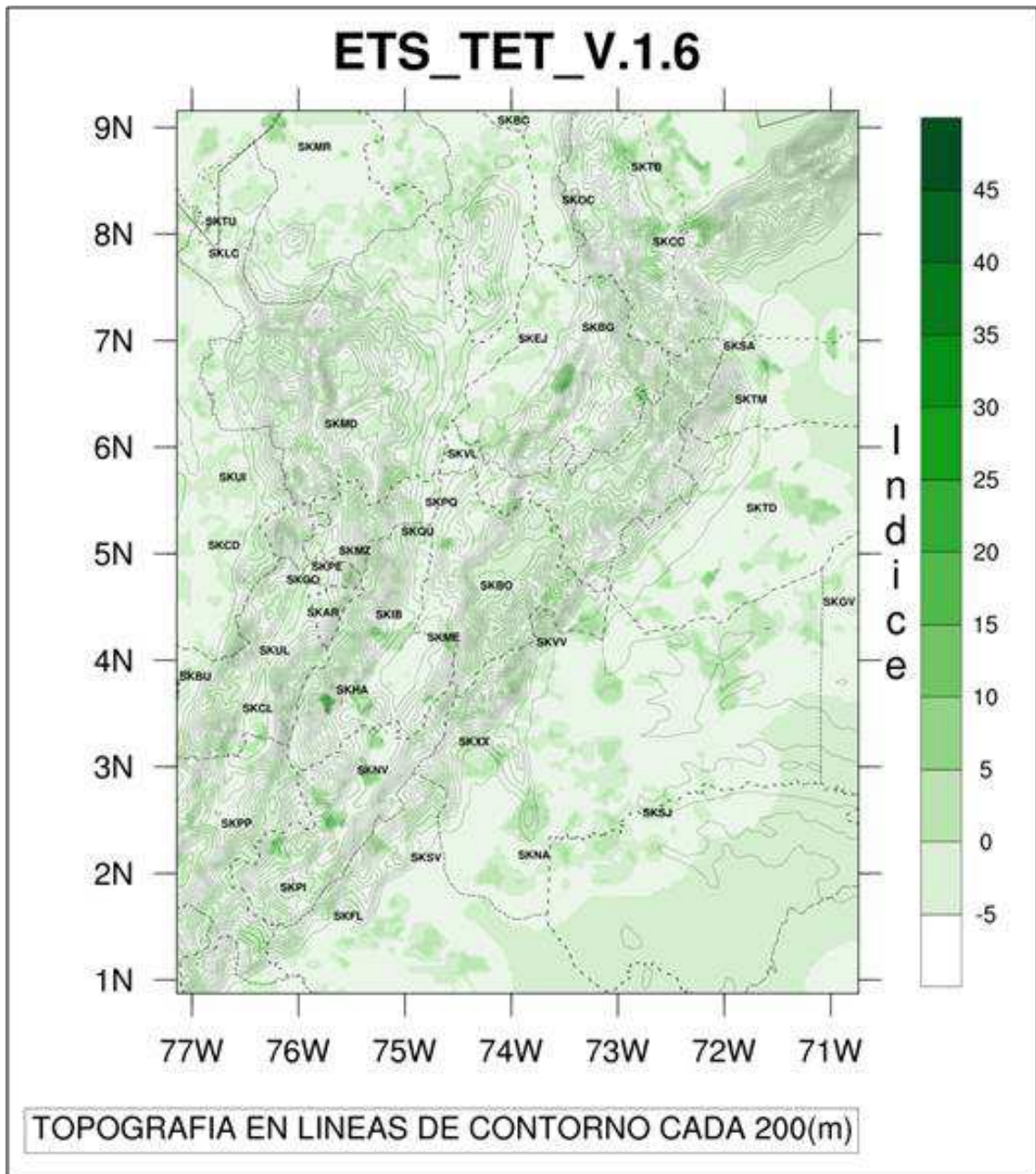


Técnicas de Estadística Tradicional con Variable Discreta

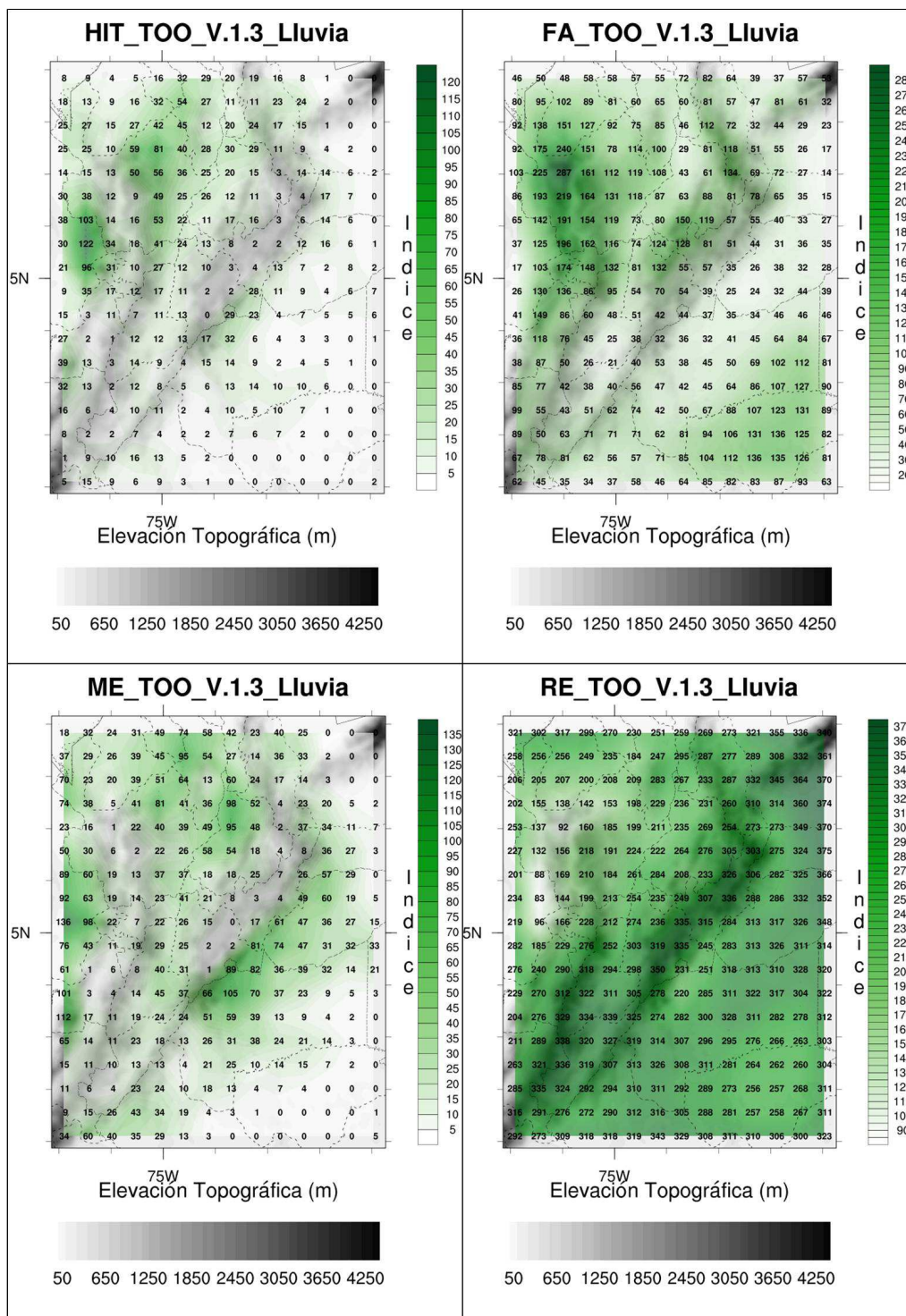


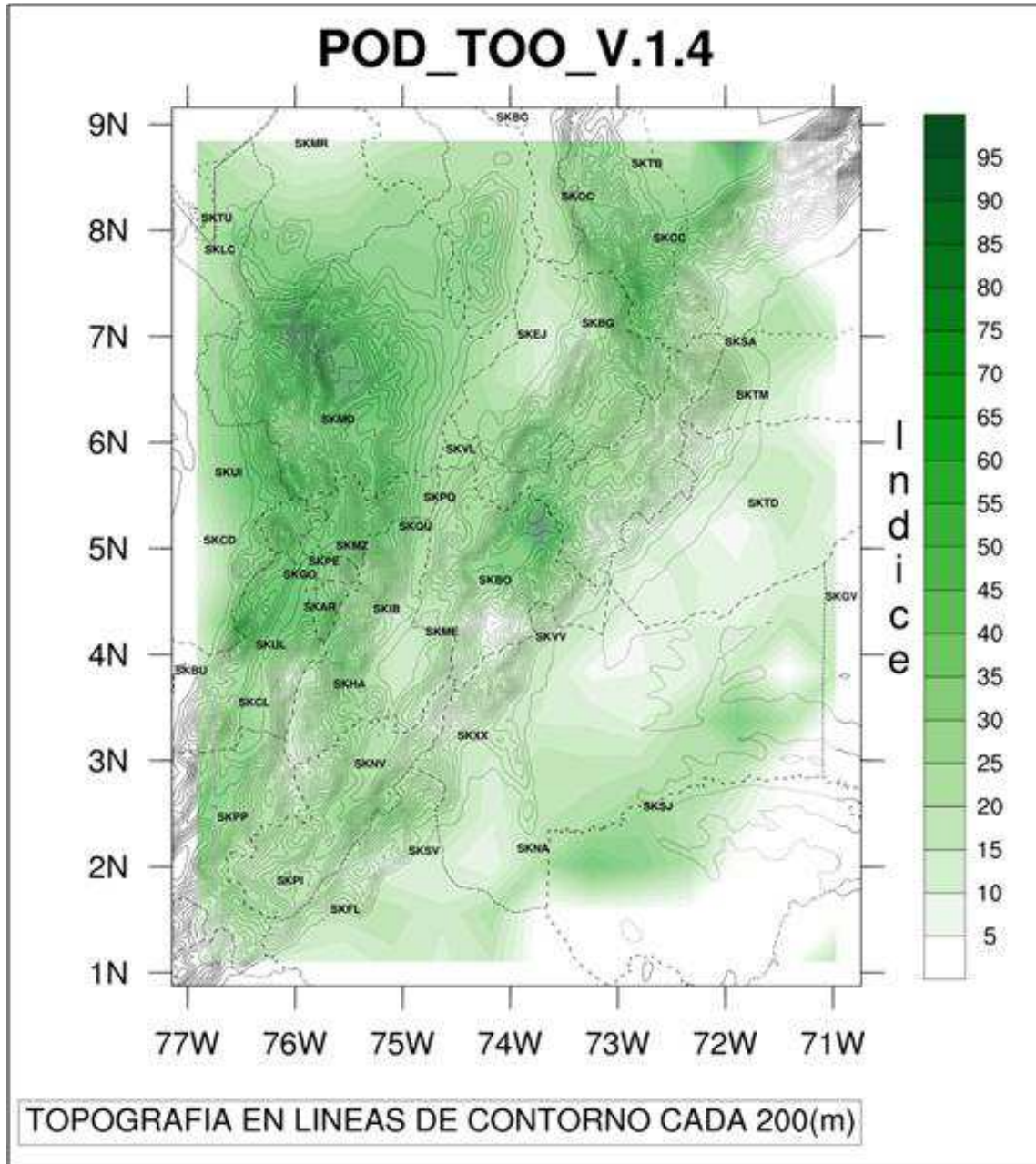


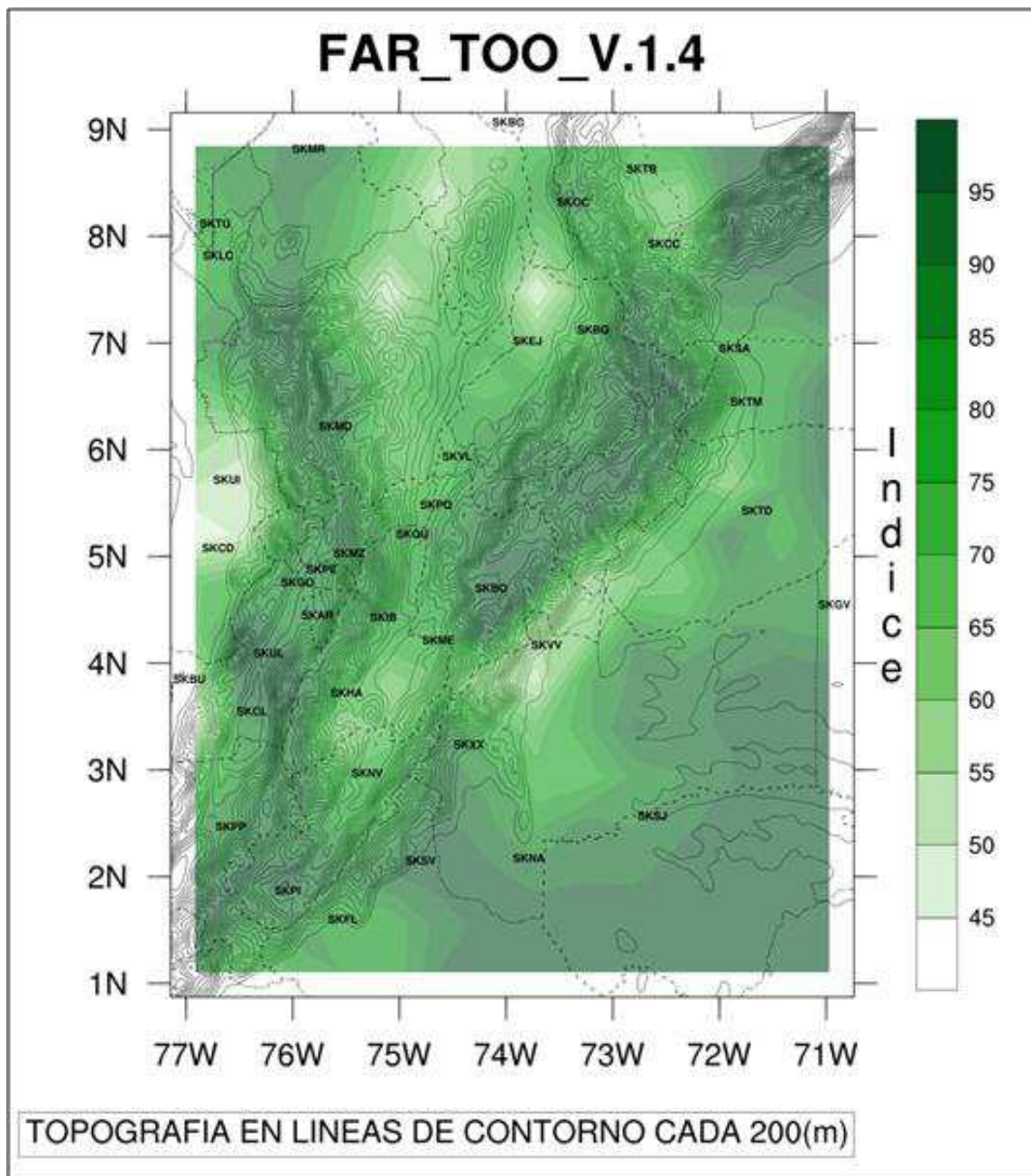


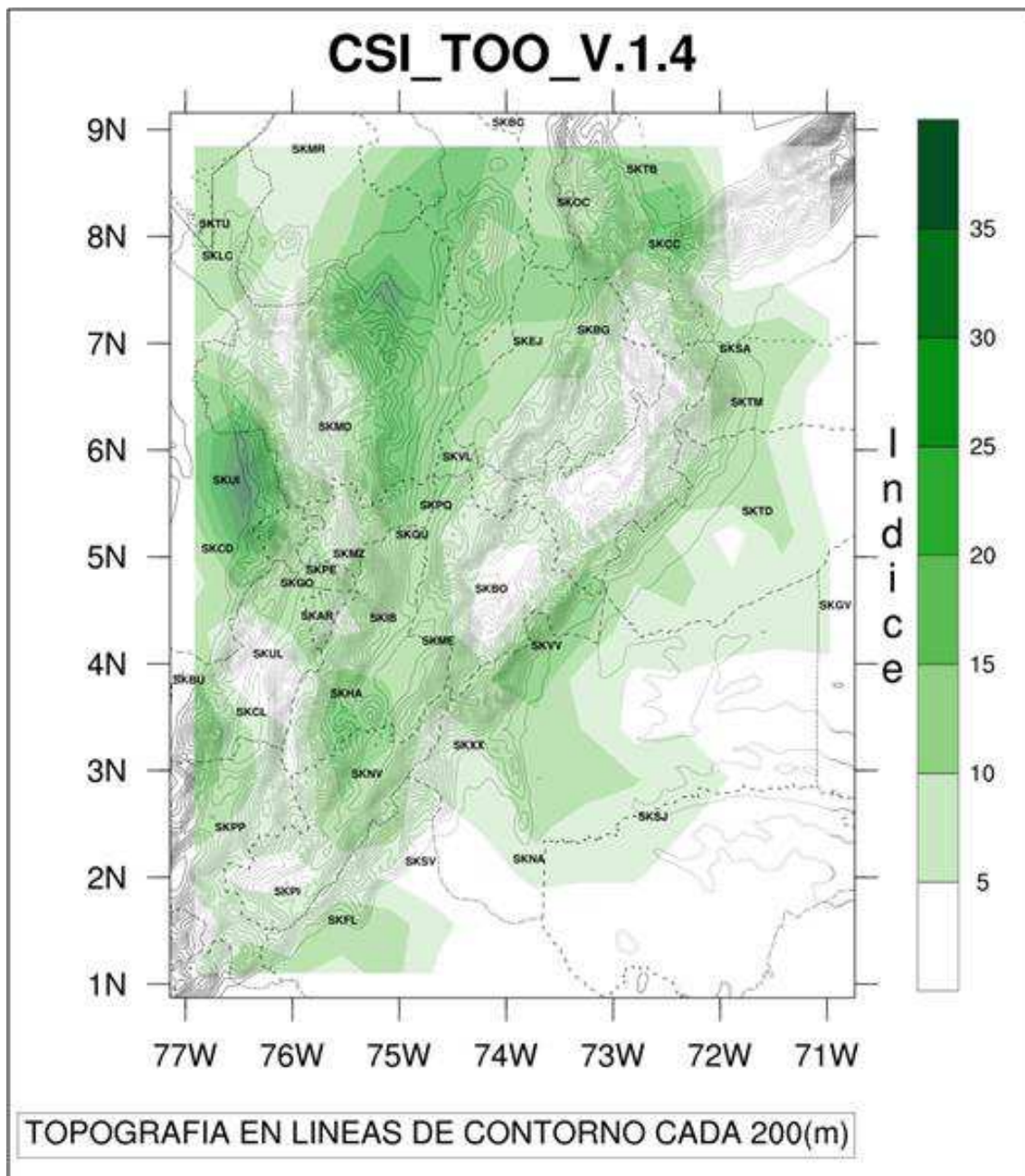


Técnicas Orientadas a Objetos con Variable Discreta









Bibliografía

- [1] Aguilar, E., J., Sigró, M. Brunet, 2009: RClimdex con funcionalidades extra de control de calidad. Manual de Uso, versión 1.0. Documento descargado del Climate Change Research Group, University Rovira i Virgili, Tarragona, España. 12 pp.
- [2] Ackerman, A. S., Toon, O. B., Stevens, D. E., Heymsfield, A. J., Ramanathan, V., & Welton, E. J. (2000). Reduction of tropical cloudiness by soot. *Science*, 288(5468), 1042-1047.
- [3] Baldwin, M. E., J. S. Kain and S. Lakshmvarahan, 2005: Development of an Automated Classification Procedure for Rainfall Systems. *Mon. Wea. Rev.* 133, 844 - 862.
- [4] Baldwin, M. E., S. Lakshmvarahan, and J. S. Kain, 2001. Verification of mesoscale features in NWP models. Preprints, *Ninth Conf.onMesoscale Processes*, Fort Lauderdale, Fl, Amer. Meteor. Soc., 255–258.
- [5] Beguería, S., & Lorente, A. (1999). Distribución espacial del riesgo de precipitaciones extremas en el Pirineo Aragonés Occidental. *Geographicalia*, 37, 17-36.

[6] Bianchi, A. R., Yáñez, C. E., & Acuña, L. R. (2005). Base de datos mensuales de precipitaciones del Noroeste Argentino. *Instituto Nacional de tecnología Agropecuaria. Proyecto Riesgo Agropecuario, Convenio Específico*, (3), 41.

[7] Brooks, H. E., & Doswell III, C. A. (1996). A comparison of measures-oriented and distributions-oriented approaches to forecast verification. *Weather and forecasting*, 11(3), 288-303.

[8] BUCARAMANGA. (2010). Lluvia histórica de 12 horas inundó a Barrancabermeja. Junio 10, 2010, de EL TIEMPO Sitio web: <http://www.eltiempo.com/archivo/documento/CMS-7735715>

[9] Cáceres, L., R.H., 2007: "Validación de la precipitación simulada por el modelo Weather Research And Forecasting (WRF) en el sur del territorio colombiano", Universidad Nacional de Colombia, tesis maestría, 120 p.

[10] Case, J. L., Manobianco, J., Lane, J. E., Immer, C. D., & Merceret, F. J. (2004). An objective technique for verifying sea breezes in high-resolution numerical weather prediction models. *Weather and forecasting*, 19(4), 690-705.

[11] Caviedes, C. (2001). El Niño in history: storming through the ages. Gainesville, Florida: University Press of Florida.

- [12] Clark, A.J., W.A. Gallus Jr. and T.C. Chen 2007: Comparison of the Diurnal Precipitation Cycle in Convection-Resolving and Non-Convection Resolving Mesoscale Models. *Mon. Wea. Rev.* 135, 3456- 3473.
- [13] Clause, R., Facy L. (1972). *Las nubes*, Ediciones Martínez Roca S.A., Barcelona, España.
- [14] Colle, B.A., K. J. Westrick & C.F. Mass 1999. Evaluation of MM5 and Eta-10 Precipitation Forecasts over the Pacific Northwest during the Cool Season, *WeaAnd Forecasting*, 14, p. 137- 154.
- [15] Davis, C., B. Brown, and R. Bullock, 2006. Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, 134, 1772–1784
- [16] Davis, C., B. Brown, and R. Bullock, 2006. Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, 134, 1772–1784.
- [17] Dillon, M. E., García Skabar, Y., & Nicolini, M. (2013). Desempeño del pronóstico de modelos de alta resolución, en un área limitada: análisis de la estación de verano 2010-2011. *Meteorologica*, 38(2), 69-89.

[18] Done, J., C. A. Davis and M. Weisman 2004: The next generation of NWP: explicit forecasts of convection using the weather research and forecasting (WRF) model. *Atmos. Sci. Let.* 5: 110–117

[19] Ebert, E.E. & J.L. McBride, 2000. Verification of precipitation in weather systems: determination of systematic errors. *J. of hydrology.* 1-24.

[20] El tiempo. (2008). 84 MUERTOS POR INUNDACIONES EN BRASIL. Noviembre 26, 2008, de EL TIEMPO Sitio web: <http://www.eltiempo.com/archivo/documento/MAM-3206386>

[21] Fowle, M. A., & Roebber, P. J. (2003). Short-range (0-48 h) numerical prediction of convective occurrence, mode, and location. *Weather and forecasting*, 18(5), 782-794.

[22] Gandin, L.S., & A.H. Murphy, 1992: Equitable skill scores for categorical forecasts. *Mon Wea Rev.*, 120, 361 – 370.

[23] GEORGE P. CRESSMAN, 1959: AN OPERATIONAL OBJECTIVE ANALYSIS SYSTEM. *Mon. Wea. Rev.*, 87, 367–374.

[24] Golvez J.M., Douglas M.W., 2006: Modulation of rainfall by lake Titicaca using the WRF model. Proceedings of 8th ICSHMO. Foz do Iguacu, Brazil, April 24-28. P 745-752.

[25] Guerrero Higuera, Á. M., García Ortega, E., Matellán Olivera, V., & Sánchez Gómez, J. L. (2013). Procesamiento paralelo de los pronósticos meteorológicos del modelo WRF mediante NCL.

[26] Hanssen, A.W. & W.J.A. Kuipers, 1965: On the relationship between the frequency of rain and various meteorological parameters. Mededeelingen en Verhandelingen, No 81, Royal Netherlands Meteorological Institute, 65 pp.

[27] Hitchens, N. M., M. E. Baldwin & R.J. Trapp, 2012: An Object-Oriented Characterization of Extreme Precipitation-Producing Convective Systems in the Midwestern United States. *Mon. Wea. Rev.* 140, p.1356- 1366.

[28] Hoffman, R. N., Liu, Z., Louis, J. F., & Grassoti, C. (1995). Distortion representation of forecast errors. *Monthly Weather Review*, 123(9), 2758-2770.

[29] IBAGUÉ. (2014). Inundaciones y carros aplastados por torrencial aguacero en Ibagué. Octubre 7, 2014, de EL TIEMPO Sitio web: <http://www.eltiempo.com/archivo/documento/CMS-14649418>

[30] IDEAM. (2005). ATLAS CLIMATOLÓGICO DE COLOMBIA. Bogotá, Colombia: ISBN 958-8067-14-6.

[31] IGAC. (2010). Mapa físico e hidrográfico. Noviembre 20, 2014, de IGAC Sitio web:

[http://upload.wikimedia.org/wikipedia/commons/3/3e/Mapa de Colombia %28relieve %29.svg](http://upload.wikimedia.org/wikipedia/commons/3/3e/Mapa_de_Colombia_%28relieve%29.svg)

[32] IGAC. (2010). Mapa Hidrográfico. Noviembre 20, 2014, de IGAC Sitio web: <http://vamosporcolombia.files.wordpress.com/2011/02/hidrografia-colombia.jpg>

[33] Mejía, Hoy, S.M., 2008: “Validación de un esquema de parametrización de microfísica explícita mejorada, dentro del modelo mm5, para el pronóstico de la precipitación en Colombia”, Universidad Nacional de Colombia, tesis maestría, 91 Paginas.

[34] Mejía, I.S.. (2010). Geografía económica de la región Andina Oriental. Cartagena de Indias: ISSN 1692-3715.

[35] Nachamkin, J. E. (2004). Mesoscale verification using meteorological composites. *Monthly Weather Review*, 132(4), 941-955.

[36] Nurmi, P. (2003). Recommendations on the Verification of Local Weather Forecasts (at ECMWF Member States). Consultancy Report, ECMWF Operations Department.

[37] Ochoa, A., & Poveda, G., (2003). Diagnostics of the spatial distribution of climate change signal in Colombia's hydrometeorology.

[38] OMM, 2002: WMO Technical Document 1125, GCOS-76.

[39] Pérez, C., Poveda, G., Mesa, O., Carvajal, L. F., & Ochoa, A. (1998). Evidencias de cambio climático en Colombia: tendencias y cambios de fase y amplitud de los ciclos anual y semianual. *Boletín del Instituto Francés de Estudios Andinos*, 27(3), 537-546.

[40] Philander, S. G. H., Gu, D., Lambert, G., Li, T., Halpern, D., Lau, N. C., & Pacanowski, R. C. (1996). Why the ITCZ is mostly north of the equator. *Journal of Climate*, 9(12), 2958-2972.

[41] Poveda, G. (2004). La hidroclimatología de Colombia: una síntesis desde la escala inter-decadal hasta la escala diaria. *Rev. Acad. Colomb. Cienc*, 28(107), 201-222.

[42] Poveda Jaramillo, G., Ramírez Osorio, J. M., & Jaramillo Robledo, A. (2001). Modelación estocástica de la variabilidad anual e interanual de la humedad del suelo en Colombia bajo diferente cobertura vegetal.

[43] Poveda, G., & Rojas, W. (1997). Evidencias de la asociación entre brotes epidémicos de malaria en Colombia y el fenómeno El Niño-Oscilación del Sur. *Revista de la Academia Colombiana de Ciencias*, 21(81), 421-429.

[44] Retallack, B. J.. (1974). Compendio de Meteorología. Parte 2-Meteorología Física. En OMM-No.364(pp. 119-136). Ginebra, Suiza: ISBN 92-63-30364-7.

- [45] Rogers, R. R., & Yau, M. K. (1989). A short course in cloud physics, International series in natural philosophy.
- [46] RT. (2012). Las inundaciones más devastadoras del siglo XXI. julio 8, 2012, de RT en español Sitio web: <http://actualidad.rt.com/actualidad/view/48797-Las-inundaciones-mas-devastadoras-del-siglo-XXI>
- [47] Serna, N. J. (2013). Efecto de los fenómenos El Niño y La Niña sobre la frecuencia de los eventos de precipitación extrema en la Sabana de Bogotá, Universidad Nacional de Colombia, Tesis de Maestría. p. 19.
- [48] Skok, G., Tribbia, J., Rakovec, J., & Brown, B. (2009). Object-based analysis of satellite-derived precipitation systems over the low-and midlatitude Pacific Ocean. *Monthly Weather Review*, 137(10), 3196-3218.
- [49] Snow, J. W.. (1976) The climate of northern South America, En: Schwerdtfeger, W. (ed.), *Cimates of Central and South America*, Elsevier, Amsterdam. 295-403
- [50] STANSKI, Henry R.; WILSON, Laurence J.; BURROWS, William R. (1989). *Survey of common verification methods in meteorology*. Geneva: World Meteorological Organization.
- [51] Stevens, B. (2005). Atmospheric moist convection. *Annu. Rev. Earth Planet. Sci.*, 33, 605-643.

[52] Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J. & Midgley, B. M. (2013). IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change.

[53] Strabler, A. N., & Strabler, A. H. (1989). Geografía física. *Barcelona, Ediciones Omega SA Platô*, 2. p. 767

[54] Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment* (1967), 8(12), 1251-1256.

[55] Uribe, Cortes, A., (2012). "Estimación de la capacidad del modelo WRF para pronosticar eventos extremos asociados con altas precipitaciones en la región Andina Colombiana.", Universidad Nacional de Colombia, tesis maestría, 134 Páginas.

[56] Wang W., Huang S., Li V., 2006: Sensitive numerical simulation and analysis of rainstorm using nested WRF model. *Journal of hydrodynamics*. V.18, 578-586.

[57] Weisman, M. L, C. Davis, W. Wang, K. W. Manning and J. B. Klemp, 2008. Experiences with 0–36-h Explicit Convective Forecasts with the WRF-ARW Model. *Wea. and Forecasting*, 23, 407-437.

[58] Weiss A., Kein B., Levit C., Beldwin D., Bright E., 2005: Use of guidance from high resolution Weather Research and Forecasting (WRF) model to forecast severe convective storms in the United States. *Geophysical Research Abstracts*, v.7, 08404.

[59] Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press. p.243

[60] Wilks, D.S., 1995. *Statistical Methods in the Atmospheric Sciences: An Introduction*. Acad. Press, 467 p.

[61] WILLIAMSON, D. L. (1981). Storm track representation and verification. *Tellus*, 33(6), 513-530.

[62] Yoshikane T., Ma X., Kimura F., Mara M., 2005: Regional Climatic Simulation for hydrological model using WRF model around Yellow river Basin.

[63] Zipser, E. J., Liu, C., Cecil, D. J., Nesbitt, S. W., & Yorty, D. P. (2006). Where are the most intense thunderstorms on Earth?. *Bulletin of the American Meteorological Society*, 87(8), 1057-1071.

[64] Zuluaga, M. D., Posada, J. E., & Poveda, G. (2001). Relación entre las fases de la oscilación cuasibienal (QBO), la precipitación mensual y la radiación de onda larga emergente en Colombia. In *Memorias IX Congreso Latinoamericano e Ibérico de Meteorología, Buenos Aires, Argentina*.