

UNIVERSIDAD NACIONAL DE CORDOBA
COMISION NACIONAL DE ACTIVIDADES ESPACIALES
FACULTAD DE MATEMÁTICA, ASTRONOMÍA Y FÍSICA
INSTITUTO DE ALTOS ESTUDIOS ESPACIALES
MARIO GULICH



Universidad
Nacional
de Córdoba



**Determinación de Tasa de Precipitación
a partir de datos ATMS mediante un algoritmo
basado en Redes Neuronales Artificiales**

Trabajo de grado presentado ante la Facultad de Matemática,
Astronomía y Física como requisito para obtener el grado de Magister
en Aplicaciones Espaciales de Alerta y Respuesta Temprana a Emergencias por:

Erith Alexander Muñoz Rios

Licenciado en Física

Director

PhD. Francesco Di Paola

Consejo Nacional de Investigaciones de Italia (CNR-Italia)

Co-Director

Lic. Mario Alberto Lanfri

Instituto de Altos Estudios Espaciales "Mario Gulich"

21 DE MARZO DE 2014

Abstract

En este trabajo, se presenta un algoritmo para estimación de tasa de precipitación a partir de datos provenientes del sensor *ATMS* a bordo de la plataforma espacial Suomi-NPP. El algoritmo aprovecha la capacidad de penetración en las nubes asociada a las microondas pasivas, así como también la sensibilidad ante precipitaciones, cristales de hielo y nubes de agua precipitable, mediante el entrenamiento de redes neuronales. Con este fin se entrenan 4 redes neuronales, 2 para tierra y 2 para océano, de las cuales 1 corresponde a periodo de verano y la otra a invierno respectivamente, haciendo uso de datos simulados para los 22 canales que conforman el sensor *ATMS*. El algoritmo demuestra alto potencial para reproducir patrones de precipitación, así como una capacidad satisfactoria para la estimación de la magnitud de tasa de precipitación.

Pacs 07.05.Mh, 07.05.Rm, 07.05.Tp, 07.57.Hm, 07.87.+v, 93.30.Hf, 93.85.Pq, 94.05.Dd, 94.05.Sd, 95.30.Jx.

Keywords: Microondas Pasivas, Estimación de Precipitación, Teledetección de la Atmósfera, Redes Neuronales.

Estructura del Trabajo

El presente trabajo está conformado por 5 capítulos y 1 apéndice, en este sentido a continuación se presenta una descripción general del contenido de cada una de las partes del trabajo en aras de proporcionar al lector un bosquejo general de esta obra, y por ende comprender el objeto de cada parte en relación al desarrollo de los objetivos de esta investigación.

CAPITULO 1: INTRODUCCIÓN

En este capítulo se presenta la preparación del proyecto de investigación, lo cual implica la descripción del problema asociado a la necesidad de desarrollar algoritmos de estimación de precipitación basados en sensores satelitales. Por otro lado, se plantea como objetivo general el desarrollo del mencionado algoritmo, desglosando el mismo en 6 objetivos específicos. Adicionalmente se hace una breve descripción sobre la características de la investigación al presentar la justificación, limitaciones, delimitación y alcance de la investigación, esto con la finalidad de proporcionar al lector una idea bastante clara sobre el enfoque específico del trabajo.

CAPITULO 2: FUNDAMENTOS TEÓRICOS

El capítulo 2 contiene un detallado análisis de antecedentes relacionados con el uso de sensores satelitales para la estimación de tasas de precipitación, así como también algunos trabajos cuyas metodologías han involucrado el uso de Redes Neuronales Artificiales. Por otro lado, se desarrollan desde el punto de vista conceptual tópicos relacionados con el uso de Redes Neuronales Artificiales y principios físicos asociados a la teledetección de la atmósfera, los cuales sirven de referencia para el diseño metodológico de la investigación y para el análisis y conclusiones en base a los resultados que tengan lugar a objeto de esta investigación.

CAPITULO 3: METODOLOGÍA DE LA INVESTIGACIÓN

Este capítulo comprende la estrategia llevada a cabo para el logro del objetivo general del trabajo. En primer lugar, se presenta el diseño del algoritmo de estimación de tasa de precipitación, basado en el entrenamiento de 4 redes neuronales; 2 para píxeles de océano y 2 para píxeles de tierra, una por cada estación equinoccial tomando como referencia el hemisferio norte. Cabe destacar que este proceso implica un proceso de selección de canales del sensor *ATMS* (en base a los principios físicos descritos en el capítulo 2) que son posteriormente usados como datos de entrada a las diferentes redes neuronales. Para el entrenamiento de las redes se ha usado datos simulados de temperatura de brillo para los 22 canales del sensor *ATMS*, la forma de obtención de estos datos también es presentada en este capítulo. Finalmente se presentan algunos casos de estudios y la metodología utilizada para indagar sobre la calidad de las estimaciones reportadas por el algoritmo desarrollado.

CAPITULO 4: RESULTADOS Y ANÁLISIS

Se presentan los resultados mas relevantes relacionados con el proceso de entrenamiento de las Redes Neuronales Artificiales, y la comparación de las salidas de tasas de precipitación arrojadas por el algoritmo con datos de estaciones pluviométricas y radares

meteorológicos. Posteriormente, haciendo referencia a la metodología planteada en el capítulo 3, disertar sobre los resultados obtenidos.

CAPITULO 5: CONCLUSIONES Y RECOMENDACIONES

Destacar las conclusiones de la investigación en referencia a la precisión de la salida del algoritmo, aportes realizados, novedades incorporadas en el sensor *ATMS* para la estimación de la precipitación. Y finalmente escribir recomendaciones que permitan mejorar los resultados y técnicas planteadas en este trabajo

APÉNDICE A: COMPILACIÓN Y EJECUCIÓN DEL ALGORITMO DE ESTIMACIÓN DE PRECIPITACIÓN BAJO AMBIENTE LINUX

Este apéndice contiene un breve manual para el uso del algoritmo desarrollado, que incluye una guía que explica los pasos de ejecución desde el momento que se descargan los datos del sensor *ATMS* hasta la salida de los mapas de precipitación.

Índice general

1. INTRODUCCIÓN	1
1.1. Planteamiento del Problema	1
1.1.1. Precipitación	1
1.1.2. Medición de Tasas de Precipitación (TP)	2
1.1.3. Estimación de Precipitación usando Sensores Remotos	3
1.2. Justificación de la Investigación	6
1.3. Objetivos de la Investigación	7
1.3.1. Objetivo General de la Investigación	7
1.3.2. Objetivos Específicos	8
1.4. Alcance, limitaciones y delimitación de la Investigación	8
1.4.1. Alcance	8
1.4.2. Limitaciones	9
1.4.3. Delimitación	9
1.4.4. Colofón	10
2. FUNDAMENTOS TEÓRICOS	11
2.1. Antecedentes de la Investigación	12

2.1.1.	Estimación de Precipitación usando Sensores Remotos Satelitales	12
2.1.2.	Características Generales del Sensor <i>ATMS</i>	21
2.1.3.	Estimación de Precipitación mediante Redes Neuronales Artificiales usando datos satelitales	23
2.2.	Principios Físicos de la Estimación de Precipitaciones mediante Microondas Pasivas	35
2.2.1.	Estructura de la Atmósfera Terrestre	36
2.2.2.	Composición de la Atmósfera Terrestre	38
2.2.3.	La función de Planck y la Teoría de Radiación de Cuerpos Negros	39
2.2.4.	Ley de Stefan-Boltzmann	42
2.2.5.	Temperatura de Brillo (BT)	42
2.2.6.	La Ecuación de Transferencia Radiativa (<i>ETR</i>)	43
2.3.	Redes Neuronales Artificiales (<i>ANNs</i>)	44
2.3.1.	Estructura de una <i>ANN</i>	45
2.3.2.	Modelo Matemático de una Neurona Artificial	46
2.3.3.	Perceptrón de una Capa	47
2.3.4.	Perceptrón Multi-Capas	49
2.4.	Entrenamiento de <i>ANNs</i>	51
2.4.1.	Inicialización de la <i>ANN</i>	52
2.4.2.	Entrenamiento mediante Propagación de Errores hacia Atrás (<i>BP</i>)	53
2.5.	Características Generales de las Librerías Open NN y ALGLIB	54
2.5.1.	Características de ALGLIB	54
2.5.2.	Características de Open NN	55
3.	METODOLOGÍA DE LA INVESTIGACIÓN	56
3.1.	Algoritmo de Estimación de Tasa de Precipitación (<i>RR</i>)	58
3.1.1.	Diseño de la Red Neuronal	59
3.1.2.	Observación de la Atmósfera a partir de los canales <i>ATMS</i>	63
3.1.3.	Selección de Canales <i>ATMS</i>	70
3.2.	Datos de Entrenamiento	72
3.3.	Casos de Estudio para el Análisis Comparativo de la Salida de Tasa de Pre- cipitación del Algoritmo	72

4. RESULTADOS Y ANÁLISIS	74
4.1. Selección de canales para el entrenamiento de las ANNs	74
4.1.1. Combinación de Canales para el Algoritmo de Estimación de Tasa de Precipitación sobre píxeles de superficie en tierra	75
4.1.2. Combinación de Canales para el Algoritmo de Estimación de Tasa de Precipitación sobre píxeles de Océano	77
4.2. Proceso de Entrenamiento de las ANNs	78
4.2.1. Entrenamiento de la Red Neuronal para la estimación de precipita- ciones sobre píxeles de tierra en época de verano	78
4.2.2. Entrenamiento de la Red Neuronal para la estimación de precipita- ciones sobre píxeles de océano en época de verano	84
4.2.3. Entrenamiento de la Red Neuronal para la estimación de precipita- ciones sobre píxeles de tierra en época de invierno	88
4.2.4. Entrenamiento de la Red Neuronal para la estimación de precipita- ciones sobre píxeles de océano en época de invierno	90
4.2.5. Colofón	92
4.3. Casos de Estudio	94
4.3.1. CASO A.1: Precipitación sobre Kansas, Oklahoma, Misuri y noroes- te de Texas del 29 de Abril de 2012	94
4.3.2. CASO A.2: Precipitación sobre Oklahoma, Kansas, Misuri e Illinois del 30 de Abril de 2012	99
4.3.3. CASO B.1: Evento de precipitación 9 de Marzo de 2013	103
4.3.4. CASO B.2: Evento de precipitación 10 de Marzo de 2013	107
4.3.5. CASO C.1: Evento de precipitación 8 de Mayo de 2013	109
4.3.6. CASO C.2: Evento de precipitación 9 de Mayo de 2013	111
5. CONCLUSIONES	114
5.1. Conclusiones	114
5.1.1. Sobre la selección de canales <i>ATMS</i> para el entrenamiento de redes neuronales para la elaboración del algoritmo de estimación de tasa de precipitación	114

5.1.2. Sobre el uso de Redes Neuronales para el diseño de Algoritmos de Estimación de Tasa de Precipitación	115
5.1.3. Sobre las Salidas del Algoritmo	117
5.2. Recomendaciones	118
5.3. Próximas Investigaciones	119
REFERENCIAS	121
A. COMPILACIÓN Y EJECUCIÓN DEL ALGORITMO DE ESTIMACIÓN DE PRECIPITACIÓN BAJO AMBIENTE LINUX	131
A.1. Descarga de Datos <i>ATMS</i>	132
A.2. Lectura de Ficheros HDF5, conversión a Temperatura de Brillo y Creación de archivos de datos de entrada para el algoritmo	134
A.2.1. Conversión a <i>BT</i> de los 22 Canales	135
A.2.2. Conversión a <i>BT</i> de canales pre-seleccionados	137
A.3. Calculo de la tasa de precipitación mediante la red neuronal	138
A.4. Elaboración de los archivos <i>ascii</i> para la generación de mapas de precipitación	139
A.5. Generación de los mapas de precipitación	140
A.6. Tiempo Estimado de Ejecución del Algoritmo	140

Índice de figuras

1.1. Estimación de Tasa de Precipitación sobre Venezuela usando datos del sensor AMSU-B, mediante el algoritmo 183-WSL.	5
2.1. Rangos de aplicación y sensibilidad de los canales del <i>TRMM</i> para medición de parámetros relacionados a procesos de precipitación (Strangeways, 2007)	16
2.2. Funciones de Pesos de Temperatura para los canales del <i>AMSU-A</i> (Chen et al., 2006)	19
2.3. Funciones de Pesos del Vapor de Agua para los canales del <i>AMSU-B</i> (Chen et al., 2006)	20
2.4. Estructura de la Red Neuronal usada por Bellerby et al. (2000)	24
2.5. Diagrama de Dispersión entre la salida de la <i>ANN</i> y el <i>MM5</i> a una resolución de $5, 2^\circ$ por Leslie et al. (2008)	33
2.6. Estructura térmica de la atmósfera con sus respectivas capas (Brasseur and Salomon, 2005).	37
2.7. Logaritmo de la Radianza Espectral $B_\lambda(T)$ graficado respecto al logaritmo de la longitud de onda λ asociada a un cuerpo de $T = 6000K$, temperatura típica de la fotosfera solar, y también graficado contra un cuerpo de $T = 288K$ que es la temperatura media de la superficie terrestre (Andrews, 2010)	41

2.8. Modelo matemático de una neurona (Xingui and Shaohua, 2010).	46
2.9. Modelo del Perceptron de una Capa (Hsieh, 2009).	47
2.10. Funciones de activación: a) Escalón , b) Lineal, c) Logística, d) Hiperbólica (Mas and Flores, 2008).	48
2.11. <i>MLP</i> , se muestra la capa de entrada de datos, una capa oculta, y una de capa de salida (Hsieh, 2009).	50
3.1. Etapas para el cumplimiento del objetivo general del proyecto de investigación.	57
3.2. Diagrama de Flujo del algoritmo para la estimación de tasa de precipitación.	58
3.3. Diseño secuencial de la topología de red neuronal de dos capas ocultas . . .	60
3.4. Transmitancia Vertical en la Atmósfera en el rango microondas (www.amsu.cira.colostate.edu)	64
3.5. Temperatura de Brillo para los canales 1 hasta el 6 del sensor <i>ATMS</i> ; los círculos negros resaltan cristales de hielo asociados a desarrollos de nubes convectivas, los azules están relacionados con precipitación o nubes de agua líquida y los círculos rojos a nubes sobre pixeles de océano	66
3.6. Temperatura de Brillo para los canales 7 hasta el 12 del sensor <i>ATMS</i>	67
3.7. Temperatura de Brillo para los canales 13 hasta el 18 del sensor <i>ATMS</i> . De acuerdo a la figura 2.2, los canales 13,14 y 15 tienen poca relación con los mecanismos asociados a la precipitación, mientras que en el resto de las imágenes los círculos negros corresponden a cristales de hielo en nubes convectivas, los azules precipitación y nubes de agua líquida, y en púrpura nubes sobre océano	68
3.8. Temperatura de Brillo para los canales 19 hasta el 22 del sensor <i>ATMS</i> . En los círculos negros se aprecian cristales de hielo asociados a nubes de alto desarrollo vertical, mientras que los púrpuras se relacionan con sistemas de nubes sobre océano	69
4.1. Comparación de la salida de la red neuronal con los datos de entrenamiento para la estimación de precipitación sobre pixeles de tierra en época de verano	80

4.2. Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en <i>Open NN</i> del caso <i>LS_1</i>	81
4.3. Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en <i>Open NN</i> del caso <i>LS_2</i>	82
4.4. Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en <i>Open NN</i> del caso <i>LS_3</i>	83
4.5. Comparación de la salida de la red neuronal con los datos de entrenamiento para la estimación de precipitación sobre pixeles de océano en época de verano.	85
4.6. Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en <i>Open NN</i> del caso <i>OS_1</i>	86
4.7. Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en <i>Open NN</i> del caso <i>OS_2</i>	87
4.8. Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en <i>Open NN</i> del caso <i>OS_3</i>	87
4.9. Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento <i>LW_1</i>	89
4.10. Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento <i>LW_2</i>	89
4.11. Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento <i>LW_3</i>	90
4.12. Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento <i>OW_1</i>	91
4.13. Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento <i>OW_2</i>	91

4.14. Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento <i>OW_3</i>	92
4.15. Precipitación acumulada para 24 horas del 29 de Abril de 2012 en Estados Unidos. (Fuente: http://water.weather.gov/precip/)	94
4.16. Precipitación horaria para las 09:00 horas UTC, del 29 de Abril de 2012 en los estados de Kansas, Oklahoma, Misuri y Texas de los Estados Unidos. (Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)	96
4.17. Tasa de Precipitación Instantánea, para la escena <i>ATMS</i> barrida entre las 08:24 y 8:32 UTC del 29 de Abril de 2012, estimada mediante la red <i>LS_1</i>	96
4.18. Tasa de Precipitación Instantánea sobre el noroeste del estado de Texas para las 08:24 UTC del 29 de Abril de 2012, estimada mediante la red <i>NEXRAD</i>	97
4.19. Tasa de Precipitación Instantánea sobre el norte del estado de Oklahoma y sur de Kansas para las 08:28 UTC del 29 de Abril de 2012, estimada mediante la red <i>NEXRAD</i>	98
4.20. Precipitación acumulada para 24 horas del 30 de Abril de 2012 en los estados de Texas, Kansas, Oklahoma, Misuri y Illinois de los Estados Unidos. (Fuente: http://water.weather.gov/precip/)	101
4.21. Precipitación horaria para las 09:00 horas UTC, del 30 de Abril de 2012 en los estados de Texas, Kansas, Oklahoma, Misuri y Illinois de los Estados Unidos. (Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)	101
4.22. Tasa de Precipitación Instantánea, para la escena <i>ATMS</i> barrida entre las 08:00 y 8:08 UTC del 30 de Abril de 2012.	102
4.23. Precipitación acumulada para 24 horas del 09 de Marzo de 2013 en los Estados Unidos. Se han resaltado tres zonas de precipitación, en las cuales se aprecian precipitaciones máximas superiores a los 25mm (Fuente: http://water.weather.gov/precip/)	103
4.24. Precipitación horaria para las 09:00 horas UTC, del 9 de Marzo de 2013 en los Estados Unidos. Se muestran 3 zonas de interes de precipitación las cuales son comparadas con las estimaciones del algoritmo (Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)	104

4.25. Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 08:30 y 8:38 UTC del 29 de Abril de 2012. Lo zona 3 muestra precipitaciones sobre el norte y sureste de Iowa con tasas de precipitaciones máximas de 12mm/h y 11mm/h respectivamente, mientras que la zona 2 muestra una parte de la precipitación sobre Misuri con un máimo de precipitación de 8mm/h 104

4.26. Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 08:38 y 8:46 UTC del 29 de Abril de 2012. Se aprecia la segunda parte del patrón mostrado en el mapa de precipitación horaria por la *NWS*, en donde en el foco de precipitación que se aprecia en Oklahoma tiene una tasa de precipitación superior a 15mm/ 105

4.27. Precipitación horaria para las 20:00 horas UTC, del 9 de Marzo de 2013 en los Estados Unidos. Para este instante la precipitación sobre los estados de Iowa, Misuri, Oklahoma y Texas continua con un ligero corrimiento hacia el sureste (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/) 105

4.28. Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 19:58 y 20:06 UTC del 9 de Marzo de 2013. Se aprecia que el patrón de precipitación sobre Texas, Oklahoma y Misuri es bien reproducido por el algoritmo, con tasas máximas de precipitación de 11mm/h, 8mm/h y 11mm/h respectivamente. 106

4.29. Precipitación acumulada para 24 horas del 10 de Marzo de 2013. Se aprecian altas precipitaciones sobre gran parte de la región central del país, de las cuales resaltan las máximas tasas de precipitación sobre los estados de Oklahoma y Arkansas, las cuales superan los 50mm (Fuente:<http://water.weather.gov/precip/>) 107

4.30. Precipitación horaria para las 09:00 horas UTC, del 10 de Marzo de 2013 en los Estados Unidos. Se aprecian 4 zonas de precipitación con máximos de precipitcion que varían desde 12,5 mm (0.50 pulgadas) hasta 19 mm (0.75 pulgadas) (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/) 107

4.31. Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 08:14 y 08:22 UTC del 9 de Marzo de 2013. Se observa buena tendencia en la estimación de tasa de precipitación para las zonas de interés resaltadas, sin embargo también se aprecian una cantidad considerable de píxeles mal clasificados como precipitantes en comparación con el mapa de precipitación horario presentado por la *NWS*. 108

4.32. Precipitación acumulada para 24 horas del 8 de Mayo de 2013. Para este escenario la adquisición de datos <i>ATMS</i> estuvo orientada a la región central del país, lo cual comprende los estados de Kansas, Nebraska, Dakota del Sur e Iowa (Fuente: http://water.weather.gov/precip/)	109
4.33. Precipitación horaria para las 09:00 horas UTC, del 8 de Mayo de 2013 en los Estados Unidos. Se han etiquetado 4 zonas de precipitación para ser comparados con la salida del algoritmo(Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)	109
4.34. Tasa de Precipitación Instantánea, para la escena <i>ATMS</i> barrida entre las 08:06 y 08:14 UTC del 8 de Mayo de 2013. En la barrida de esta escena se observa la primera parte del patrón observado en el mapa de precipitación horaria, cabe destacar la buena reproducción del patrón y la buena estimación de los niveles de tasa de precipitación.	110
4.35. Tasa de Precipitación Instantánea, para la escena <i>ATMS</i> barrida entre las 08:14 y 08:22 UTC del 8 de Mayo de 2013. Segunda parte del patrón observado por la <i>NWS</i> , donde se aprecia nuevamente una reproducción bastante satisfactoria del mapa de precipitación, aún cuando se observan píxeles mal clasificados principalmente sobre los estados de Nueva México y Colorado.	110
4.36. Precipitación acumulada para 24 horas del 9 de Mayo de 2013 en los Estados Unidos. (Fuente: http://water.weather.gov/precip/)	111
4.37. Precipitación horaria para las 09:00 horas UTC, del 9 de Mayo de 2013 en los Estados Unidos. Se seleccionan 3 zonas de precipitación para la comparación sobre los estados de Oklahoma, Misuri y Iowa (Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)	111
4.38. Tasa de Precipitación Instantánea, para la escena <i>ATMS</i> barrida entre las 07:49 y 07:57 UTC del 9 de Mayo de 2013. Se aprecia una clara tendencia a reproducir en buena forma los patrones de precipitación, de igual forma los niveles de tasa de precipitación estimados son satisfactorios. Sin embargo se aprecia una sobrestimación de precipitación al sur de Kansas.	112
4.39. Precipitación horaria para las 20:00 horas UTC, del 9 de Mayo de 2013 en los Estados Unidos. De la cual se seleccionan dos focos de precipitación centrados sobre los estados de Texas e Illinois respectivamente (Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)	112

4.40. Tasa de Precipitación Instantánea, para la escena <i>ATMS</i> barrida entre las 19:09 y 19:17 UTC del 9 de Mayo de 2013. Se aprecia buena definición de los patrones de precipitación niveles de tasa de precipitación satisfactorios	113
A.1. Interfaz proporcionada por el NOAA-CLASS para la adquisición de datos <i>ATMS</i> , parte superior	132
A.2. Interfaz proporcionada por el NOAA-CLASS para la adquisición de datos <i>ATMS</i> , parte inferior	133
A.3. Visión general de una carpeta contentiva de archivos de datos <i>ATMS</i> descargados via ftp a partir del sistema CLASS de la NOAA	134
A.4. Visión general de la carpeta contentiva de los archivos .dat generados a partir de la ejecución del script Lanzador_HDF5_DAT_Open NN_channel.sh	137

Índice de cuadros

2.1. Frecuencia centrales de los canales <i>AMSU-A</i> (Staelin and Chen, 2000)	18
2.2. Frecuencia centrales de los canales <i>AMSU-B</i> (Staelin and Chen, 2000)	19
2.3. Frecuencia centrales de los canales <i>ATMS</i> (Surussavadee and Staelin, 2010)	22
2.4. Gases de concentración fija que componen la Atmósfera (Andrews, 2010) .	38
2.5. Gases de concentración variable en la Atmósfera (Andrews, 2010)	39
3.1. Primera fase de selección de canales del sensor <i>ATMS</i> , en la misma son des- cartados aquellos canales que requieren de un proceso de corrección antes de poder ser usados, y aquellos canales cuya información aporta poca infor- mación relacionada con el proceso de precipitación, en base a las funciones de peso. En este sentido una decisión SI implica que el canal pudiera ser usado en el algoritmo, mientras que un NO indica que el uso de dicho canal es descartado	71
4.1. Diseño de Experimentos para el entrenamiento de la <i>ANN</i> para el algoritmo de estimación de tasa de precipitación sobre tierra en periodo de Verano (el LS son siglas en inglés que hacen referencia a Land-Summer, lo cual se interpreta como dato para pixel sobre tierra en época de Verano)	76

4.2. Diseño de Experimentos para el entrenamiento de la ANN para el algoritmo de estimación de tasa de precipitación sobre tierra en periodo de Invierno (el LW son siglas en inglés que hacen referencia a Land-Winter, lo cual se interpreta como dato para pixel sobre tierra en época de Invierno)	76
4.3. Diseño de Experimentos para el entrenamiento de la ANN para el algoritmo de estimación de tasa de precipitación sobre océano en periodo de Verano (el OS son siglas en inglés que hacen referencia a Ocean-Summer, lo cual se interpreta como dato para pixel sobre océano en época de Verano)	77
4.4. Diseño de Experimentos para el entrenamiento de la ANN para el algoritmo de estimación de tasa de precipitación sobre océano en periodo de Invierno (el OW son siglas en inglés que hacen referencia a Ocean-Winter, lo cual se interpreta como dato para pixel sobre océano en época de Invierno)	77
4.5. Resumen de resultados de coeficientes de correlación para las redes neuronales de estimación de tasa de precipitación sobre píxeles de tierra en época de verano	84
4.6. Resumen de resultados de coeficientes de correlación para las redes neuronales de estimación de tasa de precipitación sobre píxeles de océano en época de verano	88
4.7. Mejores resultados de coeficientes de correlación obtenidos por el proceso de entrenamiento de las redes neuronales	93

CAPÍTULO 1

INTRODUCCIÓN

1.1. Planteamiento del Problema

1.1.1. Precipitación

Una de las variables meteorológicas mas difíciles de medir, por su alta variabilidad espacial y temporal, y a su vez de mayor relevancia, por su alto impacto, es la precipitación. Por este motivo, históricamente han sido diversas las técnicas implementadas para su medición y/o estimación. Dichas técnicas han estado en cada momento limitadas por la falta de una comprensión teórica satisfactoria sobre su dinámica y los mecanismos físico-químicos involucrados, así también como por el estado general de la tecnología disponible para el diseño de sistemas de estimación y medición.

La importancia de estimar o medir tasas de precipitación se atribuye, entre otros fac-

tores, a que es un elemento determinante dentro del modelo del ciclo hidrológico, el cual permite la permanencia y distribución de aguas dulces en el planeta, siendo éste un mecanismo primordial para el desarrollo y sustento de la vida. Por otro lado, dada la vinculación de la precipitación con el ciclo hidrológico, muchos sistemas de producción agrícola dependen en gran medida de la capacidad de pronosticar con suficiente tiempo de antelación valores de tasa de precipitación, y no sólo ésta, sino además los ciclos interanuales de precipitación, es decir la definición de temporadas, valores mínimos y máximos, tiempos de sequías, etc.

El pronóstico de la precipitación es de vital importancia en diversas áreas, como por ejemplo: para el diseño y desarrollo de sistemas de alerta y respuesta ante el riesgo de emergencia o desastres provocados por precipitaciones, como lo es el caso de inundaciones; en la planificación y control de tráfico aéreo, influye de forma determinante en la eficiencia de sistemas de comunicaciones inalámbricas y tal vez con menor carácter de vitalidad pero sin menospreciar la importancia, para la planificación de calendarios de eventos sociales, deportivos, culturales y políticos (Muñoz and Falcón, 2008).

1.1.2. Medición de Tasas de Precipitación (TP)

Convencionalmente, la medición de tasa de precipitación, es realizada mediante pluviómetros localizados en estaciones de mediciones terrestres. Existen diversos tipos de pluviómetros, pero en esencia todos miden la cantidad de agua caída expresada en milímetros de altura. Éste tipo de medidas son realizadas en tiempo real, lo que quiere decir que se realiza durante el evento de precipitación y por lo tanto por si solas no son mediciones suficientes para la elaboración de pronósticos. Sin embargo, la medición regular de tasas de precipitación mediante pluviómetros permite la construcción de series temporales de datos, climatológicamente representativos, que en base a técnicas estadísticas supervisadas crean la posibilidad de elaborar pronósticos de precipitación.

Ahora bien, la asimilación de datos de tasa de precipitación, que sirvan de entrada para modelos de pronósticos a partir de estaciones terrenas meteorológicas presenta serias dificultades. En primer lugar, puede mencionarse que la densificación espacial de dichas estaciones

colectoras de datos, en la mayoría de los casos es insuficiente, debido principalmente a limitaciones económicas y en algunos casos a la imposibilidad de construir estas estaciones en lugares remotos y de difícil acceso, para los cuales no resulten factible el mantenimiento operativo y la colecta de los datos. En segundo lugar, cabe destacar que generalmente los pluviómetros no registran la evolución temporal de la precipitación, sino el volumen precipitado durante un intervalo de tiempo.

En respuesta a las dificultades mencionadas en el párrafo anterior, se han desarrollado diversas metodologías alternativas para generar bases de datos operativas de tasa de precipitación. Por ejemplo, una técnica bastante empleada para resolver la insuficiencia de densificación de datos es el uso de métodos de interpolación, lo cual permite la definición de resoluciones espaciales requeridas a partir de la densificación de los datos obtenidos por las estaciones meteorológicas mediante métodos matemáticos, sacrificando para esto, calidad y representación de los datos.

En este contexto, otras técnicas importantes se basan en el uso de sensores remotos desde plataformas espaciales, las cuales han mostrado un gran potencial para la generación de datos meteorológicos por las características específicas que este tipo de mecanismos presentan tanto en resolución temporal como en resolución espacial, lo que permite una mejor caracterización climatológica tanto en escala global como local, según [Schonhuber et al. \(2008\)](#).

1.1.3. Estimación de Precipitación usando Sensores Remotos

En la actualidad, una importante cantidad de satélites están dedicados a la medición de variables atmosféricas. Satélites geoestacionarios como los Satélites Geoestacionarios de Operaciones Ambientales (GOES-NASA) ([Menzel and Purdom, 1994](#)) o Meteosat de Segunda Generación (MSG-ESA) ([Schmetz et al., 1998](#)), han aumentado notablemente la capacidad para hacer estimaciones sobre microestructuras de nubes y procesos de precipitación desde plataformas geoestacionarias ([Levizzani et al., 2001](#)). Recientemente, las observaciones de tope de nubes a partir de plataformas de órbitas geoestacionarias, son llevadas a cabo

con una resolución espacial que varía entre 1 y 8 km (Houl et al., 2008), esta es una resolución espacial que prácticamente no es factible de ser proporcionada mediante métodos convencionales de medición.

Además de los satélites geoestacionarios, diversas misiones de plataformas polares han sido llevadas a cabo para profundizar las capacidades de adquisición de datos meteorológicos. Entre los cuales, a modo de ejemplo, se puede mencionar la serie de satélites de la Administración Nacional para la Atmósfera y el Océano de los Estados Unidos de Norte América (NOAA, por sus siglas en inglés), los cuales cuentan con una gran variedad de sensores de observación de la atmósfera y de la Tierra. La mayoría de estas plataformas son usadas para proyectos de investigación y desarrollo de conocimientos para incrementar las aplicaciones espaciales en la descripción y monitoreo de fenómenos atmosféricos. Particularmente, en lo que respecta a datos y productos de precipitación, no son usados con frecuencia para pronóstico meteorológico operacional debido a la baja resolución temporal asociada a satélites de este tipo de órbitas ¹, sin embargo estos datos son buena fuente de entrada para modelos de predicción numérica del tiempo.

Los satélites dedicados al estudio de variables atmosféricas poseen sensores que reúnen características específicas para operar y cumplir la misión para la que dichas plataformas han sido desarrolladas. Los sensores en plataformas meteorológicas tienen canales de operación bien definidos dentro de bandas del visible (VIS), infrarrojo (IR) y microondas (MW) en el espectro electromagnético. La selección de dichos canales para el desarrollo del sensor, depende de las variables meteorológicas que se esperan medir. Un ejemplo de resultados obtenidos debido a la combinación de datos entre estas bandas del espectro es el trabajo realizado por Di Paola et al. (2012).

En este sentido, Kidd (2001) presentó una compilación de investigaciones desarrolladas con el objetivo de estimar tasas de precipitación a partir de satélites geoestacionarios, usando sensores en bandas ópticas e infrarrojo. Si bien es cierto que los resultados reportados representan en general buenas aproximaciones, hay que destacar que este tipo de sensores, basados en VIS y IR presentan una seria dificultad para relacionar características de los topes

¹12 horas generalmente

de las nubes con la precipitación observada en superficie, además debido a los altos niveles de extinción de hidrometeoros que tienen en estos rangos de longitudes de onda, resulta difícil que este tipo de sensores puedan suministrar información dentro de las nubes.

A diferencia de estos, los sensores que operan en MW en regiones de ondas milimétricas y submilimétricas proporcionan la capacidad de obtener gran información sobre parámetros dentro de las nubes, precipitación y a su vez perfiles de hidrometeoros (Houl et al., 2008). Diversos métodos han sido propuestos para estimar precipitación a partir de sensores satelitales de MW (Di Tomaso et al., 2009), por ejemplo en la figura 1.1 se muestra una estimación de tasa de precipitación para escenarios de tormenta severa realizada sobre Venezuela a partir del algoritmo 183-WSL basado en datos del sensor AMSU-B y desarrollado por Laviola and Levizzani (2011).

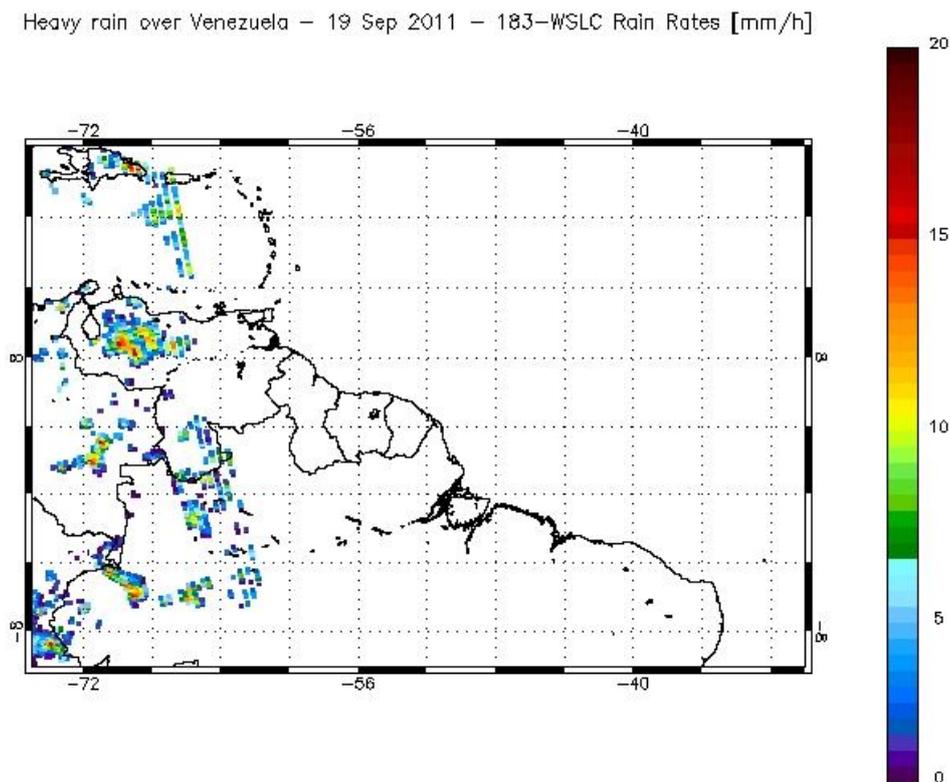


Figura 1.1: Estimación de Tasa de Precipitación sobre Venezuela usando datos del sensor AMSU-B, mediante el algoritmo 183-WSL.

Los sensores usados en microondas se dividen en activos y pasivos, los activos contie-

nen su propia fuente de energía para irradiar los blancos bajo estudio mientras que los pasivos son sensibles a las longitudes propias de radiación del blanco. Considerando que este trabajo se fundamenta en datos de microrondas pasivas, es importante mencionar que hay dos tipos de sensores pasivos de acuerdo a las características propias de medición, en primer lugar se mencionan los sensores utilizados para construir perfiles verticales de la atmósfera (Sounders) y en segundo lugar se tienen generadores de imágenes (Imagers), los cuales permiten obtener escenas de un blanco para diferentes longitudes de onda. Ahora bien, los Imagers operan en las regiones ventana (Channel Windows) del espectro de MW, para estos canales la atmósfera es casi transparente porque escapa de las líneas de absorción del oxígeno y el vapor de agua, lo que permite observar el contenido de agua en columnas de la atmósfera. Los sensores pasivos Sounders operan en líneas de absorción de MW para generar perfiles tanto para la temperatura como del contenido de vapor de agua de la atmósfera, sin embargo trabajos recientes muestran algunos usos para detectar precipitaciones en temporadas frías y también medir de forma indirecta precipitaciones ligeras sobre la superficie terrestre (Kidd, 2001, Levizzani et al., 2007).

En base a los argumentos planteados en este apartado, en este trabajo se propone la generación de datos de precipitación a partir del sensor ATMS, como una solución a la insuficiencia de datos basados en mediciones desde estaciones meteorológicas convencionales de baja densificación. Para esto se desarrolla un algoritmo que permita la estimación indirecta de tasas de precipitación a partir del entrenamiento de una Red Neuronal Artificial.

1.2. Justificación de la Investigación

En el presente plan de investigación se propone el desarrollo de un algoritmo para la estimación de tasas de precipitación, basado en Redes Neuronales Artificiales (ANNs, por sus siglas en inglés) y usando el radiómetro ATMS a bordo del satélite Suomi-NPP. En este sentido cabe destacar, que se han seleccionado datos provenientes de sensores remotos por la regularidad de la resolución espacial que proporcionan. El sensor ATMS, representa una herramienta de gran potencialidad para la descripción de procesos microfísicos de nubes, lo

cual sugiere una gran capacidad para la medición indirecta de precipitación. Además, este sensor reúne características de radiómetros como AMSU-A y AMSU-B/MHS de manera simultánea, lo cual basado en investigaciones destacadas sobre estos sensores (Laviola and Levizzani, 2011), le atribuye gran versatilidad para el propósito general planteado.

Respecto al uso de *ANNs* se puede comentar que, hasta el presente han sido reportados diversos trabajos, que han tenido como objetivo el desarrollo de algoritmos para la estimación de precipitaciones usando *ANNs* (Leslie et al., 2008, Sorooshian et al., 2008, Tapiador et al., 2004). La ventaja de usar *ANNs* reside en que pueden ser entrenadas mediante diversos métodos, utilizando datos de Temperatura de Brillo para los diferentes canales proporcionados por los sensores y valores de precipitaciones respectivos, este entrenamiento tiene como efecto que la *ANN* aprenda de forma implícita las relaciones entre los datos de entradas y salidas, para a partir de ella hacer procesos de estimación de precipitación. Esto representa una aproximación operativamente aceptable, ya que evita las dificultades inherentes al desarrollo de algoritmos basados en parámetros y relaciones físicas entre las variables.

En general, las estimaciones de tasa de precipitación por satélite se utilizan para la aplicación en regiones remotas, donde no hay presencia de estaciones pluviométricas. Sin embargo, la metodología planteada en esta investigación, podría servir de base para procesos de asimilación de datos satelitales a partir del sensor ATMS a bordo del Suomi-NPP para modelos numéricos de predicción del tiempo, modelos de alerta temprana a desastres naturales provocados por precipitaciones, modelos de incendios, entre otros.

1.3. Objetivos de la Investigación

1.3.1. Objetivo General de la Investigación

Desarrollar un algoritmo basado en Redes Neuronales Artificiales para estimación de tasa de precipitación, a partir de la Sonda de Microondas de Tecnología Avanzada (ATMS) a bordo del Satélite Suomi-NPP

1.3.2. Objetivos Específicos

1. Definir el estado del arte de sistemas de estimación de precipitación a partir de sensores remotos y la implementación de Redes Neuronales Artificiales
2. Establecer un criterio para la selección de los canales que se utilizarán como datos de entrada a la Red Neuronal, en base a las características espectrales del sensor ATMS.
3. Diseñar una estrategia para el entrenamiento de un sistema de Red Neuronal Artificial usando datos simulados de Temperatura de Brillo para frecuencias características del sensor ATMS y valores de tasa de precipitación respectivos.
4. Desarrollar un Algoritmo para la estimación de tasas de precipitación a partir de la Red Neuronal Artificial (ANN) previamente entrenada.
5. Calibrar el Algoritmo desarrollado usando datos reales de temperatura de Brillo del sensor ATMS en conjunto con valores de precipitación previamente validados.
6. Comparar la tasa de precipitación estimada mediante el algoritmo desarrollado, con mediciones obtenidas mediante pluviómetros en superficie u otra fuente de datos validados, como por ejemplo, modelos de predicción numérica del tiempo.

1.4. Alcance, limitaciones y delimitación de la Investigación

1.4.1. Alcance

La propuesta de investigación planteada, representa una solución al problema de estimar tasas de precipitación a partir de datos espaciales. Con este propósito, se entrenan redes neuronales artificiales que permitan resolver de forma implícita y mediante un algoritmo de aprendizaje los procesos físicos involucrados en el proceso de formación de precipitación, lo

cual quiere decir que las características físicas del fenómeno son tratadas de forma indirecta.

Por otro lado, las estimaciones de precipitación derivadas de este trabajo, se espera puedan servir como fuente de datos, tanto para estudios de índole climático, como para modelos meteorológicos. En este sentido, se descarta su uso para la elaboración operativa de pronósticos debido a la baja resolución temporal de la plataforma polar Suomi-NPP.

1.4.2. Limitaciones

Dentro del posible conjunto de limitaciones asociadas al diseño de la metodología planteada en este trabajo, vale la pena destacar que el sensor ATMS, se encuentra a bordo del satélite de órbita polar Suomi-NPP, el cual tiene observación de un mismo punto sobre la superficie terrestre en intervalos aproximados de 12 horas, sin embargo esto es variable y depende de la latitud, por ejemplo en los polos, se tiene una pasada cada 1 hora y 40 minutos. En las latitudes medias alrededor de 2 veces al día. Esto restringe la posibilidad de usar las estimaciones de precipitaciones que tengan lugar a partir de ésta investigación a su uso operativo desde el punto de vista meteorológico

1.4.3. Delimitación

El diseño del algoritmo de precipitación involucra el entrenamiento de 4 redes neuronales, que corresponden 2 para píxeles de tierra y 2 para píxeles que pertenecen a océano, y en cada caso una corresponde a condiciones de invierno y otra a verano. Por este motivo, el proceso de entrenamiento de las ANNs, contempla conjuntos de datos de diferentes estaciones temporales (verano e invierno), tanto para superficie como sobre océano, de tal forma que sean consideradas las variaciones climáticas que tienen lugar dentro de ciclos interanuales de forma implícita.

1.4.4. Colofón

Una vez desarrollada en el presente capítulo toda la información de carácter metodológica relacionada con esta investigación, es decir la descripción del problema, el objetivo general de la investigación, justificación, limitaciones, delimitaciones y alcance, en el siguiente capítulo se desarrollan y mencionan los diferentes aspectos teóricos sobre los cuales se fundamenta este trabajo.

CAPÍTULO 2

FUNDAMENTOS TEÓRICOS

En este capítulo se cubren áreas muy importante para el desarrollo de los objetivos planteados en la investigación. En primer lugar se presenta una compilación sobre los trabajos, mas relevantes, realizados en aplicaciones de sensores remotos y redes neuronales artificiales para estimar tasas de precipitación, así como los resultados y conclusiones que sirven de referencia para este trabajo. En segundo lugar, se muestra un estudio del contenido teórico en materia de teledetección de la atmósfera y redes neuronales artificiales, el mismo fundamenta y justifica el diseño de la presente investigación

2.1. Antecedentes de la Investigación

2.1.1. Estimación de Precipitación usando Sensores Remotos Satelitales

Muchos autores coinciden en que el uso de observaciones satelitales para estimar precipitaciones tuvo su origen a partir del lanzamiento del satélite de órbita polar *Television and Infrared Observation Satellite (TIROS-1)* en abril de 1960 (Barrett, 1997, Kidd, 2001, Liu et al., 2002), fecha a partir de la cual han sido desarrolladas diversas misiones espaciales orientadas a proporcionar productos satelitales que reúnan las características técnicas necesarias para la obtención de variables y parámetros atmosféricos de interés climático o meteorológicos. Estas características están fuertemente relacionadas al esfuerzo por optimizar el aprovechamiento de datos satelitales en el espectro óptico (VIS), infrarrojo (IR) y microondas (MW) (Barrett and Beaumont, 1994, Kidd and Levizzani, 2011). A continuación se presenta una breve descripción de las potencialidades de cada uno de estos tipos de sensores para estimar parámetros y variables atmosféricas.

Monitoreo Satelital de Precipitaciones desde sensores IR/VIS

El tipo de información atmosférica que puede ser obtenida desde datos satelitales, está relacionado directamente con las características de la resolución temporal, espacial, radiométrica y espectral de la plataforma de observación. Por ejemplo, el uso integrado de sensores con canales en bandas del espectro óptico y visual proporcionan información sobre toques de nubes, lo cual ha demostrado gran capacidad para la caracterización y clasificación de nubes en aras de obtener mediciones indirectas sobre tasas de precipitación en la superficie de la Tierra.

En este contexto es importante destacar que las investigaciones orientadas a la medición de precipitación a partir de datos satelitales generalmente tienen como objetivo el desarrollo, implementación y validación de métodos de medición indirecta de precipitación

y de otras variables meteorológicas. Un ejemplo de este tipo de métodos son las estimaciones de precipitación basadas en *IR/VIS*, las cuales son llevadas a cabo relacionando la temperatura de tope de nube derivada del sensor con la probabilidad de precipitación en superficie (Prigent, 2010).

Los sensores satelitales *IR* y *VIS* son, por lo general de tipo pasivo, sin embargo debido al hecho de que la Tierra y su atmósfera emiten radiación *IR* las 24 horas del día, los sensores *IR* son independientes de las características externas de iluminación, a diferencia de los sensores *VIS* que capturan la luz solar reflejada. La dependencia a la luz solar por parte de los sensores *VIS* compromete seriamente la capacidad de operación de este tipo de sensores en escenarios nocturnos. Unido a esto, debido a la necesidad de observar los cambios que tienen lugar dentro de escalas climatológicas o meteorológicas, se debe garantizar resoluciones temporales adecuadas.

Barrett (1997) sugiere que las bases para el monitoreo de precipitación satelital están referidas por un lado por Lethbridge (1967) quien consideró relaciones generales entre el brillo de nubes, temperatura y precipitación, y por otro lado por Barrett (1970) quien preparó mapas de precipitación mensuales basados en cartas de nubes satelitales reportando mejores resultados que aquellos generados desde estaciones en superficie, donde la calibración fue llevada a cabo a partir de datos de estaciones climáticas (Barrett and Martin, 1981). Estos estudios pioneros han precedido a una gran gama de aplicaciones de sensores remotos que han tenido lugar hasta el presente en el campo de la meteorología y climatología.

Monitoreo Satelital de Precipitaciones desde sensores de MW pasivas

Generalmente a los sensores de microondas pasivas (*PMW*, por sus siglas en inglés) se les atribuye dos ventajas importantes, la primera es que son sensibles a la frecuencia de absorción/emisión de las gotas de lluvia, y por otro lado, muestran gran penetración en la mayoría de tipos de nubes de precipitación facilitando la observación de hidrometeoros líquidos y de hielo, lo cual se traduce en mejoras sensibles a la capacidad de estimación de Tasa de Precipitación (New et al., 2001).

Históricamente se han empleado y desarrollado dos familias de sensores para la percepción remota en *PMW* cuya elección depende del objeto de estudio a realizar. La primera familia esta integrada por los sensores de escaneo, mejor conocidos en inglés como *Microwave Imager* (*MI*, por sus siglas en inglés), mientras que la segunda familia de sensores esta conformada por las *Microwave Sounder* (*MS*, por sus siglas en inglés). La diferencia principal que existe es que los *MI* permiten observar a la Tierra en las frecuencias de ventana atmosférica, mientras que los *MS* permiten la observación en las frecuencias de absorción de la atmósfera utilizando las banda de absorción de O₂ o vapor de agua.

De acuerdo a [Kidd \(2001\)](#) los primeros experimentos con *PMW* tuvieron lugar en 1972 a partir del lanzamiento del Radiómetro de Microondas de Escaneo Eléctrico (*ESMR-5*, por sus siglas en inglés); este sensor *MI* de orbita polar contó con un canal de frecuencia de operación en 19GHz, y poco despues fué seguido por *ESMR-6* en 1975 operando a 37GHz. Los resultados obtenidos a partir de estos sensores permitieron deducir grandes ventajas de la aplicación de *PMW* en cuanto a la interacción directa de los sensores con los parametros relacionados con la precipitación, lo cual sirvió de antecedente para el lanzamiento del Radiómetro de Microondas de Escaneo Multicanal (*SMMR*, por sus siglas en inglés) en 1978, el cual permitió colectar datos a partir de 5 valores de frecuencias, con doble polarización(6.6, 10.69, 18.0, 21 y 37 GHz).

Desde 1987 el Programa de Satélites Meteorológicos de Defensa (*DMSP*, por sus siglas en inglés) de Estados Unidos a usado el sensor *SSM/I* (Sensor Especial de Microondas, por sus siglas en Inglés), a bordo de sus diversas misiones espaciales. [Levizzani \(2003\)](#) mencionó que en base a [eBellerby et al.\(2000\)](#)ste sensor se hicieron la mayoría de los algoritmos de estimación de tasa de precipitación usando *PMW* en la década de los 90, debido principalmente a que contaba con polarización dual para 3 valores de frecuencia(19.35, 37.0 and 85.5 GHz) y un canal de polarización vertical(22.235GHz).

Dos ejemplo de algoritmos basados en el sensor *SSM/I* fueron presentados por [Wilheit et al. \(1994\)](#) y [Smith et al. \(1998\)](#), en los cuales se muestran diferentes enfoques de estimación de precipitación basados en principios físicos. A pesar de las claras ventajas disponibles a partir del uso de *PMW* para la estimación, estos trabajos reportaron dos desventajas rele-

vantes, la primera es la baja resolución espacial, y en segundo lugar el hecho de que este tipo de sensores están ligados a plataformas de órbita baja, lo cual limita seriamente la resolución temporal del sensor para fines operativos.

Posterior al lanzamiento de plataformas con el sensor *SSM/I*, [Kummerow et al. \(1998\)](#) reportó un aumento en las capacidades de los sensores de microondas pasivas a partir del lanzamiento de la Misión para Mediciones de Precipitaciones Tropicales (*TRMM*, por sus siglas en inglés) en 1997, de la cual cabe destacar que parte de un esfuerzo integrado entre la *NASA*, la Agencia Nacional para el Desarrollo Espacial de Japón (*NASDA*, por sus siglas en inglés) y la Agencia Aeroespacial y de Exploración de Japón (*JAXA*, por sus siglas en inglés).

En este sentido, el incremento de las capacidades de estimación de precipitación usando *PMW* mencionado por [Kummerow et al. \(1998\)](#), está relacionado con el desarrollo del sensor *TMI* (Sensor de Microondas Imager para el *TRMM*) el cual es una versión mejorada del *SSM/I*, que además contaba con un canal adicional para 10.7GHz y mayor resolución espacial para el canal de 85.5GHz . Por otro lado el *TRMM* tenía incorporado sensores en el *VIS* e *IR* y un radar de precipitación (*PR*) de microondas activas que permitieron el desarrollo de algoritmos optimizados para la estimación de precipitación, como es el caso del trabajo presentado por [Iguchi et al. \(2000\)](#).

En este contexto es importante destacar la importancia de comprender los mecanismos de transferencia de radiación asociados a los diversos parámetros relacionados con la precipitación. Por ejemplo, [Quartly et al. \(2002\)](#) concluyeron que las técnicas de estimación de precipitación usando *PMW* de baja frecuencia ($10\text{-}30\text{GHz}$), se basan en las propiedades de absorción/emisión de gotas de lluvia, mientras que en el caso de *PMW* de alta frecuencia la radiación ascendente emitida por la superficie es dispersada por los sistemas de nube, conllevando a la atenuación de la Temperatura de Brillo (*BT*, por sus siglas en inglés) medida por el sensor. De esta forma se resalta la influencia de las características intrínsecas radiativas de la atmósfera para la medición de sus parámetros mediante técnicas de teledetección.

En la figura 2.1 se aprecian las principales aplicaciones de cada canal del *TRMM* en función a la sensibilidad para medir los diferentes parámetros del proceso de precipitación, de la misma no solo se destacan principios físicos basados en la interacción sensor *TRMM*

con la atmósfera, sino además la frecuencia de Microondas con la que opera cada canal y el comportamiento observado en dicha interacción. En este sentido la primera deducción relevante es la diferencia del nivel de emisividad entre superficie de tierra (de 0.7 a 0.9) y océano (inferior a 0.4), la cual tiene un fuerte impacto en las metodologías que deben aplicarse para estimar tasa de precipitación según sea el caso.

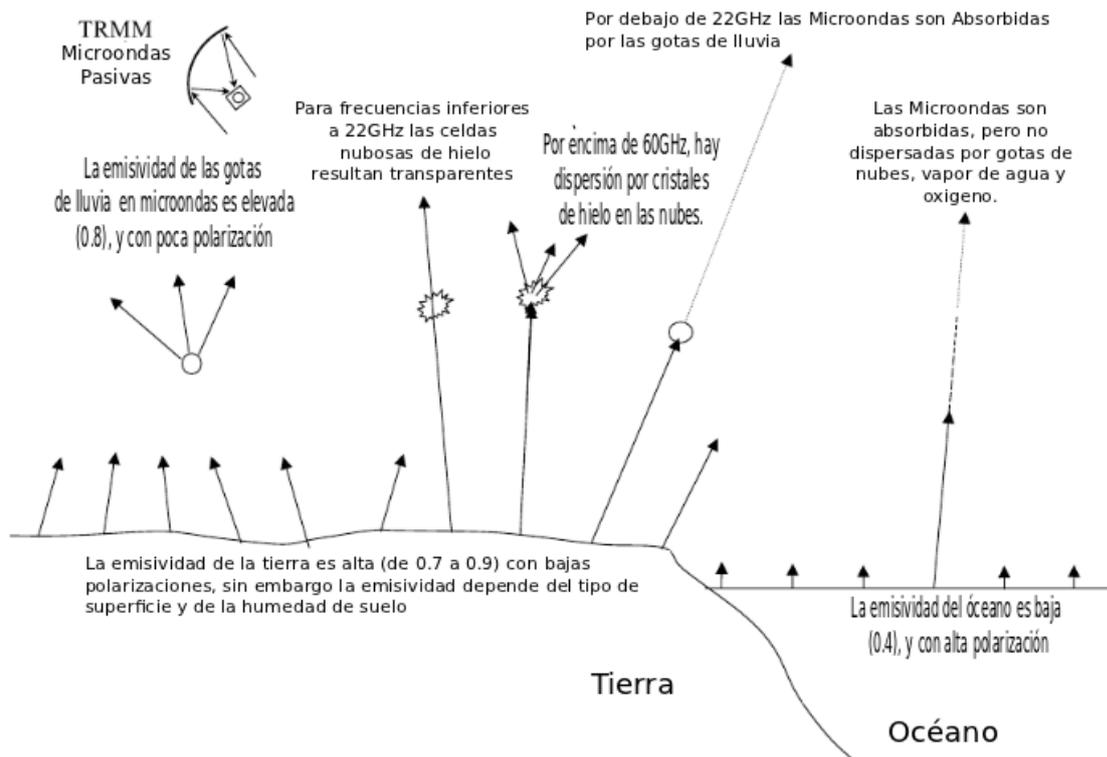


Figura 2.1: Rangos de aplicación y sensibilidad de los canales del *TRMM* para medición de parámetros relacionados a procesos de precipitación (Strangeways, 2007)

En el caso de observación sobre superficie de Tierra, a pesar de la alta emisividad por parte de las gotas de lluvia (0.8), resulta generalmente imposible su identificación debido a la radiación de fondo introducida en la escena como consecuencia de la alta emisividad de la superficie. Por este motivo, los algoritmos de estimación de precipitación sobre Tierra tienden a despreciar los canales de ventana (canales sensibles a la emisión por parte de la superficie), optando por los canales opacos, los cuales son pocos sensibles a la emisión de la superficie y facilitan la identificación de la emisión de las gotas de lluvia.

Adicionalmente se resalta en la figura 2.1 que para frecuencias inferiores a 22GHz los

cristales de hielo son transparentes, lo cual deja abierta la posibilidad de estimar precipitación por debajo de nubes convectivas, ya que existe una alta absorción por parte de las gotas de lluvia para este rango de frecuencias. En contraste, para frecuencias superiores a 60GHz los valores de dispersión por los cristales de hielo en las nubes son importantes, de tal forma que permiten la evaluación de hidrometeoros y estructuras en nubes convectivas.

Por otra parte, para el caso de observación sobre superficie de océano debido a la baja emisividad (inferior a 0.4), los efectos asociados a la emisión por superficie son menos relevantes, ya que la emisividad por parte de las gotas de lluvias es mayor y por lo tanto fácilmente separable. Esto implica que el algoritmo para estimar precipitaciones en píxeles de océano pueden involucrar canales de ventana.

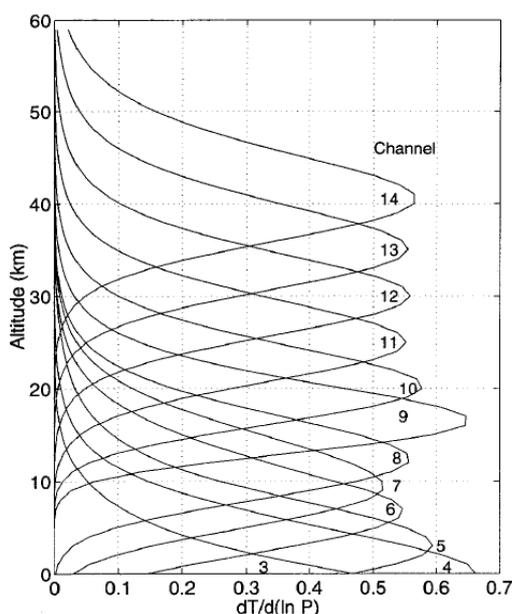
En otro orden de ideas, retomando el tema de los sensores *MI* y *MS*, es preciso destacar que de acuerdo a [Strangeways \(2007\)](#) el primer sensor *MS* fué el denominado *Unidad de Sondeo de Microondas (MSU)*, por sus siglas en inglés) el cual funcionó de forma operativa desde 1979, a bordo del satélite de órbita polar *TIROS-N* bajo la administración de la *NOAA*. Esta unidad fué ampliamente mejorada en resolución y rendimiento derivando en 1998 en el lanzamiento de la Unidad Avanzada de Sondeo de Microondas A y B (*AMSU-A* y *AMSU-B*, por sus siglas en inglés respectivamente) por parte de la *NOAA*.

<i>Canal</i>	<i>Frecuencia Central(s)(GHz)</i>	<i>Ancho de Banda(GHz)</i>
1	23,80	2 x 0.125
2	31,40	2 x 0.080
3	50,30	2 x 0.080
4	52,80	2 x 0.190
5	53,59	2 x 0.168
6	54,40	2 x 0.190
7	54,94	2 x 0.190
8	55,50	2 x 0.155
9	57,29 = f_{LO}	2 x 0.155
10	$f_{LO} \pm 0,217$	2 x 77
11	$f_{LO} \pm 0,3222 \pm 0,048$	4 x 0.035
12	$f_{LO} \pm 0,3222 \pm 0,022$	4 x 0.015
13	$f_{LO} \pm 0,3222 \pm 0,010$	4 x 0.008
14	$f_{LO} \pm 0,3222 \pm 0,0045$	4 x 0.003
15	89,00	2 x 1000

Cuadro 2.1: Frecuencia centrales de los canales AMSU-A (Staelin and Chen, 2000)

Según Staelin and Chen (2000), AMSU observa desde orbita polar con 20 canales distribuidos desde 1.6 hasta 13mm de longitud de onda. El mismo consiste dos unidades completamente independientes: AMSU-A, la cual observa a 15 valores de frecuencia hasta 90GHZ con 50km de resolución espacial aproximadamente, y AMSU-B que observa en 5 valores de frecuencia comprendidas entre 88-191GHZ con resolución espacial de 15km aproximadamente. A objeto de referencia, en el cuadro 2.1 se muestran las frecuencias centrales y el ancho de banda para los canales del AMSU-A, el simbolo f_{LO} representa la frecuencia de oscilación local del canal 7, mientras que en el cuadro 2.2 se muestran para la unidad AMSU-B.

Canal	Frecuencia Central(GHz)	Ancho de Banda(GHz)
1	89	2 x 1
2	150	2 x 1
3	183,31±1	2 x 0.5
4	183,31±3	2 x 1
5	183,31±7	2 x 2

Cuadro 2.2: Frecuencia centrales de los canales *AMSU-B* (Staelin and Chen, 2000)Figura 2.2: Funciones de Pesos de Temperatura para los canales del *AMSU-A* (Chen et al., 2006)

Para una mejor comprensión sobre la información contenida en los cuadros 2.1 y 2.2, en las figuras 2.2 y 2.3 se muestran las funciones de peso para las unidades *AMSU-A* y *AMSU-B* respectivamente¹. Para el caso del sensor *AMSU-A* en la figura 2.2 se aprecia que a partir del canal 5 no es sensible a la radiación emitida por la superficie, lo cual se atribuye a la atenuación ocasionada por el oxígeno.

Por otro lado, para el sensor *AMSU-B* se observa que a partir del canal 3 es despreciable la sensibilidad a la radiación emitida por la superficie, esto se asocia a la absorción por parte

¹Para referencias y explicación sobre las funciones de pesos leer Kuo-Nan (1980)

del vapor de agua. A partir de esta simple observación es notable como las funciones de peso permiten separar por una parte los canales ventana de los opacos, así como también conocer las sensibilidades de cada canal a los diferentes niveles de la atmósfera.

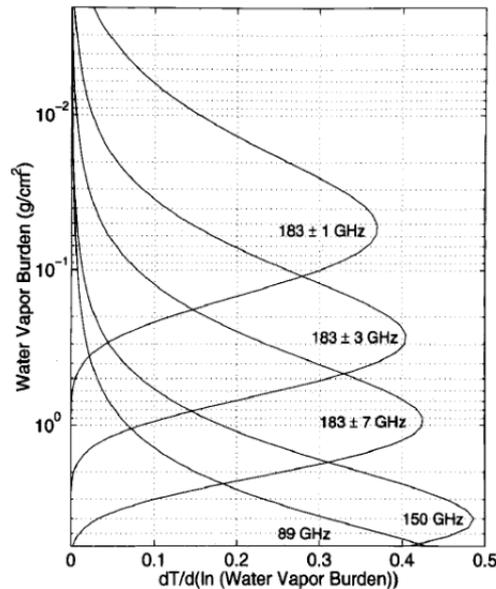


Figura 2.3: Funciones de Pesos del Vapor de Agua para los canales del *AMSU-B* (Chen et al., 2006)

Desde el lanzamiento del *AMSU-A* y *AMSU-B*, han sido numerosas las publicaciones realizadas resaltando las capacidades para estimar precipitación a partir de estos sensores². Un ejemplo reciente es el desarrollo del Algoritmo de estimación de precipitación *WSL-183* desarrollado por Laviola and Levizzani (2011), su nombre hace referencia a que se basa en la fuerte línea de absorción del vapor de agua alrededor de la frecuencia $183,31\text{GHz}$ del *AMSU-B* (Water Strong Line at 183 GHz).

El algoritmo *WSL-183* está conformado por dos módulos, uno para operaciones sobre tierra y otro sobre océano, que incluyen un proceso de separación de precipitaciones estratiformes y convectivas, para luego generar como salida una estimación de precipitación total. Entre los aportes más relevantes proporcionados por el algoritmo *WSL-183* destacan la capacidad para clasificar entre tres tipos de nubes stratiformes y tres de nubes convectivas a partir

²Para mayor información respecto a investigaciones y productos de estimación de precipitación basadas en *AMSU*, referirse a Ferraro et al. (2000), Staelin and Chen (2000), Chen and Staelin (2003) o Chen et al. (2006)

de umbrales establecidos con las temperaturas de brillos proporcionadas por los canales 3,4 y 5³ del *AMSU-B*. Por otro lado, los resultados de validación reportados por [Laviola and Levizzani \(2011\)](#) son alentadores.

2.1.2. Características Generales del Sensor ATMS

El sucesor del *AMSU* está representado por la *Tecnología Avanzada de Sondeo de Microondas (ATMS)*, por sus siglas en inglés), este sensor, de 22 canales comprendidos entre 23.80GHz y 183GHz, fué lanzado en el año 2011 a bordo del Sistema Nacional de Satélites Operacionales Ambientales de Orbita Polar (*NPOESS*) a partir de la alianza conformada por la *NASA* y la *NOAA*. De acuerdo a [Surussavadee and Staelin \(2010\)](#), el *ATMS* difiere del *AMSU* en la mejora que sugiere para estimar precipitación, por la inclusión de tres canales nuevos y la supresión de uno, también al incremento en resolución espacial para los canales de 50 hasta 58GHz, lo que deriva en una mayor amplitud de *swath*⁴ y el uso de muestreo espacial de Nyquist para los canales por debajo de 90GHz que evita discretización de la imagen.

El cuadro 2.3 permite identificar, al compararlo en forma conjunta con los cuadros 2.1 y 2.2, la inclusión de los canales 4, 19 y 21, y también la ausencia del canal 15, el cual estaba presente en el *AMSU-B*. También es importante mencionar entre las ventajas del *ATMS*, es el hecho de que en un solo sensor se han reunido las características esenciales tanto del sensor *AMSU-A* como del *AMSU-B*. Particularmente, [Surussavadee and Staelin \(2010\)](#) consideran que la inclusión de los canales de $183\pm 1,8GHz$ y $183\pm 4,5GHz$ mejoran considerablemente tanto la estimación de vapor de agua como la de precipitación, mientras que el canal de 51.76GHz mejora la captura de información sobre superficie y estimación de precipitaciones estratiformes, así como también la construcción de perfiles de temperatura troposféricos.

³ver cuadro 2.2

⁴Longitud de captura lateral para el barrido del sensor

<i>Canal</i>	<i>Frecuencia Central(GHz)</i>	<i>Resolución Nadir(km)</i>
1	23,80	75
2	31,40	75
3	50,30	33
4	51,76	33
5	52,80	33
6	53,59	33
7	54,40	33
8	54,94	33
9	55,50	33
10	$f_{LO} = 57,29$	33
11	$f_{LO} \pm 0,217$	33
12	$f_{LO} \pm 0,322 \pm 0,048$	33
13	$f_{LO} \pm 0,322 \pm 0,022$	33
14	$f_{LO} \pm 0,322 \pm 0,010$	33
15	$f_{LO} \pm 0,322 \pm 0,045$	33
16	89	33
17	166,31	15
18	$183,31 \pm 7$	15
19	$183,31 \pm 4,5$	15
20	$183,31 \pm 3,0$	15
21	$183,31 \pm 1,8$	15
22	$183,31 \pm 1,0$	15

Cuadro 2.3: Frecuencia centrales de los canales *ATMS* ([Surussavadee and Staelin, 2010](#))

Estas mejoras sugieren una compensación ante la pérdida de resolución espacial en los canales de 23,80, 31,40 y 88, 2GHz.

2.1.3. Estimación de Precipitación mediante Redes Neuronales Artificiales usando datos satelitales

Metodologías basadas en Redes Neuronales Artificiales con sensores IR/VIS

La aplicación de Redes Neuronales en metodologías de estimación de precipitación desde datos satelitales puede atribuirse, entre otras causas, al hecho de que el uso de Redes Neuronales Artificiales ha demostrado alto rendimiento en procesos de regresión y clasificación no-lineal (Freeman and Skapura, 1991, Skapura, 1996). Por este motivo diversos trabajos en el área de estimación de precipitaciones u otras variables atmosféricas como temperatura, han sido reportados hasta el presente fundamentados en diferde calibración o conversión entes implementaciones de Redes Neuronales Artificiales (Hsu et al., 1997, Tsin-tikidis et al., 1997).

En este orden de ideas, Bellerby et al. (2000) usó una red neuronal de perceptrón multicapa con propagación hacia adelante (*FFMLP*, por sus siglas en Inglés) para desarrollar un algoritmo de estimación de tasa de precipitación basado en una combinación de datos ópticos multiespectrales *GOES* con datos del radar de precipitación (*PR*) a bordo de la Misión de Medición de Precipitación Tropical (*TRMM*), obteniendo de este modo una resolución espacial de $0,12^\circ$. Los datos de entrenamiento para la red, fueron preparados mediante la combinación de mediciones *PR* con cuatro bandas *GOES*. De igual modo, fueron incorporados datos sobre textura de nubes y tasas de cambios relacionados con el proceso de estimación, dicha data fue derivada a partir de información estadística sobre múltiples pixeles *GOES* y su análisis con mediciones históricas de precipitación.

Se reportó un coeficiente de correlación aproximado de 0,47 cuando la salida de la red fue contrarrestada con datos de precipitación provenientes de mediciones pluviométricas. La red neuronal usada por Bellerby et al. (2000) se muestra en la figura 2.4, en la cual se aprecia que tanto para la capa de entrada como la de salida, se ha utilizado una función de activación lineal, mientras que para las capas ocultas se ha empleado una función sigmoideal. La configuración de la *FFMLP* usada consta de un conjunto de datos conformado por 45 entradas, 200 neuronas artificiales sigmoideales en la primera capa oculta, la cual está conectada a

una segunda capa oculta de 100 neuronas. Finalmente la tasa de precipitación instantánea es obtenida a partir de una combinación lineal de la segunda capa oculta y la salida.

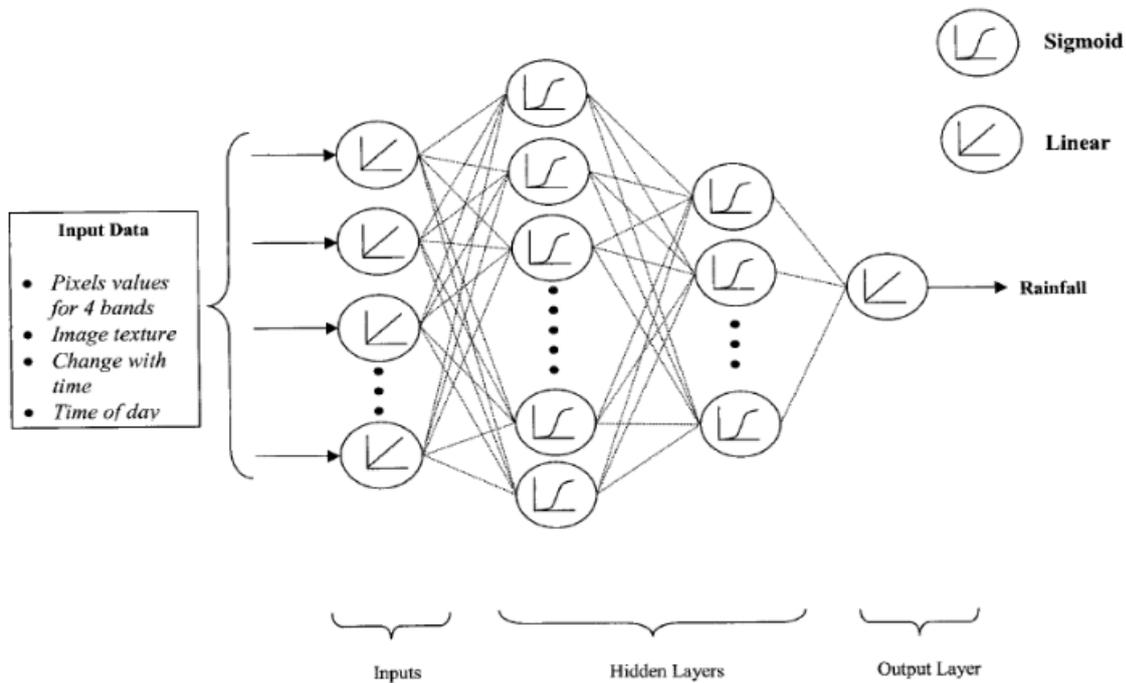


Figura 2.4: Estructura de la Red Neuronal usada por Bellerby et al. (2000)

La FFMLP fué entrenada mediante el algoritmo *Simple Gradient Descent Approach* (GDA), en la cual los errores absolutos definidos para las neuronas de la capa de salida son devueltos a las capas ocultas (*Backpropagation*), y los mismos van corrigiendo los pesos hasta obtener niveles de errores esperados o pre-definidos. Bellerby et al. (2000) comparó resultados de correlación para un conjunto de datos de entrenamiento con una red neuronal de una capa oculta y también hizo lo propio con una red neuronal de dos capas ocultas. Para esto consideró dos áreas que denominó A y B respectivamente, además realizó la evaluación para dos fechas distintas denotadas como fecha 1 y fecha 2 respectivamente. Entre los resultados mostrados se menciona que para el área de estudio A, para la fecha 1 se obtuvo un coeficiente de correlación de 0,58 para la red neuronal de dos capas, delante de 0,51 mostrado por la red de una capa, mientras que para la fecha 2, los coeficientes de correlación fueron 0,59 y 0,54 respectivamente.

Para el área de estudio B, se obtuvo para la red de dos neuronas 0,55 de correlación

en la fecha 1 mientras que la red de una capa mostró una correlación de 0,51. En la fecha 2 [Bellerby et al. \(2000\)](#) reportó un coeficiente de correlación de 0,46 para la red de dos capas y 0,40 para la red de una capa. A partir de estos resultados se concluyó que para entrenar una red de neuronas compleja contra un conjunto de datos con ruido, es esencial comprobar si las funciones no lineales representadas tienen aplicabilidad general, para esto una vez entrenada la red debe contrarrestarse la salida obtenida con un segundo conjunto de datos que no haya sido empleado en el proceso de entrenamiento. Por otro lado, la red de dos capas mostró en todos los casos mayor correlación con la data de entrenamiento que la red de una capa.

Ahora bien, otra contribución importante del trabajo de [Bellerby et al. \(2000\)](#), es el hecho de que compararon la salida de tasa de precipitación dada por el algoritmo desarrollado con resultados proporcionados por el Índice de Precipitación (*GPI*) del *GOES* ([Arkin and Meisner, 1987](#)), así como también con resultados obtenidos mediante estaciones terrestres pluviométricas de la red conformada por *TRMM-Large-Scale Biosphere-Atmosphere (LBA)*, la cual consta de 40 pluviómetros distribuidos para cubrir 13 celdas individuales en superficie con una resolución de $0,04^\circ$. En primer lugar se obtuvo que la FFMLP implementada proporcionó mayor correlación que el *GPI* en todas las evaluaciones mostradas. Por otro lado, los resultados reportados por el algoritmo muestran una correlación de 0,46; 0,54; 0,62 y 0,94 para 30 min, 1 hora, 3 horas y 24 horas de precipitación acumulada respectivamente, registrada por los pluviómetros para la fecha 1. Por otra parte, cabe destacar que resultados similares fueron obtenidos para la fecha 2 de observación.

Estos resultados representan un buen precedente en el desarrollo de sistemas de estimación de precipitación satelital (*SEPS*) para sensores *VIS* e *IR*, además los autores recomiendan que mediante el refinamiento de la red neuronal y haciendo más extensiva la validación de estimación es posible obtener resultados más precisos. Desde el punto de vista de la aplicación de teledetección para el desarrollo de *SEPS*, es importante destacar que [Bellerby et al. \(2000\)](#) seleccionó 4 canales del *GOES* basado en la capacidad de los mismos para identificar diversas características de las nubes.

En este contexto y considerando la relación establecida por [Lovejoy and Austin \(1979\)](#) entre la banda *VIS* y el espesor de nubes y precipitación, este canal fue utilizado como

dato de entrada para la *FFMLP*. De igual modo fue incluido el canal de $3,9\mu m$, ya que de acuerdo a [Scorer \(1987\)](#) contiene información sobre el estado de contenido de agua en la nube, además en conjunción con el canal de $10,7\mu m$ puede contribuir a distinguir nubes no precipitantes ([Vicente, 1996](#)), por otro lado este último canal proporciona información sobre la temperatura de tope de nubes. El cuarto canal utilizado fue el de $6,7\mu m$ debido a que esta fuertemente relacionado con la humedad en niveles altos de la troposfera.

El trabajo de [Bellerby et al. \(2000\)](#), ha servido de referencia para posteriores investigaciones sobre aplicación de redes neuronales en *SEPS*, entre los cuales se destaca el aporte de [Grimes et al. \(2003\)](#) quienes desarrollaron una metodología para la estimación operacional diaria de precipitación usando una red neuronal artificial en conjunto con datos provenientes de imágenes *TIR-Meteosat* (*Infrarojo Térmico, por sus siglas en inglés*) y modelos de análisis de campos basados en *NWP* (*Predicción Numérica del Tiempo, por sus siglas en inglés*) del *ECMWF* (*Centro Europeo para el Pronóstico del Tiempo a medio Alcance*).

Considerando que la mayor parte de algoritmos de *SEPS* que usaban datos geostacionarios *TIR* para la fecha estaban basados en *CCD* (*Duración de Nubes Frías*), como un ejemplo relevante se puede mencionar el *GPI* ([Arkin, 1979](#)), [Grimes et al. \(2003\)](#) llevaron a cabo una primera aproximación en el desarrollo de una red neuronal de arquitectura *49-10-1*, que denominó *TAMANN1*⁵, en la cual se utilizó una función de transferencia sigmooidal logística y cuyos datos de entrada fueron basados en datos *TIR*, *CCD* y parámetros de *NWP*. También es importante mencionar que los pesos de la red fueron ajustados mediante el algoritmo estándar de propagación hacia atrás (*Backpropagation*) y el rendimiento y efectividad de la red fue medida mediante la diferencia media cuadrática (*msd*).

La evaluación y validación de *TAMANN1* fue llevada a cabo mediante la comparación con el algoritmo de estimación de precipitación *TAMCCD*, el cual fue desarrollado por los propios autores basados en datos *TIR* y *CCD*. Se reportó que la calidad de las salidas de *TAMANN1* fueron inferiores a la de *TAMCCD*, lo cual fue atribuido a que el número de datos para el entrenamiento de los 500 pesos asociados a la arquitectura de la red fueron

⁵Esto se entiende como: una red neuronal de tres capas: una de entrada de 49 neuronas, una capa oculta de 10 neuronas y una capa de salida de una neurona

insuficientes. Con la intención de obtener mejores resultados para la red neuronal se aplicó la técnica de *podado de la red* (mejor conocido como *pruning*) reportada por [Weigend et al. \(1991\)](#), la cual consiste en la minimización del número de pesos durante el entrenamiento de la red.

De acuerdo con [Grimes et al. \(2003\)](#), el efecto de aplicar podado sobre la red es suprimir pesos redundantes en los nodos. El proceso de podado fue verificado a partir de la evaluación de las salidas luego de la adición de un conjunto de datos de entrada, a la ANN, compuesto de datos aleatorios. Se reportó que los datos *TIR* así como sus variaciones espaciales pueden ser eliminados si se proporcionan datos *CCD* para diferentes rangos de temperatura. Por otro lado, tomando como referencia el trabajo de [Bellerby et al. \(2000\)](#) se asignó una capa oculta a la arquitectura de la ANN resultando en la forma $30-6-4-1^6$ que denominaron *TAMANN2*, para la cual fueron reportados resultados superiores a los alcanzados por *TAMANN1* y *TAMCCD*. La mejora observada por parte de *TAMANN2* sobre *TAMANN1* se atribuye a los cambios realizados en la arquitectura, a la inclusión de rangos adicionales de *CCD* y a el proceso de podado de la ANN. Ambos métodos tienden a sobre estimar precipitaciones superiores a $10mm$, sin embargo *TAMANN2* lo hace en menor magnitud y por otra parte también tiene mejor ajuste para el caso de precipitaciones de intensidad menor.

Otro aporte interesante a la aplicación de ANN a *SEPS* es el trabajo de [Hong et al. \(2004\)](#), el cual propone una mejora al algoritmo *PERSIANN* presentado por [Hsu et al. \(1997\)](#) mediante la integración de un sistema de clasificación de nubes (*CCS*), este algoritmo fue denominado por [Hong et al. \(2004\)](#) como *PERSIANN-CCS*. [Hsu et al. \(1997\)](#) propuso un sistema de estimación de precipitaciones desde información proveniente de sensores remotos usando redes neuronales artificiales (*PERSIANN*), en el que usó como entrada datos *TIR*, sus derivadas espaciales y también un parámetro que permite clasificar la superficie como tierra, oceano o costa. La ANN es usada para discriminar entre tasas de precipitación debida a diferentes patrones o tipos de nubes mediante el uso de mapas de características auto-organizadas (*SOM*).

⁶Esto es: una red neuronal de 4 capas: una de entrada de 30 neuronas, una primera capa oculta de 6 neuronas, una segunda capa oculta de 4 neuronas y una capa de salida de una neurona

De acuerdo con [Hong et al. \(2004\)](#) una de las limitaciones de la arquitectura *PERSIANN* es que estima precipitación basada en información de píxeles locales, en este sentido para superar esta limitación el algoritmo *PERSIANN-CCS* extrae características locales y regionales de nubes usando el canal de $10,7\mu m$ en cuatro pasos: En primer lugar, separa imágenes de nubes en distintos parches de nubes; luego extrae características de las nubes, incluyendo temperatura, geometría y textura; organiza los núcleos de parches de nubes en subgrupos; finalmente calibra la temperatura de tope de nube y las relaciones de precipitación, temperatura de brillo y tasa de precipitación, para los grupos de nubes clasificados usando datos de precipitaciones de radar corregidos con estaciones pluviométricas. Los autores concluyeron que *PERSIANN-CCS* es una herramienta disponible para cuantificar distribuciones de precipitación sobre tierra, y que en futuros trabajos presentarían soluciones para estimaciones sobre océanos.

Metodologías basadas en Redes Neuronales Artificiales con sensores *MW*

La ventaja del uso de información de *MW* en lugar de *IR* para la estimación de precipitación reside en el hecho de que las gotas de lluvia afectan directamente la emisión de *MW* de la Tierra hacia la atmósfera. Por este motivo [Tapiador et al. \(2004\)](#) plantean el entrenamiento de una red neuronal artificial para extraer la relación entre siete canales del sensor de *MW* *SSM/I* y la precipitación, evaluando el rendimiento de la red en base a la comparación con datos de referencia ⁷.

Desde el lanzamiento de la Unidad Avanzada de Sondeo de Microondas (*AMSU*, por sus siglas en inglés) a bordo de los satélites de la Administración Nacional para la Atmósfera y el Océano (*NOAA*, por sus siglas en inglés), se han publicado diversas metodologías para estimar variables atmosféricas a partir de este tipo de sensor. Una gran parte de dichas publicaciones han estado enfocadas hacia el uso de *ANNs* como una solución práctica a problemas relacionados con la alta complejidad asociada a la determinación de variables atmosféricas. Por ejemplo, [Staelin and Chen \(2000\)](#) usaron *ANNs* para estimar precipitación,

⁷Es importante destacar que el objetivo de este trabajo está enfocado a estimar precipitaciones desde sensores basados en *MW* pasivas, en todo caso, para información relacionada al uso de radares activos para la estimación de precipitación es recomendable consultar a [Tomassetti et al. \(2008\)](#)

en el orden de capturar la no-linealidad que relaciona a las radianzas del AMSU con la tasa de precipitación.

Esta ANN opera en la Temperatura de Brillo (*BT*, por sus siglas en inglés) absoluta cercana a las frecuencias 183 ± 1 y 183 ± 7 GHz, el coseno del ángulo de escaneo y las perturbaciones de la Temperatura de Brillo del AMSU-A⁸ introducidas por la presencia de hidrometeoros cerca a las frecuencias 52,8; 53,6 y 54,4 GHz. En primer lugar hicieron uso de datos a 183 ± 7 GHz para demarcar las zonas con probabilidad de precipitación, expresando que en zonas con *BTs* alrededor de 260K generalmente son zonas de lluvia, y de este modo generar mascararas que filtren estas zonas de interés; luego corrigieron el ángulo de escaneo de los canales 4 al 7 usando dos ANNs⁹ con el objetivo de estimar *BT* en dirección nadir.

Ahora bien, el siguiente paso consistió en estimar la tasa de precipitación, para esto diseñaron una *FFMLP* compuesta por 7 variables de entrada, una capa oculta de 3 neuronas, una de 2 neuronas y la capa de salida de una neurona, que opera en valores de radianzas entre 183 ± 1 y 183 ± 3 GHz, las perturbaciones negativas cercanas a 54GHz del AMSU-A en los canales 4 al 6, el coseno del ángulo de escaneo y una constante asociada a unidades. En todos los nodos se usó la tangente hiperbólica de la suma de sus siete o tres entradas pesadas, de acuerdo sea la capa. [Staelin and Chen \(2000\)](#) aseguran que el algoritmo implementado por la ANN esta descrito por el conjuntos de pesos encontrados. Entre los resultados reportados resalta el hecho de que en la primera capa oculta, conformada por tres neuronas, la primera neurona demostró alta sensibilidad para estimar precipitación en la troposfera alta cuando son usados los canales 3 del AMSU-B, 5 y 6 del AMSU-A y por otro lado se tiene un peso negativo para el canal 4 AMSU-A, el cual es sensible a precipitaciones a bajas altitudes en la troposfera. Por otro lado se reportó que la segunda y tercera neuronas mostraron alta sensibilidad para precipitaciones originadas a alturas media y bajas respectivamente.

Entre los resultados mas destacables de [Staelin and Chen \(2000\)](#), resalta el hecho de que pese a la diferencia en resolución espacial entre AMSU-A y AMSU-B, cuando se usa

⁸El sensor AMSU consiste de dos unidades de medición independientes: AMSU-A, la cual observa en 15 frecuencias diferentes hasta 90GHz con 50km de resolución espacial y AMSU-B que observa en 5 frecuencias entre 88-191GHz con 15km de resolución espacial.

⁹Una red para correcciones sobre superficie en tierra y la otra para correcciones sobre el océano

simultáneamente frecuencias cercanas a 54GHz y 183GHz, es posible obtener estimaciones de precipitación altamente consistentes con respecto al Sistema de Radares Meteorológicos de alta resolución Doppler de Nueva Generación (*NEXRAD*, por sus siglas en inglés) administrado por el Servicio Nacional Meteorológico (*NWS*, por sus siglas en inglés) adscrita a la *NOAA* de los Estados Unidos de Norte América.

[Blackwell and Chen \(2005\)](#) plantearon que las técnicas de estimación basadas en *ANNs* se han vuelto muy común en teledetección atmosférica de alta resolución debido a la simplicidad, flexibilidad y habilidad para representar precisamente relaciones estadísticas multi-dimensionales complejas. En este sentido reportaron el desarrollo de un algoritmo de estimación de precipitaciones basados en el trabajo de [Staelin and Chen \(2000\)](#) usando *ANNs*. Si bien es cierto este trabajo incluye igualmente datos del sensor *AMSU*, la novedad presentada por [Blackwell and Chen \(2005\)](#) incluye la adición de datos de entradas correspondientes a componentes principales (*PC*, por sus siglas en inglés) de perfiles de vapor de agua. En el algoritmo se ha estimado al canal 183 ± 7 GHz como el primordial para la detección de precipitación, debido a su alta sensibilidad a este párametro y a su vez por la baja sensibilidad a la superficie.

La sensibilidad a la superficie asociada al canal 183 ± 7 GHz puede aumentar en regiones secas y frías, en tales casos [Blackwell and Chen \(2005\)](#) recomiendan usar un canal mas opaco como el 183 ± 3 GHz, y si aun este sensor es sensitivo la opción sería el canal 183 ± 1 GHz. También se reportaron resultados sobre la comparación de las salidas del algoritmo con el sistema *NEXRAD*, de lo cual se concluyó una alta compatibilidad para ambas técnicas en cuanto a la localización de sistemas de precipitaciones severas. El algoritmo fue evaluado, incluso durante el desarrollo del huracán de categoría 5 Isabel (2003), mostrando claramente el ojo del huracán conforme a los resultados obtenidos por *NEXRAD*.

Los trabajos presentados por [Staelin and Chen \(2000\)](#) y [Blackwell and Chen \(2005\)](#) usaron para el proceso de entrenamiento datos proporcionados por el *NEXRAD*, para lo cual [Surussavadee and Staelin \(2008a\)](#) considera que se desprecian precipitaciones no-convectivas debido a que los datos provenientes de radares involucran gran dificultad para comprender la distribución de hidrometeoros en la atmósfera. En este orden de ideas [Surussavadee and](#)

[Staelin \(2008a\)](#) presentaron un algoritmo basado en *ANNs* para estimar precipitación, también basado en el sensor *AMSU* y entrenado a partir de datos provenientes del modelo de predicción numérica del tiempo (*NWP*, por sus siglas en inglés) *MM5* (Modelo Meso-escalar 5, por sus siglas en inglés) para superficies reales en todos los ángulos de escaneo. Las principales novedades presentadas por [Surussavadee and Staelin \(2008a\)](#) incluyen el uso de *PC* para reducir los efectos de superficie y correcciones de las *BT* nadir basadas en *ANNs*, entre otras.

Partiendo de que el comportamiento de la absorción por parte de la precipitación está relacionado con el perfil de temperatura en la atmósfera, [Surussavadee and Staelin \(2008a\)](#) han desarrollado dos algoritmos para la estimación de la precipitación, uno para el uso en superficie y otro para ser usado sobre océano, ambos basados en píxeles de longitud y latitud. Sobre océano, la absorción debida a precipitación puede distinguirse de la absorbida por el vapor de agua usando los canales cercanos a 22.2 y 183GHz que son los canales de resonancia del vapor de agua. Los autores reportan que todas las *ANNs* usadas tienen de estructura dos capas ocultas, la primera de 10 neuronas y la segunda con 5 neuronas y ambas usan como función de activación la tangente hiperbólica.

El algoritmo de entrenamiento usado fue el de Levenberg-Marquardt y los bias y pesos fueron inicializados mediante el método Nguyen-Widrow. Luego de evaluar la correlación entre cada canal de entrada a la *ANN* con la salida, se obtuvo que los canales más determinantes para estimar la tasa de precipitación en superficie son, para *AMSU-A* desde el 4 hasta el 8, mientras que para *AMSU-B* es el canal 183 ± 7 GHz. Sin embargo los autores reportan que la importancia relativa de cada canal varía con el tipo de precipitación observado, resultados sobre la evaluación y validación de este algoritmo fueron reportados por [Surussavadee and Staelin \(2008b\)](#).

Entre los resultados más relevantes reportados se tiene el hecho de que la precisión de las estimaciones de precipitación mostraron independencia del ángulo zenith y por consiguiente los ángulos de escaneo fueron presentados como el promedio; también que la estimación sobre tierra es ligeramente menos precisa que sobre océano, particularmente a tasas de precipitación baja donde la superficie permanece visible; también se evidencia subestima-

ción de tasas de precipitación en eventos severos, lo cual fue relacionado con la atenuación generada por los cristales de hielo asociados a grandes desarrollos convectivos. Cabe destacar que se evaluó la estimación de tasa de precipitación en superficie para 106 eventos de tomenta a partir de un diagrama de dispersión entre la verdad de campo del MM5 y la salida AMSU/MM5 obteniendo un coeficiente de correlación de 0,70.

De esta breve referencia de trabajos se distinguen algunos rasgos de gran importancia para la estimación de precipitación en base a *ANNs*, tal como la necesidad de datos confiables y representativos para el proceso de entrenamiento, establecimiento de criterios para la selección de variables de entrada, topología de la *ANN* así como algoritmos de entrenamiento. El enfoque de estos trabajos en el sensor *AMSU* es un buen antecedente para la presente investigación por la gran similitud que tiene con el sensor *ATMS*, el cual es fuente de datos para el presente estudio.

En referencia a *ATMS*, [Leslie et al. \(2008\)](#) desarrollaron y presentaron un estudio de validación para un algoritmo de estimación de precipitación basado en un enfoque estadístico que involucró un extenso trabajo de procesamiento de imágenes y la construcción de un estimador basado en una *FFMLP*. El entrenamiento de la *FFMLP* se realizó con datos de radianza *ATMS* simulados, para esto se usó el Modelo de Resolución de Nubes (*CRM*, por sus siglas en inglés) del *MM5* para generar cantidades termodinámicas de la atmósfera, tales como perfil de humedad y de hidrometeoros, las mismas sirven de entrada al Algoritmo de Transferencia de Radiación (*RTA*, por sus siglas en inglés) cuya salida son las radianzas *ATMS* simuladas.

El algoritmo basado en *ANN* fue entrenado para estimar tasa de precipitación a partir del *MM5*. En el proceso de entrenamiento fueron usados los 22 canales del *ATMS* y, en todos los casos se usó una red que evaluó la estimación de tasa de precipitación en superficie para 106 eventos de tomenta a partir de un diagrama de dispersión entre la verdad de campo del *MM5* y la salida AMSU/MM5 obteniendo un coeficiente de correlación de 0,70 usando una capa oculta con 10 neuronas y como función de activación la tangente hiperbólica. Por otra parte los pesos y bias fueron inicializados siguiendo el método de Nguyen-Widrow y el algoritmo de entrenamiento fue el método de Levenberg-Maquardt para un máximo de 100

épocas en todos los casos. Los resultados obtenidos se muestran en la Fig. 2.5, en los mismo se aprecia en general una buena tendencia.

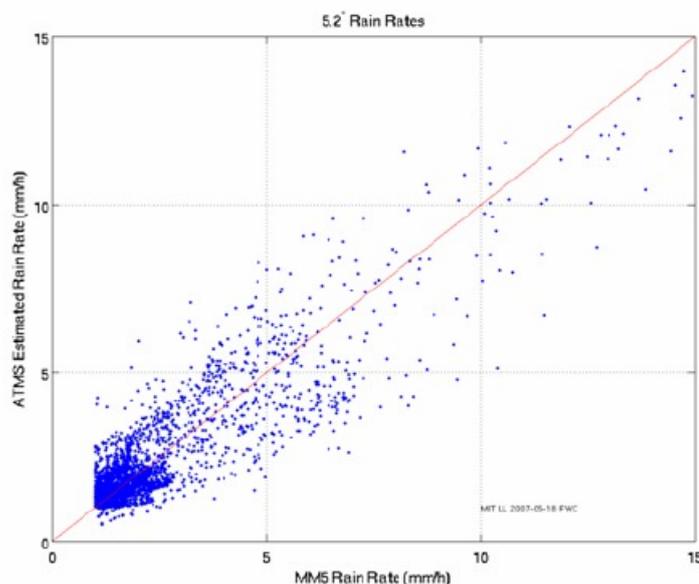


Figura 2.5: Diagrama de Dispersión entre la salida de la ANN y el MM5 a una resolución de 5, 2° por Leslie et al. (2008)

Entre las principales conclusiones reportadas por Leslie et al. (2008) resaltan el hecho de haber fundamentado su trabajo en la metodología de Staelin and Chen (2000) y la buena precisión alcanzada, la cual fué de 0.83 mm/h estimado a partir del error cuadrático medio (RMS, por sus siglas en inglés).

Metodologías basadas en Redes Neuronales Artificiales con fusión de sensores IR/VIS y MW

El objetivo de los métodos de fusión es combinar la capacidad de muestreo temporal de plataformas geostacionarias, con las mediciones directas sobre precipitación proporcionadas por sensores de microondas pasivas (LEO, por sus siglas en inglés) portados en satélites de órbita baja. A pesar de los problemas relacionados con este método, asociados principalmente a las diferencias en resoluciones temporales y espaciales, usualmente muestra mejor rendimiento y precisión que el uso de datos IR/VIS o MW por separado (Tapiador et al.,

2004).

Marzano et al. (2005) presentaron un algoritmo que integra datos *IR* con datos *MW* para estimar precipitación, el cual denominaron *MICRA* (The Microwave Infrared Combined Rainfall Algorithm) y está basado en la integración estadística de data complementaria *GEO-IR* (*GEO*, *Orbita Geosincrónica sobre la Tierra*) y *LEO-MW* (*LEO*, *Orbita Baja sobre la Tierra*). Además presentaron un estudio comparativo de *MICRA* con un algoritmo basado en *ANN* denominado *NEREMIS* (*Algoritmo de estimación de precipitación usando sensores MW e IR*).

NEREMIS es una *FFMLP* de cuatro capas con arquitectura 27-10-8-1 que usa como fuente de entrada datos de Temperatura de Brillo (*BT*) provenientes de sensores *GEO-IR*, así como también información sobre la latitud y longitud del escenario de observación. Los autores no reportaron detalles sobre el algoritmo de entrenamiento utilizado, sin embargo, mencionaron datos interesantes sobre la técnica de calibración. La calibración del *NEREMIS* se realizó utilizando en la salida estimaciones de tasa de precipitación derivadas de mediciones *LEO-MW*, esto con la finalidad de establecer comparaciones para *MICRA*, por consiguiente son requeridos datos *IR* y *MW* co-registrados en espacio y tiempo. Los resultados reportados por Marzano et al. (2005) mostraron una tendencia sistemática en la estimación de campos de precipitación, donde fue resaltada gran concordancia entre las estimaciones derivadas de ambos métodos, lo cual una vez más muestra resultados que hacen factible el entrenamiento de *ANN* para *SEPS*.

Rivolta et al. (2006) desarrollaron una técnica de pronóstico de corto término (*now-casting*) usando *ANN* y combinando datos *IR* y *MW*. En base a conclusiones reportadas por diversos autores sobre la mejora que ofrece la disponibilidad de datos radiométricos de *MW* para establecer la correlación física de la precipitación que tiene lugar bajo nubes observadas por el sensor (Kummerow and Giglio, 1995, Levizzani et al., 1996, Marzano et al., 2005, Tapiador et al., 2004) y considerando el hecho de que facilita la detección de estructuras de nubes y que las mediciones de *BT* han mostrado ser altamente sensitivas a la presencia de hidrometeoros líquidos y de hielo. Rivolta et al. (2006) entrenaron una *FFMLP* bajo distintas configuraciones usando como entrada imágenes *IR* tras un pre-proceso de calibración

utilizando datos del *SSM/I* (Special Sensor Microwave Imager, por sus siglas en Inglés).

[Rivolta et al. \(2006\)](#) se basaron en el uso de una estructura *FFMLP*, similar a la estructura usada por [Bellerby et al. \(2000\)](#) mostrada en la figura 2.4, y para minimizar la función de errores durante el proceso de entrenamiento se utilizó el algoritmo de gradiente descendiente con propagación hacia atrás. Con la finalidad de introducir en la *ANN* información referente a las diferentes fases asociadas a los procesos meteorológicos, el par de datos de entrada y salida del proceso de entrenamiento han sido seleccionados con distancias temporales de 6 horas para un total de 4 series de imágenes METEOSAT. Entre los aportes observados en el trabajo de [Rivolta et al. \(2006\)](#) resalta el hecho de haber usado diferentes patrones de entrenamiento para las distintas configuraciones de red diseñadas.

Las dos estructuras destacadas fueron en primer lugar una *FFMLP* con arquitectura 27-75-50-1 la cual ha sido basada en un nivel de vecindad 1 (Esto es 9 píxeles) considerando la imagen anterior relativa ($N-1$), la actual (N) y la próxima ($N+1$), la cual corresponde a la salida del pronóstico. En segundo lugar se tiene una estructura del tipo 363-15-8-1 que se basa en un nivel de vecindad 5 (121 píxeles) para las tres últimas imágenes de forma similar a la primera configuración. Los resultados reportados por [Rivolta et al. \(2006\)](#), confirmaron en términos del Error Cuadrático Medio (*RMS*, por sus siglas en inglés), que la configuración con mejor rendimiento es la *FFMLP* de estructura 363-15-8-1

2.2. Principios Físicos de la Estimación de Precipitaciones mediante Microondas Pasivas

La comprensión de los mecanismos de absorción, emisión y dispersión a partir de la interacción de ondas electromagnéticas con los gases atmosféricos, permite obtener información sobre una gran cantidad de variables asociadas a la microfísica de nubes, que es donde tiene su origen el proceso de precipitación. En este sentido, la teledetección atmosférica mediante el uso de *PMW* se fundamenta en la solución de la ecuación de transferencia radiativa (*ETR*) para modelar fenómenos de la atmósfera. Dicha ecuación puede ser deducida a partir

del modelo de capas paralelas, en donde se considera que el flujo neto de radiación en una placa z de la atmósfera esta dado a partir de la estimación del flujo ascendente y descendente en relación a la capa.

Con la finalidad de entender un mecanismo que permita abordar el problema de estimar tasas de precipitación usando *PMW*, en esta sección se presentan tópicos relacionados a los principios de índole físico relacionados con la teledetección de la atmósfera, de modo tal que en conjunto con la deducción de la *ETR*, faciliten el diseño de una metodología para elaborar el algoritmo de estimación de tasa de precipitación objeto de esta investigación.

2.2.1. Estructura de la Atmósfera Terrestre

La atmósfera generalmente es dividida de acuerdo a su perfil vertical de temperatura según como se muestra en la figura 2.6. En la misma el eje vertical izquierdo es la presión (Escrita en inglés en la figura como *pressure*), a la derecha la altura geométrica respecto al nivel del mar (*geometric height*) y en el eje horizontal la temperatura expresada en escala Kelvin y Fahrenheit.

En la figura 2.6, se destacan las 4 capas térmicas de la atmósfera: La Tropósfera (*Troposphere*), Estratósfera (*Troposphere*), Mesósfera (*Mesosphere*) y la Termósfera (*Thermosphere*), así como también 3 capas que sirven de interfaz entre cada par de capas: Entre la Tropósfera y la Estratósfera se encuentra la Tropopausa, entre la Estratósfera y Mesósfera se tiene la Estratopausa y finalmente entre la Mesósfera y la Termósfera la Mesopausa. Para una mejor descripción sobre las características físico-químicas tanto de las capas componentes como de las capas interfaces el lector puede acudir a [Seinfeld and Pandis \(2006\)](#) o [Muñoz and Falcón \(2008\)](#)

En la figura 2.6, a modo de referencia se puede observa la relación de altura entre una nube cumulonimbus y la atmósfera, de esta comparación cabe destacar que los fenómenos meteorológicos que inciden de manera directa en el proceso de precipitación estan limitados a la Tropósfera y la Tropopausa, razón por la cual en este trabajo, estas dos capas serán descritas detalladamente en virtud de tener una mayor comprensión del proceso de precipitación

y su estimación a partir del sensor ATMS.

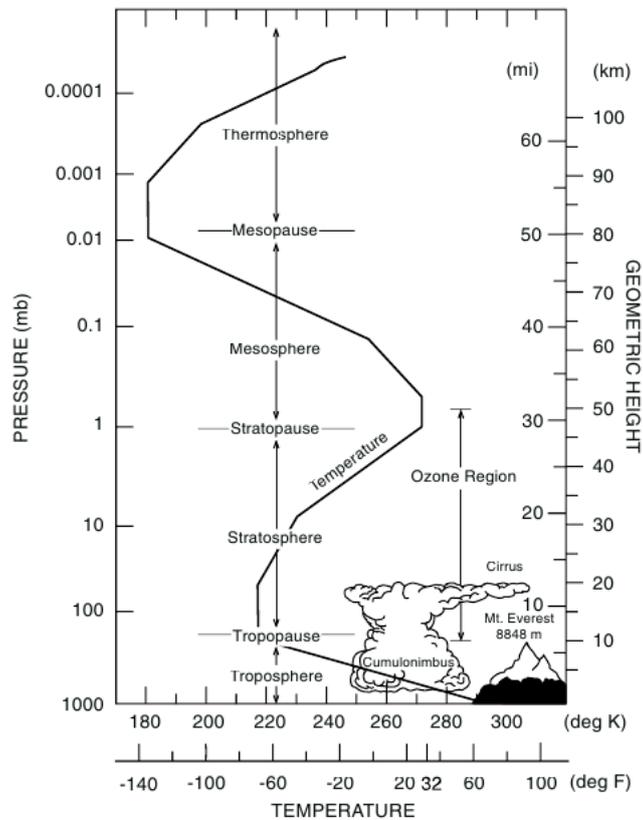


Figura 2.6: Estructura térmica de la atmósfera con sus respectivas capas (Brasseur and Solomon, 2005).

La Tropopausa, es la primera capa de la atmósfera y en esta capa se concentra cerca del 80 % de la masa total de la atmósfera (Muñoz and Falcón, 2008). La importancia desde el punto de vista meteorológico que tiene la Tropósfera radica en el hecho de que en ella tienen lugar los fenómenos microfísicos y químicos que dan lugar a la formación y de desarrollo nubes a partir de las cuales tiene su origen la precipitación. La altura de la Tropósfera alcanza en promedio 16km en el ecuador y va disminuyendo suavemente conforme a la latitud, hasta tener una altura de 8km en los polos. Esta diferencia de altura entre los polos y el ecuador se debe al transporte radiativo que tiene lugar en la atmósfera, el cual está regido por los mecanismos relacionados con la incidencia de radiación solar a la Tierra.

Considerando la naturaleza del transporte de energía que tiene lugar en la Tierra y su atmósfera, la altura de la Tropopausa varía de acuerdo a la latitud, así como también con

las estaciones climáticas y las variaciones propias que tienen lugar en escala temporal meteorológica. La Temperatura en la Tropósfera disminuye en forma lineal en relación a la altura, de acuerdo con el gradiente adiabático, el cual la Organización Mundial de Meteorología establece como $-6K/km$. Y esta disminución tiene lugar hasta la altura máxima de la Tropósfera, el cual coincide con el inicio de la Tropopausa, la cual es una capa de interfaz entre la Tropósfera y la Estratósfera que se caracteriza por un gradiente de $-2K/km$, hasta alcanzar la Estratósfera.

Es importante destacar con fines didácticos que la zona de intercepción entre la Tropopausa y la Estratosfera es claramente definida a partir de diferentes metodologías, como por ejemplo mediante el uso de radiosondas. Esta interfaz define la denominada capa de inversión térmica asociada en diferentes trabajos con el calentamiento global que parten del principio de concentraciones de gases de efecto de invernadero.

2.2.2. Composición de la Atmósfera Terrestre

Los constituyentes atmosféricos, en general, son característicos para una atmósfera y definen de modo directo sus propiedades químicas y físicas. La composición de la atmósfera de la tierra, se clasifica según la concentración en gases fijos y variables (Muñoz and Falcón, 2008), los mismos se muestran en los cuadros 2.4 y 2.5 respectivamente.

<i>Gases Fijos</i>	<i>Porcentaje</i>	<i>Partes por Millón de Volumen</i>
Nitrógeno(N_2)	78.08	780000
Oxígeno(O_2)	20.95	209500
Argón(Ar)	0.93	9300
Neón(Ne)	0.0015	15
Helio(He)	0.0005	5
Kriptón(Kr)	0.0001	1
Xenón(Xe)	0.000005	0.5

Cuadro 2.4: Gases de concentración fija que componen la Atmósfera (Andrews, 2010)

En el cuadro 2.4 se muestran los gases de concentración aproximadamente homogénea en la atmósfera, en el mismo se detalla información sobre el porcentaje de concentración y las partes por millón de volumen que ocupan. Por su parte, en el cuadro 2.5 se observan los gases de concentración variables, donde a pesar de los bajos porcentajes en que se encuentran, estos cuatro gases¹⁰ juegan un papel muy determinante en áreas como meteorología, polución de la atmósfera y en el balance de radiación general de la atmósfera. Por ejemplo, en el caso de la precipitación, las concentraciones a escala local de vapor de agua, está fuertemente relacionada con la estabilidad de la troposfera y la formación de nubes.

<i>Gases Variables</i>	<i>Porcentaje Mínimo</i>	<i>Partes por Millón de Volumen</i>
Vapor de Agua(H_2O)	0.00001- 4.0	0,1-40,000
Dióxido de Carbono(CO_2)	0.0360	≥ 360
Metano(CH_4)	0.00017	$\geq 1,7$
Ozono(O_3)	0.000003- 0.001	0.03-10

Cuadro 2.5: Gases de concentración variable en la Atmósfera (Andrews, 2010)

2.2.3. La función de Planck y la Teoría de Radiación de Cuerpos Negros

De acuerdo con Andrews (2010), la Ley para la Cavidad Radiante de Planck establece que la densidad de energía espectral $u_\nu(T)$ de un cuerpo negro que se encuentra a una temperatura absoluta T , esta dada por la ecuación 2.1:

$$u_\nu(T) = \frac{8\pi h\nu^3}{c^3(\exp[h\nu/k_B T] - 1)} \quad (2.1)$$

Donde k_B es la constante de Boltzmann, ν es la frecuencia a la que es emitido un fotón por el cuerpo negro, h es la constante de Planck, y c es la rapidez de la luz en el vacío. Ahora bien, dado que los fotones emitidos por el cuerpo negro con la energía descrita por

¹⁰Sin perdida de generalidad se han incluido en en el cuadro 2.5 los gases variables con mayor variabilidad en la atmósfera y con mayores valores de concentración media. Sin embargo, cabe destacar que se han ignorado una gran cantidad de gases que juegan un papel importante en los mecanismos de balance de energía por no estar directamente relacionados con el objeto de esta investigación

la ecuación 2.1 se mueven isotrópicamente, se tiene que la densidad de energía asociada a un grupo de fotones que se mueven a través de un diferencial de ángulo sólido $\Delta\Omega$ es $u_\nu * \Delta\Omega/4\pi$. A partir de la consideración del flujo de energía por unidad de tiempo, por unidad de área, transportada por este grupo de fotones a una rapidez c se puede demostrar que la potencia por unidad de área, por unidad de ángulo sólido, por unidad de intervalo de frecuencia para un cuerpo negro está dada por la ecuación 2.2, la cual se conoce como la función de Planck.

$$B_\nu(T) = \frac{2h\nu^3}{c^2(\exp[h\nu/k_B T] - 1)} \quad (2.2)$$

Para una mejor comprensión sobre la función de Planck en el contexto del campo de la teledetección es conveniente expresar la densidad de energía del grupo de fotones como una función de la longitud de onda λ , en lugar de la frecuencia ν . La ecuación 2.3 refleja a la densidad de energía como función de la Temperatura T del cuerpo negro (Andrews, 2010).

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5(\exp[hc/\lambda k_B T] - 1)} \quad (2.3)$$

La importancia de la ecuación 2.3 es que permite conocer la radianza espectral de un cuerpo negro con temperatura absoluta T , para mayor comprensión en la figura 2.7 se aprecia la radianza espectral para un cuerpo negro a temperatura de $6000K$ y para otro de $288K$. Debido a la amplitud de magnitudes involucradas en la graficación de la ecuación 2.3, la figura 2.7 ha sido elaborada en escala logarítmica para ambos ejes. En esta figura se hace una comparación entre un cuerpo negro a la temperatura de la fotosfera del sol y el de un cuerpo negro con temperatura típica a la temperatura media de la superficie terrestre, y se observa como la radianza espectral del cuerpo más frío, es completamente abarcado por el más caliente. También se aprecia que las longitudes de onda de emisión del cuerpo negro análogo a la tierra está en el orden del infrarrojo y microondas. De igual modo la diferencia entre los máximos en radianza espectral están descritos por la ley de desplazamiento de Wien (Kuo-Nan, 1980).

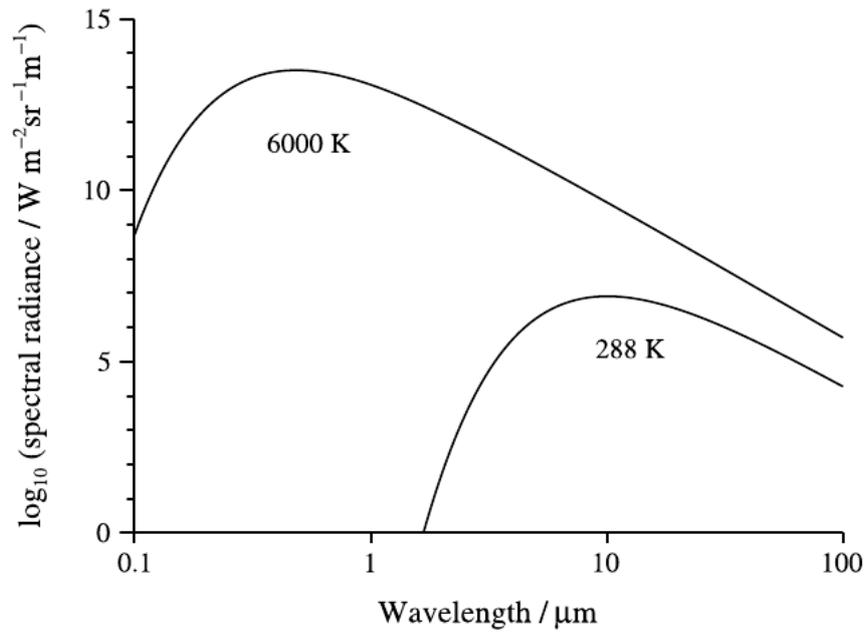


Figura 2.7: Logaritmo de la Radianza Espectral $B_{\lambda}(T)$ graficado respecto al logaritmo de la longitud de onda λ asociada a un cuerpo de $T = 6000K$, temperatura típica de la fotosfera solar, y también graficado contra un cuerpo de $T = 288K$ que es la temperatura media de la superficie terrestre (Andrews, 2010)

En la práctica, un cuerpo negro es una idealización, el cual en esencia es capaz de absorber todo el espectro de energía electromagnética que recibe, y el mismo debe cumplir con la Ley de Kirchoff para el equilibrio térmico (Andrews, 2010), de tal modo que por definición un cuerpo negro presenta un coeficiente de emisividad espectral $\varepsilon_{\nu} = \alpha_{\nu} = 1$, donde α_{ν} es el coeficiente de absorbancia. Ahora bien, el coeficiente de emisividad espectral de un cuerpo no-negro es inferior al del cuerpo negro idealizado por Planck y el mismo está dado como la razón entre la radianza espectral del cuerpo y la radianza del cuerpo negro, de tal modo que para dicho cuerpo se cumple que $\varepsilon_{\nu} \leq 1$.

A partir de la función de Planck se deriva un principio físico fundamental para la teledetección atmosférica, y esto es que a partir de la densidad de energía emitida por un objeto es completamente factible determinar la Temperatura.

2.2.4. Ley de Stefan-Boltzmann

La función de Planck, expresada en la ecuación 2.3, permite conocer la potencia por unidad de área, por unidad de ángulo sólido, por unidad de longitud de onda B_λ asociada a una señal electromagnética de longitud de onda λ y a una temperatura T , lo cual es útil cuando se quiere determinar la temperatura de un cuerpo y no existe la posibilidad de medir totalmente la radianza espectral del mismo. Bajo la condición de que exista la posibilidad de conocer totalmente la radianza espectral del cuerpo emisor, es posible integrar la función de Planck para todos los valores de λ correspondientes al espectro de emisión, obteniendo:

$$B(T) = \int_0^\infty B_\lambda(T) d\lambda = \int_0^\infty \frac{2hc^2 d\lambda}{\lambda^5 \exp[hc/\lambda k_B T] - 1} \quad (2.4)$$

Kuo-Nan (1980) demuestra que, al resolver esta integral y considerar la geometría del ángulo sólido a partir del cuerpo emisor, la densidad de flujo de energía asociada al cuerpo emisor está dado por:

$$F = \sigma T^4 \quad (2.5)$$

La ecuación (2.5) es conocida como la Ley de Stefan-Boltzmann, donde σ es la constante de Stefan-Boltzmann, cuyo valor y dimensionalidad pueden ser verificados a partir de Kuo-Nan (1980). A diferencia de la función de Planck, la ley de Stefan-Boltzmann permite determinar la Temperatura T de un cuerpo negro a partir de la densidad de flujo de energía F , siempre que se conozca la radianza espectral total del mismo.

2.2.5. Temperatura de Brillo (BT)

Dada la densidad de flujo de energía F de un cuerpo negro, se puede determinar la temperatura T del mismo a partir de la Ley de Stefan. Ahora bien, para cuerpos grises es posible asumir una temperatura de brillo de emisión definida como la temperatura a la cual

la intensidad específica observada es la misma a la correspondiente para un cuerpo negro, es decir:

$$I_\nu = B_\nu(T_{Brillo}) \quad (2.6)$$

De tal forma que se cumple que:

$$T_{Brillo} = T_b = \epsilon T \quad (2.7)$$

Donde ϵ es la emisividad intrínseca del cuerpo y T es la temperatura absoluta (Macquart, 2010)

2.2.6. La Ecuación de Transferencia Radiativa (ETR)

Los procesos radiativos que tienen lugar en la atmósferas están descritos por la *ETR*, por este motivo a continuación se presenta una deducción de su forma más general, la cual sirve de fundamento para posteriores disertaciones sobre procesos de transferencia radiativa relacionados con la teledetección de la atmósfera en *PMW* que tienen lugar en este trabajo.

En este orden de ideas, es importante mencionar que dichos procesos tienen su origen en los mecanismos de absorción, emisión y dispersión comúnmente relacionados a la interacción de la radiación electromagnética con la atmósfera y la Tierra. En este sentido, la variación de intensidad asociada a una portadora de radiación electromagnética, luego de interactuar con un material está dada por la ecuación (2.8) según Kuo-Nan (1980).

$$dI_\lambda = -k_\lambda \rho I_\lambda ds \quad (2.8)$$

Donde ρ es la densidad del material, k_λ es la extinción de masa asociada al área transversal sobre la que incide la radiación de longitud de onda λ , I_λ es la intensidad de la radia-

ción incidente y ds es el diferencial de superficie sobre la que incide la radiación. La ecuación (2.8) permite determinar la reducción de intensidad en la radiación debido a la absorción y dispersión del medio.

Por otro lado la intensidad también puede ser fortalecida por emisión por parte de la materia y también por la suma de múltiples dispersiones a la portadora de radiación, en este caso la variación de intensidad está descrita por la ecuación (2.9).

$$dI_\lambda = j_\lambda \rho ds \quad (2.9)$$

Donde j_λ tiene el mismo significado físico que k_λ en la ecuación (2.8), a diferencia de que explícitamente representa un coeficiente asociado a la contribución de intensidad debida a emisión por parte del diferencial de superficie ds y las múltiples dispersiones que se suman a la portadora de radiación I_λ . Uniendo las ecuaciones (2.8) y (2.9), se obtiene la variación de intensidad debida a absorción, emisión y dispersión debido a la interacción con materia (Ver ecuación 2.10).

$$dI_\lambda = -k_\lambda \rho I_\lambda ds + j_\lambda \rho ds \quad (2.10)$$

Definiendo a la función de fuente $J_\lambda = j_\lambda / k_\lambda$ y despejando se tiene la forma general de la *ETR*:

$$\frac{dI_\lambda}{k_\lambda \rho ds} = -I_\lambda + J_\lambda \quad (2.11)$$

2.3. Redes Neuronales Artificiales (ANNs)

Las Redes Neuronales Artificiales (ANNs, por sus siglas en inglés) representan una de las técnicas que forman parte del área de *Inteligencia Artificial*. Haykin (1999) proporcionó

el siguiente concepto para ANNs: Una Red Neuronal Artificial es un procesador distribuido dotado de alto nivel de paralelismo construido a partir de unidades de procesamiento simples, las cuales tienen una capacidad natural de almacenar conocimiento conforme a la experiencia y disponer de dicho conocimiento para su uso. De acuerdo a [Mas and Flores \(2008\)](#) la definición presentada por [Haykin \(1999\)](#) asemeja a las ANNs con el cerebro en dos aspectos:

1. El conocimiento es adquirido por la red desde su medio, a través de un proceso de aprendizaje.
2. Las fuerzas de las conexiones inter-neuronas, conocidas como pesos sinápticos, son usadas para almacenar el conocimiento adquirido.

El primer aspecto está relacionado a las características de los datos de entrada y salida, así como con los mecanismos de entrenamiento de la red, mientras que el segundo aspecto está relacionado con la estructura de la ANN. A continuación se proporciona un tratamiento formal del conjunto de tópicos asociados con ANNs usados en esta investigación.

2.3.1. Estructura de una ANN

De acuerdo con [Haykin \(1999\)](#), toda estructura compleja de ANNs está conformada por una unidad de procesamiento simple denominada *neurona*. Una neurona puede considerarse como una entidad que emite una respuesta (salida) debido a la estimulación generada (activación) por el recibimiento de señales conocidas (entradas). En otras palabras, la neurona recibe información de entrada para la cual ella emite respuestas en función de un conocimiento previo, lo cual implica que si la neurona recibe estímulos desconocidos, la respuesta pudiera ser no acorde a lo esperado. La corrección de respuestas no esperadas por parte de la neurona que recibe nuevos estímulos, se resuelve mediante la aplicación de procesos de entrenamiento.

2.3.2. Modelo Matemático de una Neurona Artificial

En la figura 2.8 se aprecia el modelo matemático de una neurona, donde considerando que j es el índice de identificación de la neurona estimulada, se tiene que $\vec{X} = x_1, x_2, x_3, \dots, x_n$ es el vector de entradas que contiene información de n neuronas interconectadas con j .

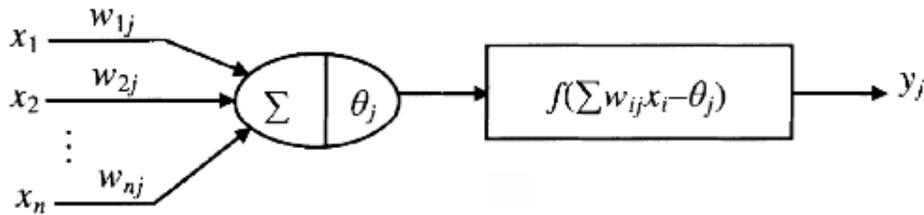


Figura 2.8: Modelo matemático de una neurona (Xingui and Shaohua, 2010).

Por otro lado w_{ij} es el peso sináptico entre cada una de las n neuronas, identificadas con $i = 1, 2, 3, \dots, n$, y la neurona j ; θ_j es el umbral de activación de la neurona j también conocido como *bias* y f es la función de activación asociada a la neurona j , generalmente se usa una función sigmoideal como la tangente hiperbólica. Finalmente la relación entre las entradas y la salida y_j de la neurona esta dada como:

$$y_j = f\left(\sum_{i=1}^n w_{ij}x_i - \theta_j\right) \quad (2.12)$$

La ecuación (2.12) representa una generalización simple de la ecuación de McCulloch-Pitts (Hertz et al., 1991), la cual hace referencia a la similitud entre el modelo biológico neuronal con el modelo matemático presentado, con la salvedad de que en este caso, la relación entrada-salida no considera dependencia temporal, es decir es una relación atemporal y por otro lado no se consideran en la salida efectos de acumulación debido a los datos de entrada, lo cual se traduce en que no se almacena información sobre estados anteriores de las entradas, sólo el momento actual (Xingui and Shaohua, 2010).

La forma en que las neuronas son dispuestas en una red determinan la arquitectura o topología de la red, la cual esta estrechamente relacionada con el algoritmo de entrenamiento

(Mas and Flores, 2008).

2.3.3. Perceptrón de una Capa

El modelo del Perceptrón fue introducido por Rosenblatt (1958, 1962) y también de forma independiente por Widrow and Hoff (1960), como un mecanismo capaz de ser entrenado en forma supervisada y a su vez como la arquitectura más simple de ANNs. En la figura 2.9 se muestra la arquitectura del perceptron, el mismo consta de una capa de entrada conectada a una capa de salida, en la capa de entrada el número de neuronas¹¹ n que es igual a la cantidad de datos de entrada x_i , mientras que en la capa de salida se tienen m neuronas de salida y_j .

La salida y_j de una neurona en términos de la combinación lineal de entradas x_i , está definida por una función de activación. Existen diferentes tipos de funciones de activación entre las cuales se destacan, por ser usados en la presente investigación, la función lineal y las funciones sigmoideas (Hiperbólica y Logística).

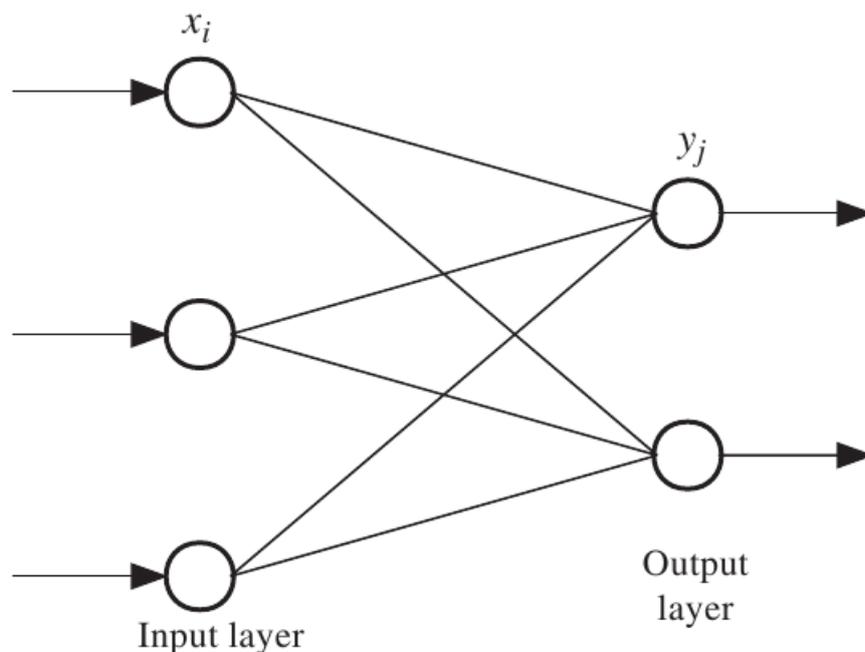


Figura 2.9: Modelo del Perceptron de una Capa (Hsieh, 2009).

¹¹También conocidas como nodos en la literatura sobre ANN

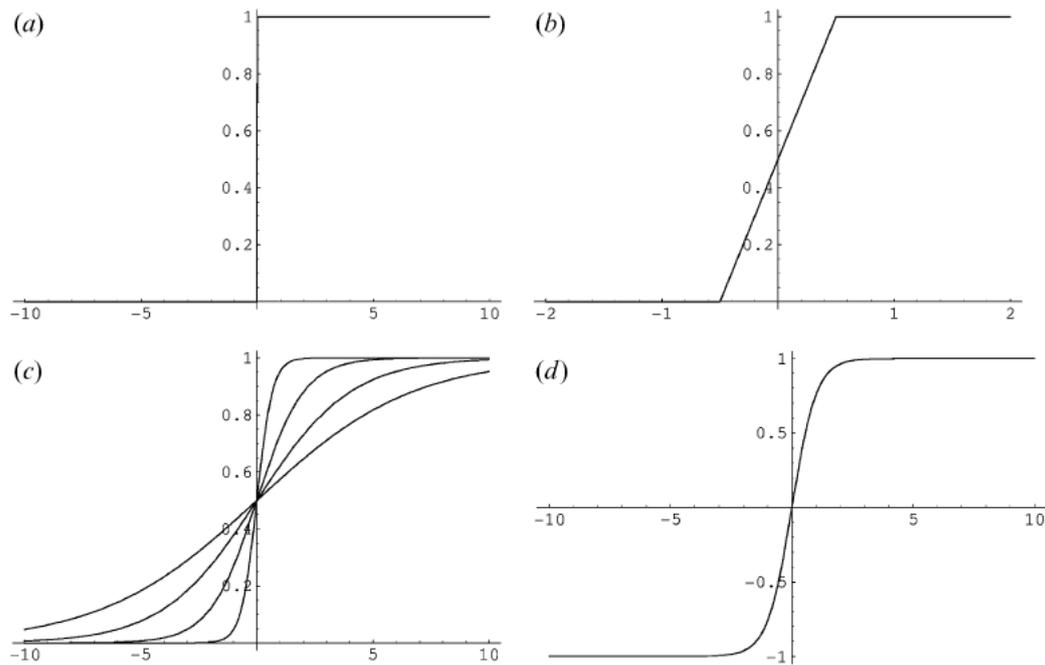


Figura 2.10: Funciones de activación: a) Escalón , b) Lineal, c) Logística, d) Hiperbólica (Mas and Flores, 2008).

La forma de las funciones sigmoidales¹² permiten simular la actividad de una neurona biológica, debido a que dada las combinaciones de los diferentes estímulos recibidos por una neurona, es posible describir estados de transición intermedios entre los estados absolutos de activación y no-activación en la respuesta neuronal. Tal y como se aprecia en la figura 2.10, la función de activación para respuestas binarias está representada en la figura 2.10.a por la función escalón, en la figura 2.10.b se muestra la función lineal, y en la 2.10.c y 2.10.d las funciones sigmoidales.

En general, cuando la salida de la ANN no es binaria, se requiere el empleo de funciones sigmoidales para modelar los diferentes estímulos transmitido por las neuronas. Las funciones sigmoidales empleadas son:

1. **Función Logística:** Esta función esta representada en la figura 2.10.c y esta dada por la ecuación (2.13), la misma tiene un valor asintótico en 0 cuando $a \rightarrow -\infty$, y un valor

¹²Son denominadas de esta forma en alegoría a que su graficación da como resultado una curva en forma de S

asintótico en 1 cuando $a \rightarrow +\infty$. Cuando se usa la función de activación logística las salidas de la ANN son interpretadas como una función discriminante, que pueden ser consideradas como probabilidades. Esta función es bastante empleada en algoritmos de clasificación

$$f(a) = \frac{1}{1 + \exp(-a)} \quad (2.13)$$

2. **Función Tangente Hiperbólica:** Esta función se muestra en la figura 2.10.d y escrita en la ecuación (2.14), la misma tiene un valor asintótico en -1 cuando $a \rightarrow -\infty$, y un valor asintótico en 1 cuando $a \rightarrow +\infty$.

$$f(a) = \tanh(a) \quad (2.14)$$

A diferencia de la función logística, esta función no muestra pesos sinápticos y bias positivos, motivo por el cual, hay preferencia por usar la función tangencial hiperbólica como función de activación en capas ocultas, ya que empíricamente ha demostrado mayor velocidad de convergencia que la función logística (LeCun et al., 1991).

En referencia al modelo del perceptrón debe añadirse que las limitaciones fueron puestas en evidencia en el trabajo de Minsky and Papert (1969), en el cual tras un estudio basado en aplicaciones de lógica operacional, el perceptrón no permitió conseguir soluciones para el problema del XOR, debido a que es un problema linealmente no separable, destacando por ende, la aplicabilidad del perceptrón exclusivamente a problemas linealmente separables.

2.3.4. Perceptrón Multi-Capas

Posterior al conocimiento de las limitaciones del modelo del perceptrón relacionadas a la linealidad de los datos de entrada y salida, se pensaba que esta dificultad podía ser superada mediante la inclusión de capas ocultas entre la capa de entrada y salida, sin embargo para la época no había ningún algoritmo para resolver ANNs multi-capas. El trabajo de Rumelhart

et al. (1986) fijó las bases que permitieron posteriormente el desarrollo del Perceptrón Multi-Capa (*MLP*, por sus siglas en inglés). En el mismo se presenta el re-descubrimiento del Algoritmo de Propagación Dirigida hacia Atrás, el cual había sido presentado previamente por Werbos (1974).

En la figura 2.11 se muestra la estructura de un *MLP* de arquitectura $x_i-h_j-y_k$ ¹³. Esta arquitectura de *ANN* es conocida como Perceptrón Multi-Capa Dirigido hacia Adelante (*FFMLP*, por sus siglas en inglés) la cual se caracteriza por el hecho de que, cada neurona de cada capa, esta conectada con cada neurona de la próxima capa. En este caso se deben implementar dos funciones de activación una para activar a las neuronas de la capa oculta al ser estimuladas por las neuronas de la capa de entradas, y otra para activar las neuronas de la capa de salida tras ser estimuladas por las neuronas de la capa oculta. En este contexto considerando en primer lugar la salida de las neuronas h_j de la capa oculta, se tiene la ecuación (2.15):

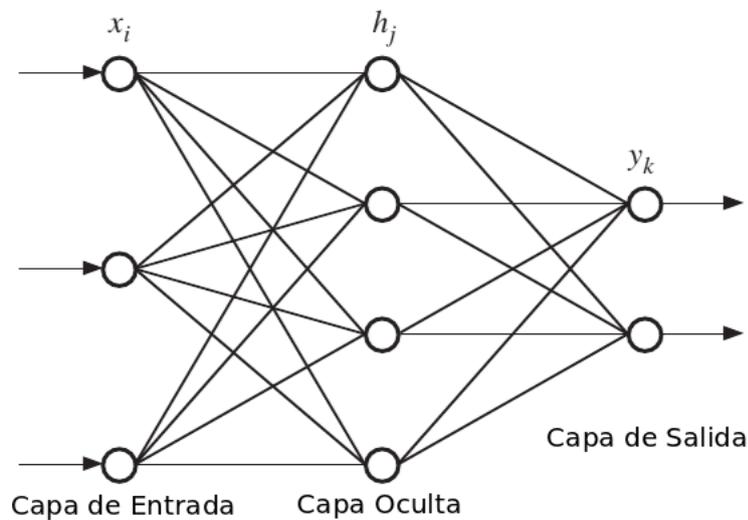


Figura 2.11: *MLP*, se muestra la capa de entrada de datos, una capa oculta, y una de capa de salida (Hsieh, 2009).

$$h_j = f \left(\sum_{i=1}^n w_{ij} x_i - b_j \right) \quad (2.15)$$

Donde, de forma similar a la ecuación (2.12), w_{ij} es el peso sináptico entre cada

¹³Una capa de entrada de i neuronas, una oculta de j neuronas y una capa de salida de k neuronas

neurona de entrada i con la j -ésima neurona de la capa oculta, f es la función de activación y b_j es el umbral de activación de la neurona de la capa oculta j . Por otra parte, la salida para la neurona y_k de la capa de salida esta dada por la ecuación (2.16).

$$y_k = g \left(\sum_{j=1}^n \check{w}_{jk} h_j - \check{b}_k \right) \quad (2.16)$$

Donde \check{w}_{jk} es el peso sináptico entre cada neurona j de la capa oculta con cada neurona y_k de la capa de salida y g es la función de activación. La extensión de la figura 2.11, en la cual se presenta un *MLP* de una capa oculta, para el caso de N capas ocultas resulta natural, siendo necesario obtener la salida en dependencia de la función de activación entre capas sucesivas, de tal forma como se mostró para h_j y y_k en este ejemplo.

2.4. Entrenamiento de ANNs

De acuerdo con [Blackwell and Chen \(2009\)](#) generalmente los problemas atmosféricos basados en ANNs están enfocados como un problema de clasificación mediante reconocimiento de patrones (*RP*, por sus siglas en inglés) o como un problema de regresión de funciones. En estimaciones de precipitación, la técnica de *RP* es empleada generalmente para identificar pixeles de precipitación, mientras que la regresión de funciones es empleada para estimar tasas de precipitación. El objetivo general de este trabajo es determinar tasa de precipitación, por lo cual el marco teórico presentado a continuación está vinculado al entrenamiento de ANNs orientadas a la regresión de funciones.

En este sentido es propicio el contexto para denotar que, se define entrenamiento de una ANN al proceso de determinar valores óptimos de pesos sinápticos y bias que maximicen la aproximación del conjunto de salidas de la ANN con el conjunto de datos que representan la meta¹⁴ en la fase de entrenamiento. Las características intrínsecas de la ANN determinan en gran medida su capacidad y propiedad para ser entrenadas, entre estas características se pueden mencionar estructura de la ANN, forma de inicialización de la red, así como fases de

¹⁴Datos objetivos

regularización que permiten la auto-evaluación en tiempo real del proceso de entrenamiento. El proceso de post-entrenamiento requiere además un proceso de evaluación de la calidad del proceso, el cual puede ser realizado mediante regresión lineal entre la salida de la red y los datos objetivos, así como también la estimación del coeficiente de correlación.

El entrenamiento de una *ANN* puede tener diferentes objetivos, en este sentido, en la presente investigación se requiere que las *ANNs* sean entrenadas a partir de un conjunto de datos de entrenamiento. De acuerdo con Hsieh (2009) el proceso de entrenamiento generalmente es llevado a cabo al minimizar una función de costo¹⁵ J , que particularmente, siguiendo la metodología planteada por Hsieh (2009), ha sido definida como un medio del error medio cuadrático (*MSE*, por sus siglas en inglés) entre la salida obtenida por la *ANN* y el conjunto de datos objetivos usado durante el proceso de entrenamiento, y matemáticamente está dado a continuación:

$$J = \frac{1}{N} \sum_{n=1}^N \left\{ \frac{1}{2} \sum_k [y_k^n - y_{dk}^n]^2 \right\} \quad (2.17)$$

En la ecuación (2.17) y_{dk} representa el dato objetivo (meta proporcionada en el entrenamiento) y y_k es la salida de la *ANN* y $n = 1, 2, 3, \dots, N$ son las observaciones o mediciones. Entonces en resumen se puede plantear que el proceso de entrenamiento de *ANNs* consiste en el empleo de algoritmos de optimización (comúnmente denominados algoritmos de entrenamiento) que permitan determinar valores w y b para los cuales la función de costo J resulte minimizada.

2.4.1. Inicialización de la *ANN*

Una de las características asociadas a la calidad del entrenamiento, es la inicialización de la red. La fase de inicialización consiste en asignar valores para los pesos sinápticos y las bias a partir de los cuales tendrá inicio el proceso de optimización llevado a cabo por el algoritmo de entrenamiento. Métodos primitivos sugieren la asignación de valores aleatorios

¹⁵También conocida como función de pérdida, función de errores o función objetivo

para este proceso, sin embargo debido a que la calidad del entrenamiento está relacionada a esta etapa de inicialización, se han desarrollado diversos métodos para facilitar el proceso de optimización de pesos y bias a partir de un mecanismo de inicialización pre-diseñado.

2.4.2. Entrenamiento mediante Propagación de Errores hacia Atrás (BP)

El termino Propagación hacia Atrás (*BP*, por sus siglas en inglés) tiene diferentes connotaciones en el campo de las *ANNs*, por una parte se suele asociar este termino a la arquitectura de la *ANN*, y por otra parte también se denomina *BP* al algoritmo de minimización de la función de costo presentada por [Rumelhart et al. \(1986\)](#), que es en el contexto en el cual este termino será empleado en la presente investigación.

De acuerdo con [Bishop \(2005\)](#) el algoritmo *BP* consiste en el entrenamiento de *MLPs* mediante la aplicación del método del Gradiente Descendiente (*DG*, por sus siglas en inglés) sobre la función de costo, sin embargo cabe destacar que de acuerdo a [Xingui and Shaohua \(2010\)](#) existen varios métodos mejorados basados en *BP*, como por ejemplo el Gradiente Conjugado, Newton, Quasi-Newton, Levenberg-Maquardt, entre otros. La optimización de pesos y bias es llevada a cabo mediante *BP* usando Gradiente Descendiente a partir de la actualización iterativa en la dirección en la que la función de costo disminuye mas rápido, es decir en la dirección negativa del gradiente.

En esta investigación se usa el algoritmo Quasi-Newton para el proceso de entrenamiento de las *ANNs*, debido principalmente a que es un algoritmo de segundo orden mejorado, lo cual implica una buena relación entre rendimiento y tiempo requerido para el entrenamiento, por otro lado este algoritmo está implementado en dos librerías de licencia abierta para C++, cuyas salidas son evaluadas en este trabajo para seleccionar la mejor salida, estas librerías son Open NN y ALGLIB.

2.5. Características Generales de las Librerías Open NN y ALGLIB

2.5.1. Características de ALGLIB

ALGLIB es un conjunto de herramientas computacionales y numéricas, orientadas a implementaciones matemáticas aplicadas. Contiene interfaces de desarrollo para la reproducción y entrenamiento de ANNs, que están compuestas por librerías disponibles para ser compiladas y usadas en varios lenguajes de programación. En este trabajo las implementaciones han sido llevadas a cabo en C++ para establecer una comparación con Openn NN, que está desarrollada exclusivamente en este lenguaje.

Dentro de las ventajas que aporta ALGLIB para implementaciones de ANNs resalta el hecho de que se pueden entrenar ANNs tanto para clasificación como para regresión lineal, por otra parte, cuenta con tres algoritmos de entrenamiento que son, el L-BFGS que es un método Quasi-Newton con costo de iteraciones fijas cuyo uso es recomendable en problemas con gran cantidad de datos, el Levenberg-Marquardt que está fundamentado en el uso de la función de error exacta Hessiana, y el método de parada temprana, el cual evita el sobreentrenamiento a partir de la finalización automática del proceso de entrenamiento mediante parámetros determinados en el algoritmo .

Una ventaja destacable de la implementación del algoritmo Quasi-Newton en ALGLIB, es el hecho de que presenta criterios de parada predefinidos, lo cual garantiza una convergencia óptima durante el proceso de entrenamiento, y al mismo tiempo proporciona 3 parámetros que permiten la modificación del criterio de parada sin perder generalidad en las características de optimización. De este modo, es posible establecer procedimientos de entrenamiento de ANNs, con pocos pasos de configuración del algoritmo y de forma optimizada¹⁶.

¹⁶<http://www.alglib.net/dataanalysis/neuralnetworks.php>

2.5.2. Características de Open NN

Open Neural Network (Open NN) es una librería basada en C++ para el entrenamiento y reproducción de ANNs. Posee herramientas para el manejo y disposición de datos, además de contar con una amplia variedad de algoritmos de entrenamiento entre los que resaltan el método del Gradiente Descendiente, Gradiente Conjugado, Newton, Quasi-Newton, Levenberg-Marquardt, entre otros. Al igual que ALGLIB proporciona herramientas para entrenar redes para clasificación y regresión lineal, pero a diferencia de ALGLIB, Open NN tiene una documentación que incluye más de 10 ejemplos de soluciones de problemas a partir del entrenamiento de ANNs por diferentes algoritmos.

La arquitectura de Open NN, y la implementación del algoritmo Quasi-Newton, permiten al usuario crear estrategias de entrenamiento a partir de la configuración de un criterio de parada por un incremento mínimo de desempeño, cumplimiento de meta, meta de gradiente normal, y también por la evaluación de decremento de generalización máxima. También es posible establecer configuraciones para llevar a cabo el proceso de entrenamiento en 3 fases, de este modo se usa un primer algoritmo para llevar a estabilidad y convergencia los bías y pesos sinápticos (Generalmente un algoritmo de primer orden como el Gradiente Descendiente), luego el segundo algoritmo que realiza el proceso de entrenamiento, y un tercer algoritmo para optimizar el proceso.

Para los efectos de esta investigación se ha usado un solo algoritmo, para evaluar las características de ajuste de cada librería en condiciones similares. De igual forma es importante mencionar que el criterio de parada utilizado fue el de la evolución de generalización, ya que evita el sobreentrenamiento de la ANN¹⁷.

¹⁷<http://www.cimne.com/flood/Links.asp>

CAPÍTULO 3

METODOLOGÍA DE LA INVESTIGACIÓN

En el presente capítulo se proporcionan detalles sobre el diseño de la metodología de investigación empleado para cumplir los objetivos planteados en este trabajo. Partiendo de los objetivos de investigación propuestos en el capítulo 1, este trabajo puede descomponerse en 6 etapas, las cuales se muestran en la figura (3.1), donde es importante destacar que la etapa 1 ya fué cubierta en el capítulo 2 de este trabajo. Por otra parte, la etapa 2 tiene por objeto evaluar las características específicas del sensor *ATMS* para los fines del desarrollo del algoritmo de precipitación, de tal forma que las características de dicho algoritmo estarán sujetas a las potencialidades intrínsecas del *ATMS*.

En referencia a la etapa 3, destaca el hecho de que se entrenan las redes neuronales que sirven de base al algoritmo de estimación de tasa de precipitación. Una vez entrenadas las redes neuronales, sigue la etapa 4, donde se diseña el algoritmo para estimar tasa de precipitación usando datos del sensor *ATMS*. Y finalmente las etapas 5 y 6 involucran pasos de calibración de las salidas de las redes neuronales tomando como referencia el ajuste de

correlación en reproducciones con datos conocidos.

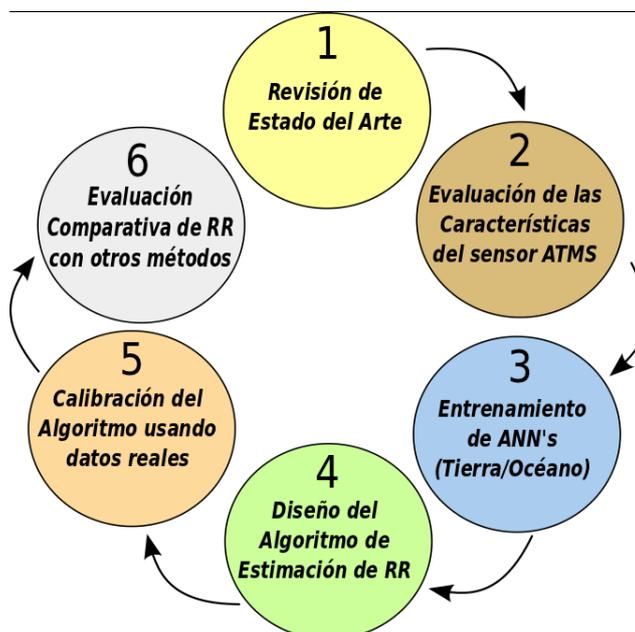


Figura 3.1: Etapas para el cumplimiento del objetivo general del proyecto de investigación.

Desde el punto de vista metodológico esta investigación se enmarca en el método de investigación descriptiva presentado por [Arias \(1999\)](#), en la cual se desarrolla un algoritmo para la estimación de tasa de precipitación a partir del entrenamiento de redes neuronales usando datos simulados del sensor *ATMS* a bordo del satélite *Suomi-NPP*. Con este fin se evalúa el entrenamiento de redes para océano y tierra, tanto para invierno como verano, mediante el uso del algoritmo *Quasi-Newton*.

La implementación de los algoritmos de entrenamiento para las *ANNs* es llevada a cabo a partir del uso de dos librerías de código abierto para C++ (**ALGLIB** y **OPENNN**). De los resultados proporcionados por cada librería se seleccionan únicamente los resultados óptimos según sea el caso; para esto, se construye una serie de experimentos computacionales, los cuales consisten en el entrenamiento de las *ANNs* para diferentes parámetros de entrenamiento, condiciones de borde y arquitectura de red, con la finalidad de obtener resultados de correlación entre los resultados de tasa de precipitación obtenido a partir de dichas redes, en relación a mediciones realizadas mediante técnicas previamente validadas.

3.1. Algoritmo de Estimación de Tasa de Precipitación (*RR*)

El algoritmo presentado en esta investigación, tiene por objeto estimar tasa de precipitación instantánea a partir de una escena del sensor **ATMS**. Está conformado por 6 módulos, los cuales pueden ser identificados en la figura 3.2, como lectura de variables, cálculo de las temperaturas de brillo (BTn), selección de canales, preparación de archivos de datos de entrada para la red neuronal, producción de la red neuronal, y elaboración de los mapas de precipitación.

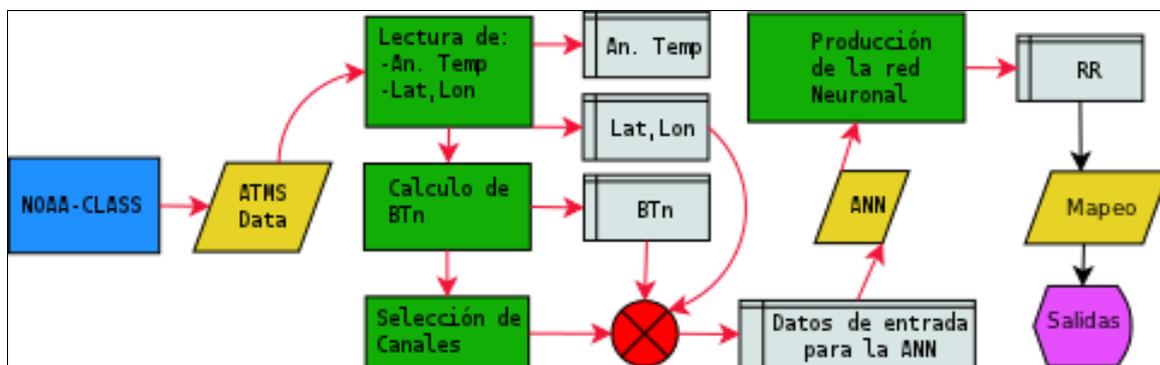


Figura 3.2: Diagrama de Flujo del algoritmo para la estimación de tasa de precipitación.

En el bloque **NOAA-CLASS** se realiza la descarga de los datos provenientes del sensor **ATMS**, **NOAA-CLASS** es una herramienta de gestión de datos proporcionada por NOAA que facilita la solicitud, selección y descarga de datos de las plataformas satelitales que administra esta institución. Los datos se descargan en formato HDF5, y los mismos vienen expresados en Números Digitales (DN, por sus siglas en inglés) y deben ser calibrados para ser expresados en valores de temperatura de brillo. Este proceso de calibración o conversión es llevado a cabo a partir de la siguiente ecuación:

$$T_b = (330 / (2^{16} - 9)) * DN \quad (3.1)$$

Donde T_b es la temperatura de brillo y DN el valor en números digitales. El proceso de lectura de los datos HDF5, así como el de calibración a temperatura de brillo, es llevado a cabo a partir de un módulo computacional desarrollado por el autor de este trabajo, basado

en la librería oficial de HDF5 para C++. Ahora bien la siguiente fase para el desarrollo del algoritmo comprende la fase de selección de canales **ATMS** para el diseño de las redes neuronales, luego de esto se entrenan las redes, se producen y se obtiene como resultado una grilla georeferenciada con valores de tasa de precipitación, que posteriormente son mapeados en una etapa de graficación.

Es importante destacar, que los datos **ATMS**, correspondientes a la temperatura de la antena¹, son descargados como un arreglo tridimensional de datos², y posteriormente son descompuestos en matrices bidimensionales para llevar a cabo de forma eficiente las operaciones de cálculo. Al final del proceso estas matrices son utilizadas para la elaboración de los mapas de precipitación instantánea. En el apéndice [A](#) se proporciona una guía detallada para la ejecución del algoritmo, en el cual se puede seguir de forma práctica cada una de las partes mostradas en el diagrama de flujo mostrado en la figura [3.2](#).

3.1.1. Diseño de la Red Neuronal

De acuerdo con [Bellerby et al. \(2000\)](#) una red neuronal de dos capas ocultas proporciona considerables niveles de complejidad, en cuanto a la no-linealidad entre los datos de entrada y salida, en relación con el costo computacional requerido para el proceso de regresión no-lineal característica del proceso de estimación de tasa de precipitación usando *ANNs* en base a datos satelitales. En este sentido, con la finalidad de emular los fenómenos físicos involucrados en el proceso de precipitación, en este trabajo se entrenan diferentes modelos de redes neuronales, todos con una capa de entrada, dos capas ocultas y una capa de salida, para evaluar la capacidad por parte de dichas redes de predecir valores de tasa de precipitación en superficie usando como datos de entrada diferentes combinaciones de canales del sensor *ATMS*.

La topología de red descrita en el párrafo anterior se mantiene en los diversos experimentos que tienen lugar en esta investigación, teniendo lugar a variaciones únicamente en el

¹temperatura de la antena es el nombre que se le asigna al conjunto de datos en número digital, asociados a la emisión de radiación del terreno que conforma una escena de **ATMS**

²Las dimensiones están dadas por longitud, latitud y los diferentes canales

número de neuronas en la capa de entradas, el cual depende de los canales que sean seleccionados como variables de entrada para la composición de cada red, y el número de neuronas en las capas ocultas. En este orden de ideas, con la primicia de proporcionar mayor información sobre la arquitectura de red neuronal usada en este trabajo, en la figura 3.3 se muestran las diferentes partes que conforman la red neuronal, las cuales se describen respectivamente a continuación.

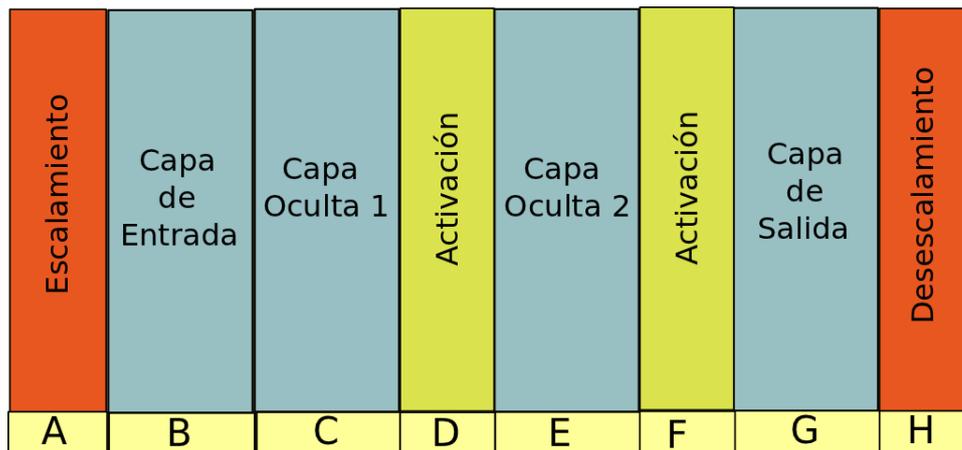


Figura 3.3: Diseño secuencial de la topología de red neuronal de dos capas ocultas

Capa de Entrada

La importancia de la capa de entrada se encuentra implícita en el hecho de que la red neuronal debe emular las relaciones físicas y matemáticas que tienen lugar entre las variables que conforman los datos de entrada y las salidas que representan las magnitudes predichas por la red neuronal. En este sentido, deben realizarse varias consideraciones con la finalidad de llevar a cabo una óptima selección de los datos de entrada, entre las que cabe resaltar en primer lugar el hecho de que dichas variables tengan poca correlación entre sí, pero alta correlación con la salida de la ANN, así como también que contengan poco contenido de ruido en el caso de que los datos provengan de instrumentos de medición ³.

Con respecto a la estructura de la capa de entrada, es importante precisar que el conjun-

³En muchos casos tienen su origen en simulaciones o mediciones indirectas sobre variables exentas de ruidos

to de variables de entrada definen el número de neuronas en esta capa. Así pues, el número de variables de entrada tiene una relación directa con el diseño de la estructura de neuronas de las capas ocultas de la red y por ende con el desempeño de una red entrenada. Es importante además resaltar que en el proceso de entrenamiento de la red neuronal, los datos de entrada están conformados por variables de entrada, representadas por los 22 canales del sensor **ATMS** y una variable de salida, que es la tasa de precipitación instantánea.

Las variables de entrada están representadas por el conjunto de canales del *ATMS* que tienen una alta relación con la posibilidad de predecir la tasa de precipitación, y por otro lado los datos de salida son valores de tasa de precipitación simulados para las diferentes combinaciones de los datos de entradas. Posterior a esta separación descriptiva sobre los datos entrada, los mismos son divididos en tres subconjuntos de datos; el primero corresponde a los datos de entrenamiento para la red⁴, el segundo a datos que sirven para la verificación en tiempo real del proceso de entrenamiento⁵ y el tercero para realizar pruebas sobre el desempeño predictivo de la red neuronal⁶.

Escalamiento y Desescalamiento de los Datos

En la figura 3.3 se puede apreciar en la etiqueta A que, antes de que los datos sean integrados a la capa de entrada, hay un proceso de escalamiento. El proceso de escalamiento puede ser llevado a cabo cuando los datos presentan linealidad en cuanto a la escala, por otro lado es conveniente hacerlo para expresar los datos de entrada en orden de magnitud cero, lo cual ayuda a una mejor aproximación por parte del algoritmo de entrenamiento.

Existen varias técnicas de escalamiento de datos tales como normalizando a media cero y varianza unitaria, normalización percentil y normalización por rango. En este trabajo se ha utilizado la normalización de media cero y varianza unitaria, el cual consiste en forzar a cada valor de la secuencia de datos a caer en una distribución uniforme centrada en cero, de tal forma que dada una serie de datos $X \in \{x_1, x_2, x_3, \dots, x_i\}$, donde x_i están dados

⁴Datos con los que se entrena la red

⁵Datos de referencia para el cálculo de error medio cuadrático

⁶Datos de prueba para la reproducción de la red neuronal

por los diferentes canales que comprende el sensor *ATMS*, el valor escalado y_i del i -ésimo componente de X , x_i esta dado como:

$$y_i = \frac{x_i - \bar{x}}{\sigma} \quad (3.2)$$

En la ecuación 3.2, \bar{x} es la media de X y σ es la desviación estándar. Asimismo, en la etiqueta H de la figura 3.3 se aprecia una capa de desescalamiento, la cual es necesaria para expresar la salida de la red neuronal en el orden de magnitud respectivo a la variable de salida, para lograr este desescalamiento, es suficiente despejar desde la ecuación 3.2 a x_i , obteniendo:

$$x_i = y_i * \sigma + \bar{x} \quad (3.3)$$

Capas Ocultas

En la arquitectura de las redes que son entrenadas para los efectos de este trabajo, se han incluido dos capas ocultas, siguiendo las recomendaciones de [Blackwell and Chen \(2005\)](#). De este modo, la arquitectura base para la evaluación del entrenamiento de redes neuronales es 10 neuronas en la primera capa oculta y 5 en la segunda, tomando como referencia el trabajo de [Chen et al. \(2006\)](#). A partir de esta arquitectura base, se evalúa el desempeño de cada red entrenada, y en función de la calidad de resultados obtenidos se plantean experimentos alternativos con diferentes números de neuronas en ambas capas ocultas.

Función de activación

Este tópico fue detalladamente explicado en el capítulo 2, donde en la figura 2.10 se muestran los cuatro tipos de funciones de activaciones más comunes dentro del campo de las redes neuronales. En este sentido, en la figura 3.3 se aprecian 2 capas de activación, una con la etiqueta D que corresponde a las neuronas de la capa oculta 1 y otra con la etiqueta

F que esta asociada a la capa oculta 2, en ambos casos se ha seleccionado como función de activación a la función hiperbólica cuya expresión puede observarse en la ecuación (2.14).

Es importante mencionar, que en la figura 3.3 ha sido ignorada una tercera capa de activación que corresponde a las neuronas de la capa de salida, esta omisión se debe al hecho de que la función de activación aplicada en esta etapa es la lineal, la cual para los efectos de este trabajo ha sido implementada en una relación uno a uno.

Capa de Salida

La capa de salida, esta compuesta por una neurona, que corresponde a la tasa de precipitación predicha por la red neuronal. Es importante destacar que los datos del sensor **ATMS** ingresados al algoritmo para una estimación, corresponden a una escena capturada por el sensor en el intervalo de tiempo de 8 minutos aproximadamente. Esto significa que, entre la primera línea que conforma la escena y la última hay una diferencia de tiempo de 8 minutos, aun así, se considera que la escena es capturada en un tiempo único, lo cual para las escalas temporales que tienen lugar en los procesos dinámicos de la atmósfera y especialmente en las estructuras verticales de las nubes resulta una aproximación aceptable.

Por lo que antecede, en referencia a los mecanismos de captura del sensor **ATMS** y a la naturaleza temporal de sus datos, se destaca que la tasa de precipitación proporcionada por la capa de salida de la red neuronal es instantánea, y espacialmente corresponde a la coordenada geográfica asociada a cada pixel que conforma la escena.

3.1.2. Observación de la Atmósfera a partir de los canales *ATMS*

La información que proporciona cada canal del sensor *ATMS* está fuertemente relacionada con el mecanismo de transferencia radiativa de la atmósfera en el rango de las microondas, por esta razón en la figura 3.4 se muestran las principales características de transmisión de microondas en la atmósfera, donde además se pueden observar las fuertes líneas de absorción por parte del Oxígeno a 54Ghz y a 118GHz y la del vapor de agua a 183Ghz, así como

los intervalos de frecuencia de los canales ventana.

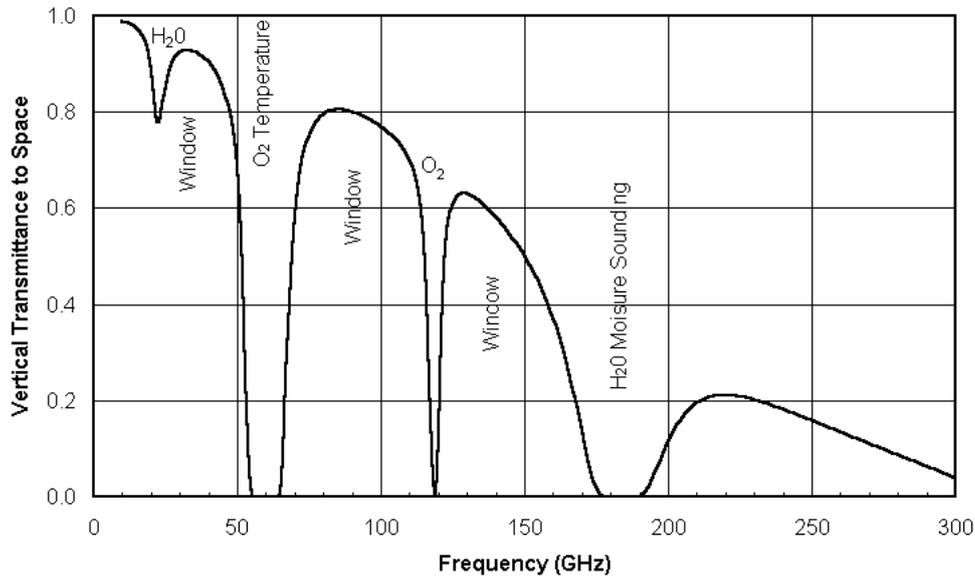


Figura 3.4: Transmitancia Vertical en la Atmósfera en el rango microondas (www.amsu.cira.colostate.edu)

En general, los canales con frecuencias cercanas a estas líneas de absorción se denominan canales opacos en virtud de que la absorción por parte de la atmósfera es bastante alta. Los canales opacos cercanos a 54GHz son útiles para la construcción de perfiles verticales de temperatura, mientras que los cercanos a 183GHz son sensibles a los perfiles verticales de vapor de agua. Por otro lado, los canales dentro de los intervalos ventana se caracterizan por altos valores de transmitancia en la atmósfera para las emisiones desde superficie.

Con la finalidad de contar con información detallada para la selección de los canales *ATMS* que se usarán en el algoritmo de estimación de precipitación, a continuación se destacan algunas características relevantes relacionadas con la potencialidad para la estimación de precipitación que ofrece cada canal del sensor *ATMS*. En este sentido se ha seleccionado como marco de observación una región centrada en Venezuela, ya que facilita la apreciación de las diferencias de sensibilidad entre océano y tierra proporcionado por cada canal. Adicionalmente es importante destacar que el escenario ilustrado en las imágenes de esta sección corresponden a la órbita ascendente con fecha del 5 de marzo del 2012 en horario de barrido

comprendido entre las 5:34pm y 5:42pm.

La figura 3.5 muestra el producto de temperatura de brillo (*BT*, por sus siglas en inglés) para los primeros 6 canales del sensor ATMS, a cada subfigura se le ha asociado una etiqueta en la cual se especifica a que canal corresponde cada imagen. Para el canal uno, el cual esta centrado en 23,80GHz dentro de una ventana atmosférica, en primer lugar cabe destacar la alta sensibilidad a las características de emisión de superficie, permitiendo identificar tipo de superficie, entre océano y tierra. Por otro lado, en los círculos negros se notan zonas de regiones frías en comparación al entorno, lo cual se atribuye a la sensibilidad de este canal a las emisiones del vapor de agua y a las nubes de agua líquida.

En el canal 2 centrado en la frecuencia de 31,4GHz, se aprecia en el círculo rojo, una zona caliente sobre océano, esta característica esta asociada a la sensibilidad de este canal a la emisión por vapor de agua y nubes de agua, lo cual hace de este canal ideal para estimar propiedades de nubes y precipitaciones sobre océano. También cabe señalar que los círculos negros en este caso, al igual que los de la imagen del canal 1, están asociados a la emisión del vapor de agua y nubes de agua líquida, con la diferencia de que ahora son mas distinguibles y de hecho se han hecho notables nuevos sistemas que no son bien definidos en el canal 1. Por esta sensibilidad asociada a la emisión del vapor de agua y de nubes de agua líquida estos dos canales son definidos como canales de ventana para el vapor de agua⁷.

Ahora bien, el canal 3 esta centrado en la frecuencia de 50.3GHz y al igual que los canales 1 y 2 es sensible a las características de superficie, en la figura 3.5 se puede distinguir claramente entre océano y superficie de tierra, así como también se aprecia en los círculos de trazos azules nubes de agua líquida sobre tierra, cabe destacar que a diferencia de los canales 1 y 2, este canal experimenta dispersiones debido a los cristales de hielo en las nubes, el resultado de esta dispersión se aprecia en los círculos de trazos negros de la figura.

⁷<https://directory.eoportal.org/eoportal/satellite-missions/s/suomi-npp>

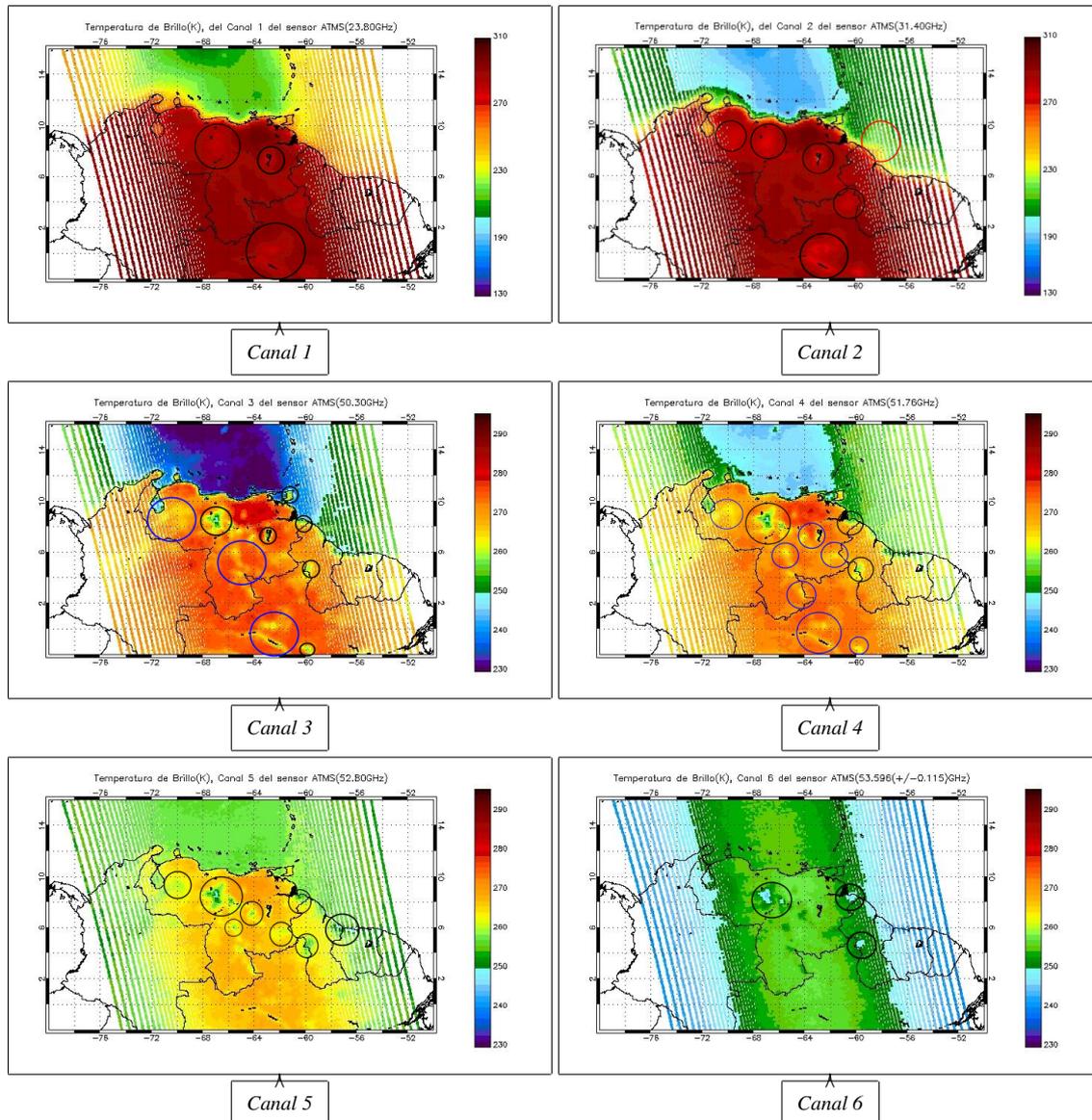


Figura 3.5: Temperatura de Brillo para los canales 1 hasta el 6 del sensor *ATMS*; los círculos negros resaltan cristales de hielo asociados a desarrollos de nubes convectivas, los azules están relacionados con precipitación o nubes de agua líquida y los círculos rojos a nubes sobre pixeles de océano

En la figura 3.5, también se aprecia para el canal 4 centrado en 51,76GHz y el canal 5 centrado en 52,80GHz una buena sensibilidad para detectar cristales de hielo en las nubes, así como también permiten discriminar nubes precipitantes o de agua líquida, pero en menor grado que los canales 1 y 2.

Por otra parte, el canal 6 (Ver tabla 2.3 para mayor información sobre este y el resto de los canales *ATMS*) muestra en los círculos negro los desarrollos verticales mas intensos de la escena, sin embargo también se aprecia un patrón de ruido transversal que se extiende desde los pixeles del centro de la escena hacia ambos extremos de la escena (*swath*). Este ruido se nota en la figura del canal 6 como una especie de gradiente de temperatura de brillo dirigido hacia ambos extremos del *swath*.

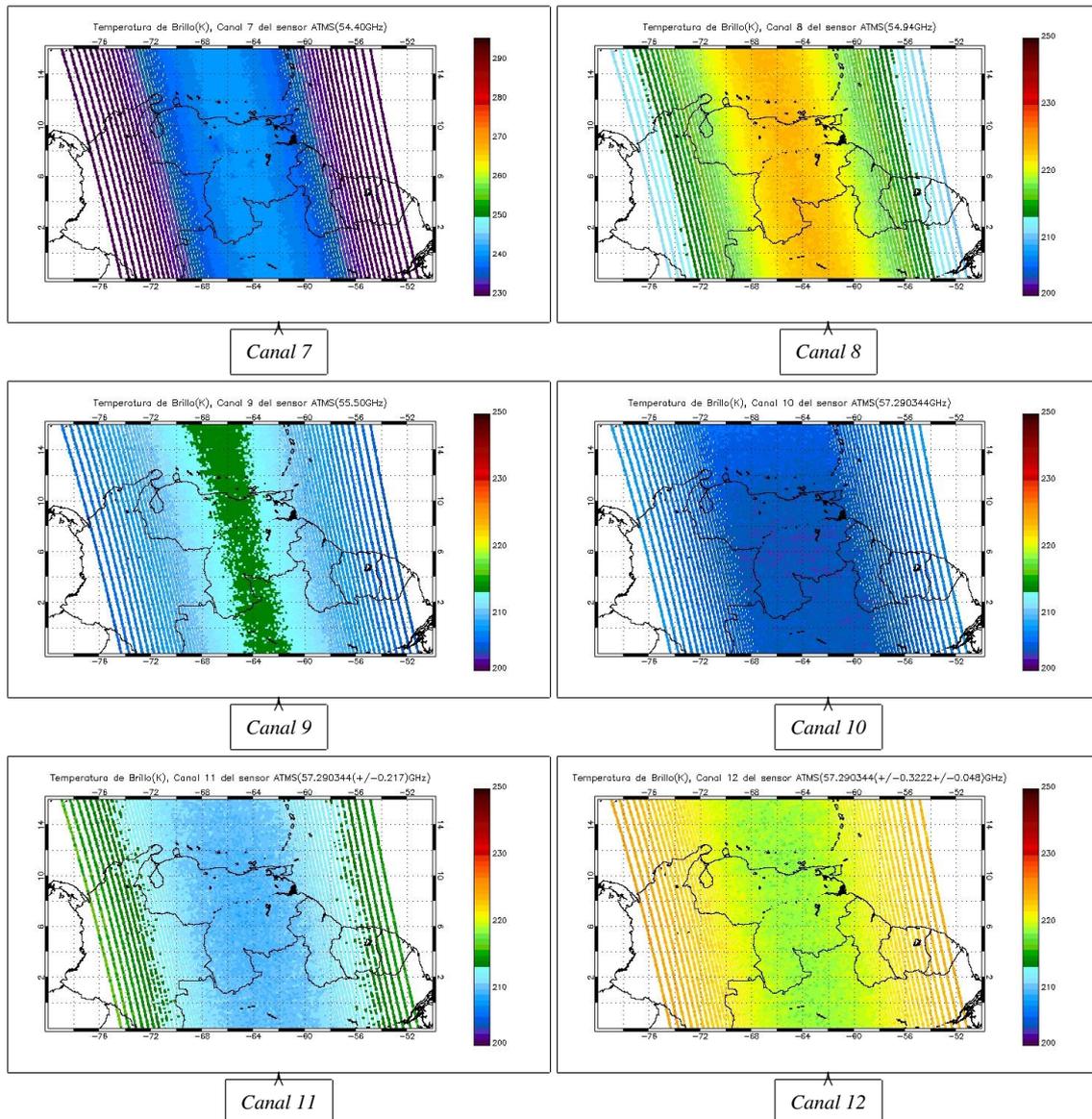


Figura 3.6: Temperatura de Brillo para los canales 7 hasta el 12 del sensor *ATMS*

Por otra parte, en la figura 3.6 se muestran los canales 7 hasta el 12, estos canales también presentan el mismo error que el canal 6, cabe destacar que entre estos canales, los

canales 6,7 y 8 desde el punto de vista técnico aportan información importante relacionada con las nubes y la precipitación, sin embargo sin una corrección previa estos canales no pueden ser incorporados al algoritmo de estimación de precipitación.

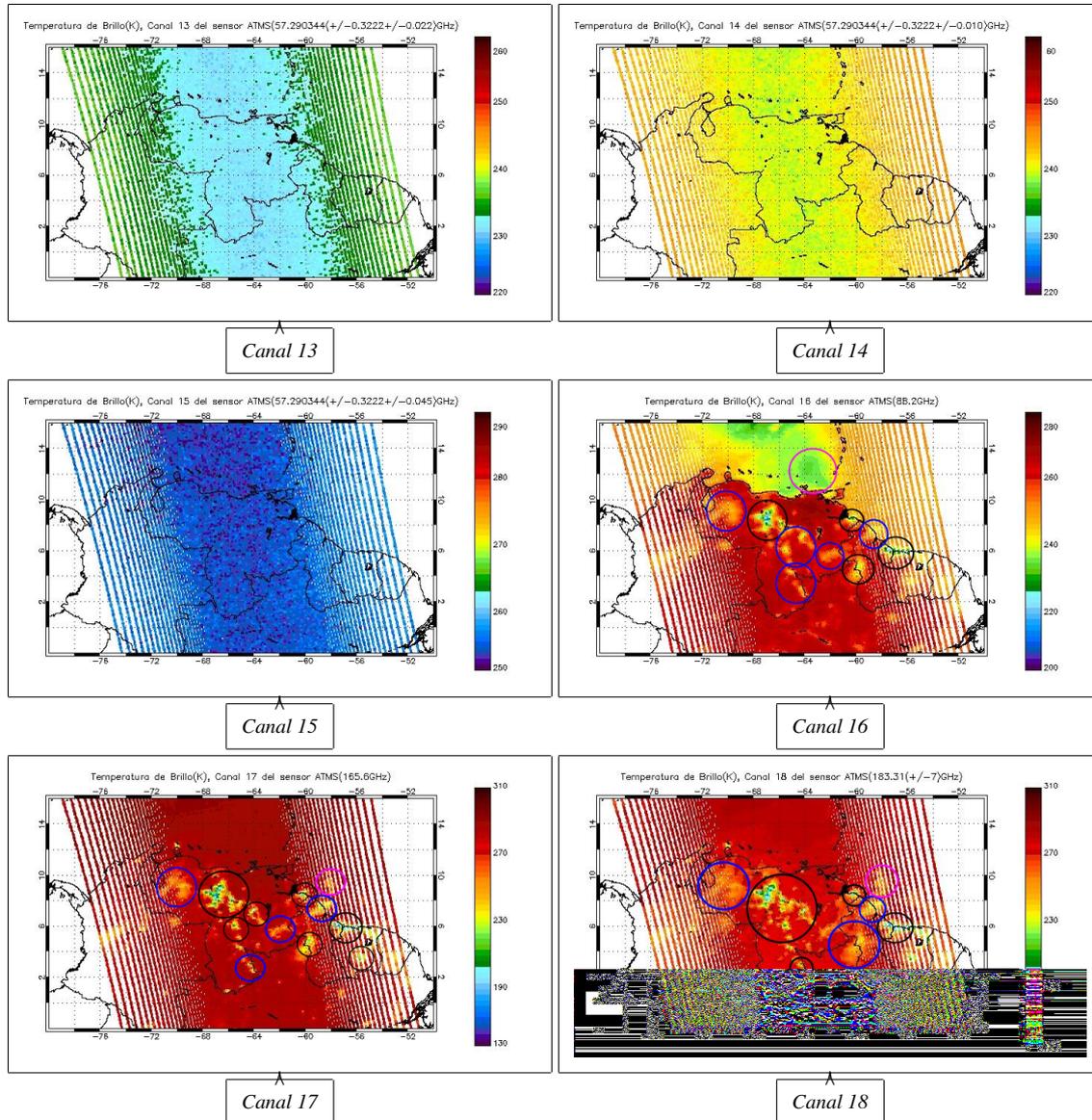


Figura 3.7: Temperatura de Brillo para los canales 13 hasta el 18 del sensor *ATMS*. De acuerdo a la figura 2.2, los canales 13,14 y 15 tienen poca relación con los mecanismos asociados a la precipitación, mientras que en el resto de las imágenes los círculos negros corresponden a cristales de hielo en nubes convectivas, los azules precipitación y nubes de agua líquida, y en púrpura nubes sobre océano

También en la figura 3.7, se muestra el conjunto de imágenes conformado por los

canales que van desde el 13 hasta el canal 18, donde se aprecia además que los canales 13 y 14 también presentan el error a los largo del *swath*. Cabe destacar que a partir del análisis de las funciones de pesos mostradas en la figura 2.2, los canales 9,10,11,12,13,14 y de igual modo el 15, tienen poca relación con la precipitación, razón por la cual estos canales no son considerados en el diseño del algoritmo.

Otra característica importante a resaltar en la figura 3.7, es que en el canal 17 y 18 nuevamente aparece el sistema de nubes sobre océano detectado por el canal 2 en la figura 3.5. En las imágenes de los canales 19 y 20 de la figura 3.8 también se aprecia el sistema nuboso sobre el mar Caribe, al nor-este de Guayana Esequiva, a diferencia de los canales 21 y 22 donde no se aprecia, esto permite deducir, que la altura del sistema de nubes formado sobre el caribe tiene un tope de nube que ronda entre los 2 y 4km.

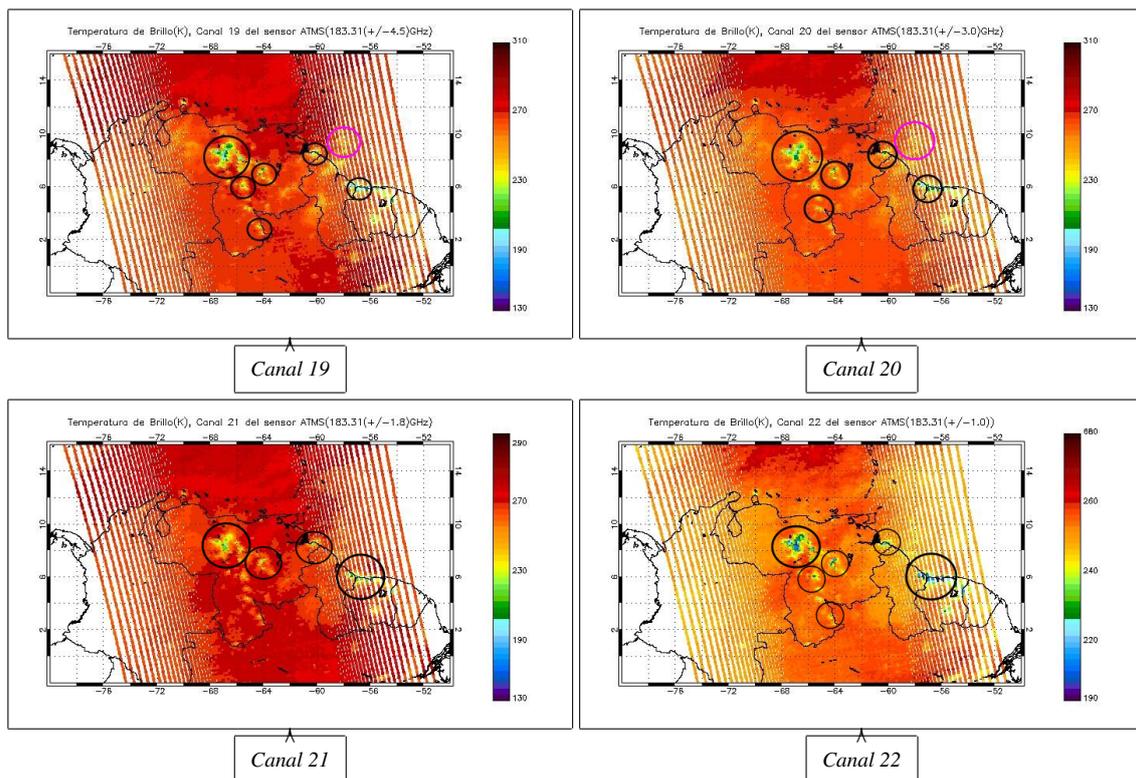


Figura 3.8: Temperatura de Brillo para los canales 19 hasta el 22 del sensor *ATMS*. En los círculos negros se aprecian cristales de hielo asociados a nubes de alto desarrollo vertical, mientras que los púrpuras se relacionan con sistemas de nubes sobre océano

3.1.3. Selección de Canales *ATMS*

El proceso de selección de los canales para el diseño del algoritmo de estimación de precipitación objeto de ésta investigación está limitado por tres aspectos, el primero es por la condición de pertenencia de los píxeles a tierra u océano, en virtud de las razones expuestas en el capítulo 2, en cuanto al principio de emisividad de la atmósfera y su relación con la emisividad de la superficie debajo de ella.

El segundo aspecto está relacionado con las funciones de peso, las cuales proporcionan detalles sobre la sensibilidad asociada a cada canal para observar diferentes características de la atmósfera, y el tercer aspecto hace referencia a la ausencia del error del gradiente de Temperatura de brillo en el *swath*. En este contexto se presenta la metodología planteada en este trabajo para hacer la selección de canales a ser combinados para el entrenamiento de las redes neuronales, tanto para el algoritmo de tierra como para el algoritmo de océano respectivamente.

Fase Previa de Eliminación de Canales

Con el objeto de facilitar el proceso de selección de canales para el proceso de entrenamiento de las redes neuronales, a continuación en la tabla 3.1, se muestran con un **NO** aquellos canales que de antemano son descartados para el diseño del algoritmo bien sea porque sus funciones de peso muestran poca sensibilidad a los procesos que están relacionados con la precipitación, o al hecho de que el canal requiere una fase de corrección antes de poder ser usados. Por otra parte en la figura 3.1 se muestran con **SI** los canales que son considerados a partir de esta primera fase de eliminación.

Es importante mencionar, que el hecho de que el canal no haya sido descartado, no implica que vaya a ser empleado en el algoritmo, su empleo en el algoritmo depende de varios factores, entre los cuales cabe destacar que muestre alta correlación con la variable precipitación, baja correlación con el resto de los canales, y también que al momento de evaluar las salidas de precipitación del algoritmo dicho canal no sea una fuente de ruidos en el producto obtenido. Adicionalmente, se puede comentar que un criterio esencial para

que un canal termine siendo implementado en el algoritmo, el mismo debe tener sensibilidad considerable para en combinación con otros canales predecir la variable de precipitación.

<i>Canal</i>	<i>Desición</i>	<i>Motivo</i>
1	NO	Alta sensibilidad a la superficie
2	NO	Alta sensibilidad a la superficie
3	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
4	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
5	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
6	NO	Alta dependencia con el ángulo de escaneo
7	NO	Alta dependencia con el ángulo de escaneo
8	NO	Alta dependencia con el ángulo de escaneo
9	NO	Alta dependencia con el ángulo de escaneo
10	NO	Relacionado con la absorción de O_2 de la estratósfera
11	NO	Relacionado con la absorción de O_2 de la estratósfera
12	NO	Relacionado con la absorción de O_2 de la estratósfera
13	NO	Relacionado con la absorción de O_2 de la estratósfera
14	NO	Relacionado con la absorción de O_2 de la estratósfera
15	NO	Relacionado con la absorción de O_2 de la estratósfera
16	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
17	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
18	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
19	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
20	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
21	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo
22	SI	Relación con precipitación, moderada dependencia al ángulo de escaneo

Cuadro 3.1: Primera fase de selección de canales del sensor *ATMS*, en la misma son descartados aquellos canales que requieren de un proceso de corrección antes de poder ser usados, y aquellos canales cuya información aporta poca información relacionada con el proceso de precipitación, en base a las funciones de peso. En este sentido una desición SI implica que el canal pudiera ser usado en el algoritmo, mientras que un NO indica que el uso de dicho canal es descartado

A partir de la tabla 3.1, se tiene un conjunto de canales conformado por un total de 10 canales disponibles para el diseño del algoritmo de estimación de tasa de precipitación, integrado por los canales 3, 4, 5, 16, 17, 18, 19, 20, 21 y 22. En el capítulo 4 se reportan resultados relacionados con las combinaciones utilizadas en los diferentes entrenamientos de redes neuronales así como la evaluación de los resultados obtenidos.

3.2. Datos de Entrenamiento

Los datos usados para el entrenamiento de las redes neuronales fueron simulados siguiendo la metodología presentada en la guía de usuarios del RTTOV v10 para la simulación de radianzas debido a la dispersión de microondas por nubes y precipitación (Hocking et al., 2012). El RTTVO v10 es un modelo de transferencia radiativa impulsado en el contexto de las aplicaciones y facilidades de los satélites EUMETSAT para la Predicción Numérica del Tiempo (NWP SAF, por sus siglas en inglés).

Este modelo contempla una interfaz independiente conocida como RTTOV-SCATT, que es proporcionada para simular radianzas de microondas afectadas por nubes y precipitación. Los efectos de dispersión asociados a hidrometeoros a frecuencias de microondas son calculados a partir de la aproximación delta-Eddington, y por otra parte RTTOV-SCATTER contiene 2 conjuntos de perfiles atmosféricos para escenarios de transferencia radiativa en cielo descubierto y escenarios de nubes y precipitación para determinar la temperatura de brillo de la atmósfera en frecuencias de microondas. Los análisis de validación de este modelo están reportados por Saunders et al. (2006).

3.3. Casos de Estudio para el Análisis Comparativo de la Salida de Tasa de Precipitación del Algoritmo

Con la finalidad de establecer un análisis comparativo de las salidas del algoritmo de estimación de tasa de precipitación, en aras de evaluar la pertinencia de los resultados arro-

gados, se ha realizado una selección de casos de precipitaciones severas que han tenido lugar sobre los Estados Unidos durante el primer semestre de 2013. El objeto del estudio de estos casos es evaluar el comportamiento del algoritmo de estimación de tasas de precipitación en escenarios de precipitación severa, en este sentido los eventos propuestos en este trabajo son de magnitudes considerables, a tal punto que en todos los casos se han reportado inundaciones y alto impacto en la población. A continuación se mencionan los casos de estudio seleccionados:

1. CASO A.1: Precipitación Severa sobre los estados de Kansas, Texas, Misuri y Oklahoma de los Estados Unidos
2. CASO A.2: Precipitación Severa sobre los estados de Kansas, Illinois, Misuri y Oklahoma de los Estados Unidos
3. CASO B.1: Precipitación sobre Estados Unidos el 09 de Marzo de 2013
4. CASO B.2: Precipitación sobre Estados Unidos el 10 de Marzo de 2013
5. CASO C.1: Precipitación sobre Estados Unidos el 08 de Mayo de 2013
6. CASO C.2: Precipitación sobre Estados Unidos el 09 de Mayo de 2013

CAPÍTULO 4

RESULTADOS Y ANÁLISIS

En este capítulo se reportan los resultados mas relevantes, pertinentes a las diferentes etapas que comprenden esta investigación. Los mismos incluyen la etapa de selección de canales para los procesos de entrenamientos de las *ANNs*, resultados relacionados con el entrenamiento propio de las *ANNs*, así como los resultados obtenidos para los eventos seleccionados para la evaluación de las estimaciones realizadas por el Algoritmo.

4.1. Selección de canales para el entrenamiento de las *ANNs*

Con la finalidad de seleccionar los canales del sensor *ATMS* para el entrenamiento de las redes neuronales que fundamentan el algoritmo de estimación de tasa de precipitación, a continuación se proporciona una disertación con el fin de justificar las diferentes combinaciones de canales a ser considerados. En este sentido, se menciona que los canales opacos (18, 19, 20, 21 y 22) son incluidos en todos los entrenamientos de redes neuronales que

tengan lugar con el fin de obtener una combinación de canales que resulte óptima.

Ahora bien, es también necesario seleccionar entre los canales restantes candidatos que en conjunto con estos 5 canales opacos permitan obtener valores óptimos de correlación entre la salida de precipitación del algoritmo con los datos de validación simulados en la etapa de post-entrenamiento de la red neuronal. Esta selección es llevada a cabo en forma separada para el algoritmo de estimación de tasa de precipitación sobre pixeles de tierra y océano respectivamente.

4.1.1. Combinación de Canales para el Algoritmo de Estimación de Tasa de Precipitación sobre pixeles de superficie en tierra

En el caso de tierra, en el capítulo 2 se mencionó que la principal desventaja del uso de canales ventana, con frecuencia de operación superior a los 22GHz para estimar precipitación, es el hecho de que la alta emisividad de la superficie terrestre dificulta la observación y distinción de la emisión de radiación por parte de las gotas de lluvia, así como también por las nubes precipitables.

En este sentido, considerando que los datos simulados son consecuencia del modelado radiativo de la atmósfera para diferentes condiciones asociadas a un evento de precipitación, una primera hipótesis para la construcción de combinaciones de canales, es el hecho de que la precipitación es una variable que resulta de procesos en cadena que tienen lugar en los diferentes niveles de la troposfera, de tal forma que se espera, que todos los canales pre-seleccionados tengan una correlación apreciable con la tasa de precipitación medida en tierra.

Entonces como consecuencia del análisis planteado en el párrafo anterior, el primer experimento de entrenamiento de red neuronal involucra el usar como datos de entrada los 10 canales preseleccionados a partir de la tabla 3.1 del capítulo 3, es decir los canales del 3 al 5 y del 16 al 22 del sensor *ATMS*. Por otra parte como caso extremo, se propone que el segundo experimento sea llevado a cabo considerando exclusivamente los canales opacos,

es decir los canales 18, 19, 20, 21 y 22. El tercer experimento, tiene como objeto suprimir los canales 3, 4 y 16, de modo tal que el conjunto de datos de entrada a la red neuronal este conformada por los canales 5, 17, 18, 19, 20, 21 y 22, esto con el fin de evitar la información sobre la superficie y las capas bajas de la atmósfera que contienen estos 3 canales.

A modo de resumen, en el cuadro 4.1 se aprecia el conjunto de experimentos diseñados para el proceso de entrenamiento, cada uno tiene asociado un número de referencia, un código para facilitar su identificación durante el entrenamiento de la red neuronal, y finalmente los canales seleccionados como datos de entrada.

<i>Número de Referencia</i>	<i>Código del Experimento</i>	<i>Combinación de Canales</i>
1	LS_1	3-4-5-16-17-18-19-20-21-22
2	LS_2	18-19-20-21-22
3	LS_3	5-17-18-19-20-21-22

Cuadro 4.1: Diseño de Experimentos para el entrenamiento de la ANN para el algoritmo de estimación de tasa de precipitación sobre tierra en periodo de Verano (el **LS** son siglas en inglés que hacen referencia a Land-Summer, lo cual se interpreta como dato para pixel sobre tierra en época de Verano)

<i>Número de Referencia</i>	<i>Código del Experimento</i>	<i>Combinación de Canales</i>
1	LW_1	3-4-5-16-17-18-19-20-21-22
2	LW_2	18-19-20-21-22
3	LW_3	5-17-18-19-20-21-22

Cuadro 4.2: Diseño de Experimentos para el entrenamiento de la ANN para el algoritmo de estimación de tasa de precipitación sobre tierra en periodo de Invierno (el **LW** son siglas en inglés que hacen referencia a Land-Winter, lo cual se interpreta como dato para pixel sobre tierra en época de Invierno)

4.1.2. Combinación de Canales para el Algoritmo de Estimación de Tasa de Precipitación sobre pixeles de Océano

Siguiendo el criterio utilizado en la sección anterior, el primer experimento para la evaluación de combinación de canales de entrada al algoritmo de estimación de tasa de precipitación sobre océano involucra inicialmente los 10 canales preseleccionados en la tabla 3.1 del capítulo 3. De igual modo, el segundo experimento está conformado por los 5 canales opacos, en virtud de evaluar la capacidad de estimación de tasa de precipitación del algoritmo, en ausencia de canales ventana o de transmitancia intermedia. Y análogamente en el tercer experimento se excluyen aquellos canales considerados ventana de tal modo que el conjunto de datos de entrada está conformado por los canales 5, 17, 18, 19, 20, 21 y 22. En el cuadro 4.3, se presenta de forma resumida la información proporcionada en este párrafo.

<i>Número de Referencia</i>	<i>Código del Experimento</i>	<i>Combinación de Canales</i>
1	OS_1	3-4-5-16-17-18-19-20-21-22
2	OS_2	18-19-20-21-22
3	OS_3	5-17-18-19-20-21-22

Cuadro 4.3: Diseño de Experimentos para el entrenamiento de la ANN para el algoritmo de estimación de tasa de precipitación sobre océano en periodo de Verano (el **OS** son siglas en inglés que hacen referencia a Ocean-Summer, lo cual se interpreta como dato para pixel sobre océano en época de Verano)

<i>Número de Referencia</i>	<i>Código del Experimento</i>	<i>Combinación de Canales</i>
1	OW_1	3-4-5-16-17-18-19-20-21-22
2	OW_2	18-19-20-21-22
3	OW_3	5-17-18-19-20-21-22

Cuadro 4.4: Diseño de Experimentos para el entrenamiento de la ANN para el algoritmo de estimación de tasa de precipitación sobre océano en periodo de Invierno (el **OW** son siglas en inglés que hacen referencia a Ocean-Winter, lo cual se interpreta como dato para pixel sobre océano en época de Invierno)

4.2. Proceso de Entrenamiento de las *ANNs*

Tal y como se mencionó en el capítulo 3, el proceso de entrenamiento de las *ANNs*, es implementado mediante dos librerías distintas programadas en C++¹, en forma paralela para cada combinación de canales. En este sentido es importante señalar que se presentan los resultados arrojados por cada librería de forma simultanea para facilitar los procesos de evaluación de calidad. A continuación se presentan los resultados obtenidos a partir de las fases de entrenamiento de las diferentes redes neuronales para los 4 escenarios planteados en esta investigación, los cuales son:

1. Estimación de Precipitación sobre píxeles de tierra en época de verano
2. Estimación de Precipitación sobre píxeles de océano en época de verano
3. Estimación de Precipitación sobre píxeles de tierra en época de invierno
4. Estimación de Precipitación sobre píxeles de océano en época de invierno

4.2.1. Entrenamiento de la Red Neuronal para la estimación de precipitaciones sobre pixeles de tierra en época de verano

El primer conjunto de entrenamiento de redes llevado a cabo es para el algoritmo de determinación de tasa de precipitación sobre píxeles de tierra en época de verano, el conjunto de datos simulados usados para este entrenamiento, involucra 25384 casos de combinaciones de los canales *ATMS* asociados con diferentes niveles de tasa de precipitación que comprenden valores entre $6,30 \times 10^{-4}$ y 61,43mm/h.

Los resultados de entrenamiento de redes, obtenidos para este algoritmo, se muestran en la figura 4.1 donde se han representado del lado izquierdo los resultados derivados a partir de la librería *ALGLIB* y del lado derecho los obtenidos mediante la librería *Open NN*, en todos los casos en el eje de las y, etiquetado con la leyenda *NN(mm/h)*, se presentan los

¹Open NN y ALGLIB

valores de tasa de precipitación estimados por la red neuronal, mientras que con la etiqueta *Database(mm/h)*, en el eje x se presenta los valores de tasa de precipitación esperados, los cuales corresponden a la tasa de precipitación calibrada para los datos de temperatura de brillo de los canales del sensor **ATMS**, que han sido usados para el entrenamiento de las redes neuronales², así como también para los datos de entrada de los experimentos planteados en la figura mencionada.

En este sentido, en las subfiguras *a* y *b* se aprecian los diagramas de dispersión para el primer experimento planteado en la tabla 4.1, es decir el experimento *LS_1*. La primera observación a destacar es que *ALGLIB* proporciona un coeficiente de correlación³ que resulta un 21,41 % más alto que *Open NN*, y este resultado puede ser observado de forma visual a partir de la regresión lineal y la línea de ajuste mostrada en cada caso, de donde resulta relevante la calidad del ajuste obtenido a partir de *ALGLIB*.

En este experimento, *ALGLIB* reportó un valor de 2,59 para el error cuadrático medio, mientras que por parte de *Open NN* la figura 4.2 muestra la diferencia de errores, durante en proceso de entrenamiento, entre los datos de salida de la red y los datos de prueba. Idealmente, en la figura 4.2 es de esperarse dos curvas superpuestas, lo cual indica que el error entre la salida de la red (curva azul) en entrenamiento reproduce perfectamente los datos de prueba (curva roja), sin embargo, en lugar de esto se tiene entonces que cuando la curva asociada a los datos de prueba (Generalization) se aproxima a los datos de salida de la red en función de las épocas se tiene un proceso de entrenamiento normal, ahora bien cuando esta curva de prueba se aleja por arriba de la curva azul (Performance), como sucede en la figura 4.2 a partir de la época 1000, se dice que la red está sobre-entrenando lo cual se traduce en pérdida de generalidad por parte del proceso de entrenamiento⁴.

²Para esta evaluación se ha usado el 100 % de los datos simulados que sirvieron de entrada para el entrenamiento de las redes neuronales

³Reportado en la figura como **R-squared**

⁴Es importante destacar que esta curva no es necesario realizarla para *ALGLIB*, debido a que el criterio de parada en esta librería es automático, y no requiere de la evaluación de ciclos de entrenamiento

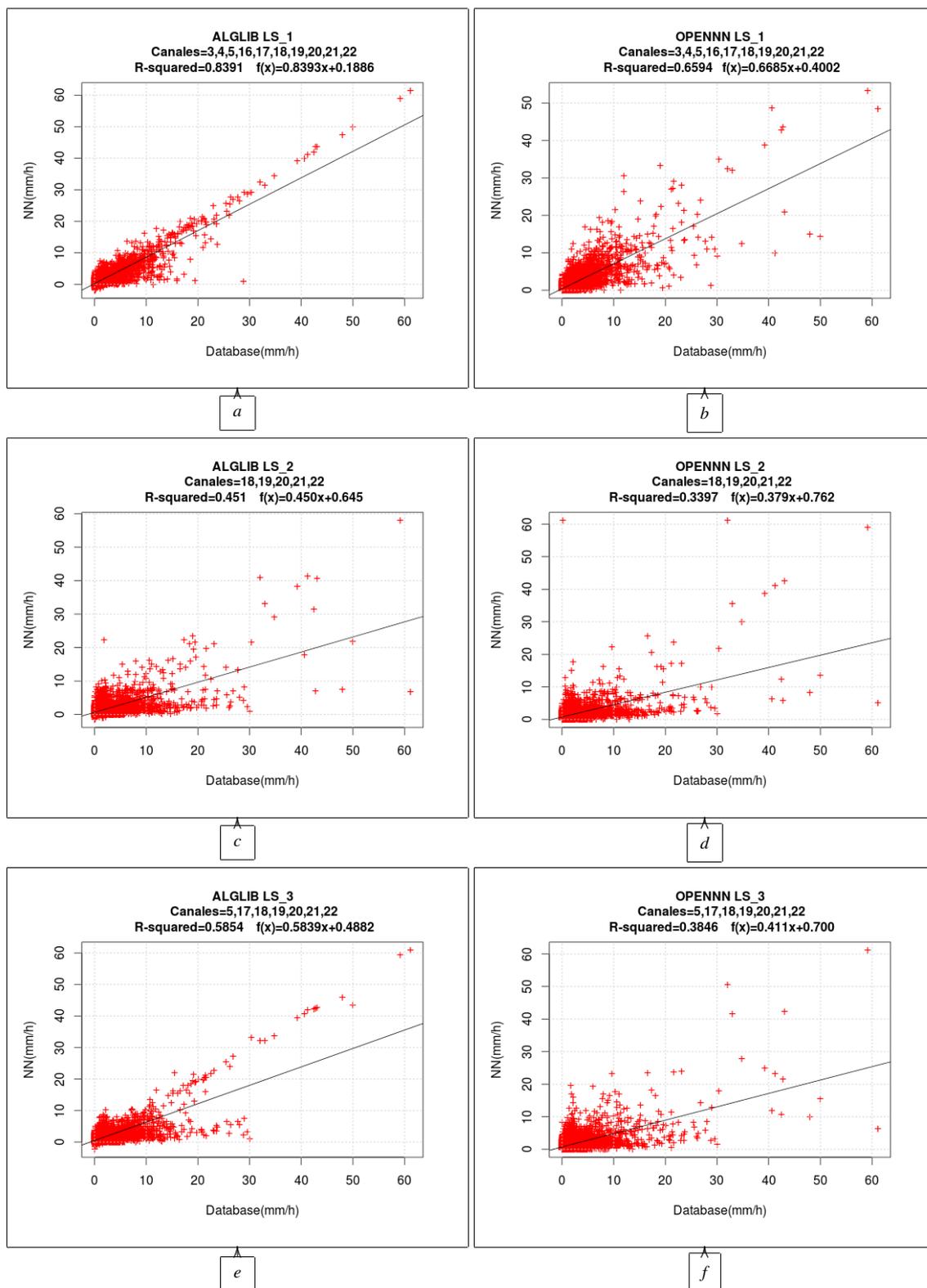


Figura 4.1: Comparación de la salida de la red neuronal con los datos de entrenamiento para la estimación de precipitación sobre pixeles de tierra en época de verano

Por otro lado, cuando la curva roja se aleja de la azul por debajo, en este caso el resultado es un sub-entrenamiento de la red. En todo caso es necesario evitar tanto el sub-entrenamiento, así como el sobre-entrenamiento de la red, en orden de llevar a cabo un buen proceso de entrenamiento. Particularmente, en este caso se tomó como pesos y bias de la red, los valores correspondientes a 1000 épocas.

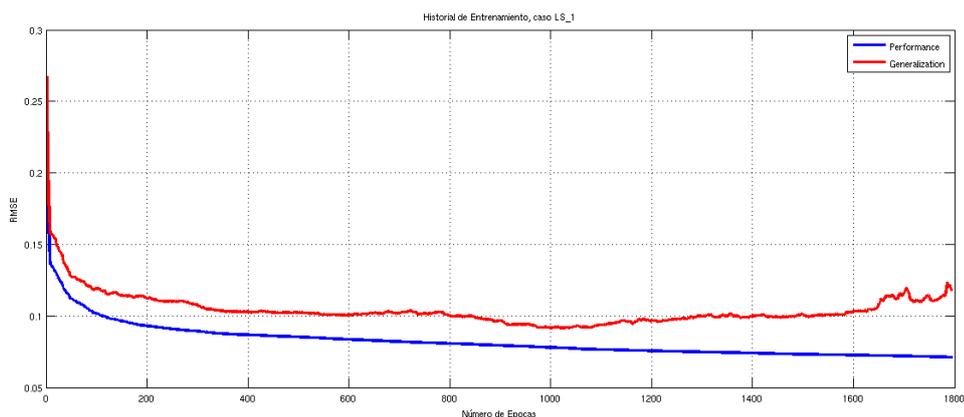


Figura 4.2: Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en *Open NN* del caso LS_1.

Ahora bien, con la finalidad de obtener resultados óptimos para los procesos de entrenamiento llevados a cabo con *Open NN*, se optó por establecer el criterio de parada a partir de la evaluación del proceso de entrenamiento, evitando sub-entrenamiento y sobre-entrenamiento durante los ciclos del proceso. Figuras del tipo 4.2, proporcionan la información suficiente para discriminar cuando el proceso de entrenamiento es óptimo, en términos de la evolución en ciclos del proceso, y por lo tanto permite definir el número del ciclo para el que se considera la red entrenada⁵.

Por otra parte, el criterio de parada del *ALGLIB* es automático, y entre sus ventajas resalta que incorpora un algoritmo que garantiza la optimización del proceso de entrenamiento

⁵La herramienta para elaborar esta gráfica no está disponible en *Open NN*, para lograr este producto fue necesaria la modificación de algunas clases en el software, para poder almacenar el historial de los parámetros de entrenamiento en cada ciclo. Esto no fue posible realizarlo en *ALGLIB* por parte del autor, ya que no resultó evidente la clase y método a modificar

a través de la evaluación del costo de iteraciones fijas. Debido a esto, en el caso de *ALGLIB*, no es necesario realizar el análisis por ciclos de entrenamiento, como se ha hecho para *Open NN* a partir de la figura 4.2.

Continuando con los experimentos de la tabla 4.1, las subfiguras *c* y *d* de la figura 4.1 corresponden al segundo experimento, el cual se ha etiquetado como *LS_2*. En este caso el coeficiente de correlación obtenido por *ALGLIB* es 24,67 % superior al de *Open NN*, y se observa mejor ajuste por parte de *ALGLIB* que además reportó para este caso un error cuadrático medio de 2,76, mientras que el ciclo óptimo considerado por *Open NN* resultó ser el 1220, cuyo criterio de selección puede deducirse a partir de la figura 4.3.

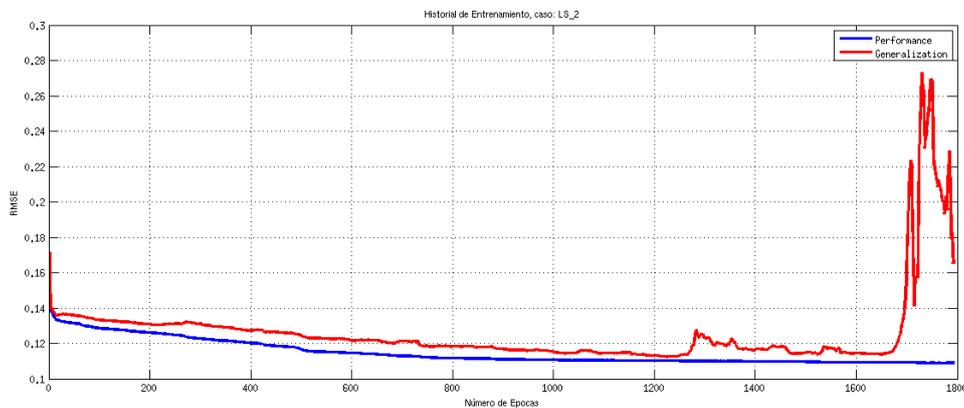


Figura 4.3: Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en *Open NN* del caso *LS_2*.

Respecto a la calidad de este resultado, se puede agregar que una correlación de 0.451 es baja y más aun considerando que esta correlación describe la capacidad de predicción de datos simulados y no de datos reales, en cuyo caso se espera disminuya aun mas dicha correlación. Este resultado se atribuye al hecho de que estos 5 canales opacos, por si solos no contienen suficiente información como para estimar precipitación debajo de las nubes de alto desarrollo vertical, ya que por la fuerte sensibilidad que poseen a los cristales de hielo en las nubes, resulta invisible la precipitación que tiene lugar debajo de estos sistemas convectivos. Por tal motivo, la inclusión de canales que permitan obtener información por debajo de las nubes convectivas, como lo son el caso de los canales 3,4 y 5, que son sensibles

a los cristales de hielo en menor medida, son una excelente opción en virtud de aumentar los niveles de predicción.

Ahora bien, en el caso del tercer experimento, etiquetado como *LS_3*, *ALGLIB* arrojó un coeficiente de correlación 34,30 % mayor que *Open NN*, lo cual hace referencia a un mejor ajuste entre las tasa de precipitación predicha por la red neuronal y los datos simulados de entrenamiento, por parte de *ALGLIB*. El error medio cuadrático reportado por *ALGLIB* es de 2,35, mientras que el número de épocas requerido por *Open NN* fue de 850. En la figura 4.4 se muestran las curvas de errores medios cuadráticos durante la fase de entrenamiento.

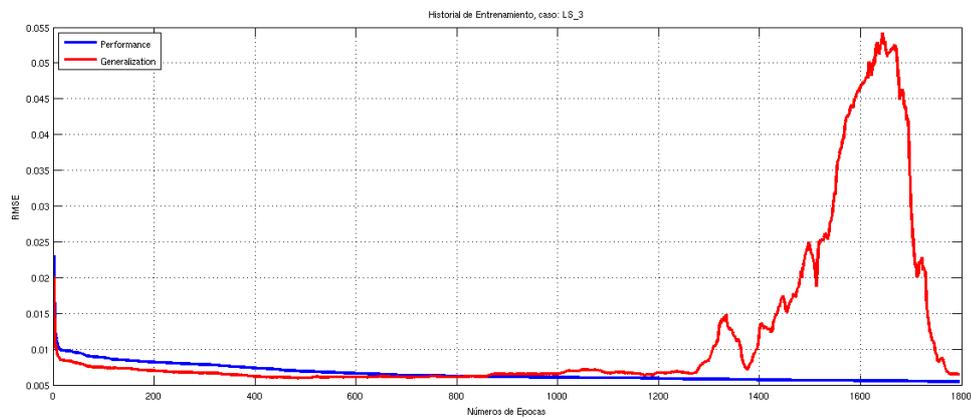


Figura 4.4: Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en *Open NN* del caso *LS_3*.

En el cuadro 4.5, se presenta un resumen de los resultados mas relevantes reportados en esta sección, donde resalta que en todos los casos la red entrenada mediante *ALGLIB* proporcionó mejor predicción que aquellas entrenadas con *Open NN*, por otro lado la mejor combinación de canales en cuanto a niveles de correlación con los datos de entrenamiento resulto ser el experimento *LS_1*.

<i>Experimento</i>	R^2 con <i>ALGLIB</i>	R^2 con <i>Open NN</i>	<i>Diferencia Porcentual(%)</i>
LS_1	0.8391	0.6594	21.41
LS_2	0.4510	0.3397	24.67
LS_3	0.5854	0.3846	34.30

Cuadro 4.5: Resumen de resultados de coeficientes de correlación para las redes neuronales de estimación de tasa de precipitación sobre píxeles de tierra en época de verano

4.2.2. Entrenamiento de la Red Neuronal para la estimación de precipitaciones sobre pixeles de océano en época de verano

Las combinaciones de canales para el entrenamiento de redes neuronales para la estimación de tasa de precipitación sobre píxeles de océano en época de verano se pueden apreciar en la tabla 4.3, donde se muestran los tres experimentos diseñados. Los resultados obtenidos a partir de cada una de las tres redes neuronales entrenadas, para un conjunto de datos de entrenamiento compuesto por 9006 casos, se aprecian en la figura 4.5 donde nuevamente se pueden apreciar resultados tanto para los entrenamientos llevados a cabo mediante *ALGLIB*, así como también por *Open NN*.

En este sentido, las subfiguras *a* y *b* de la figura 4.5 muestran los resultados del experimento *OS_1* para el entrenamiento con *ALGLIB* y *Open NN* respectivamente. En la figura se aprecia un coeficiente de correlación de 0.9066 para las salidas de la red entrenada por *ALGLIB*, lo cual representa un 19,18 % de mayor correlación respecto al 0.7327 reportado por *Open NN*. Por parte de *ALGLIB*, se obtuvo un error cuadrático medio de 3.06, mientras que el historial de error cuadrático medio para el entrenamiento con *Open NN* se aprecia en la figura 4.6, de donde se tiene que la época 600 es óptima para detener el entrenamiento.

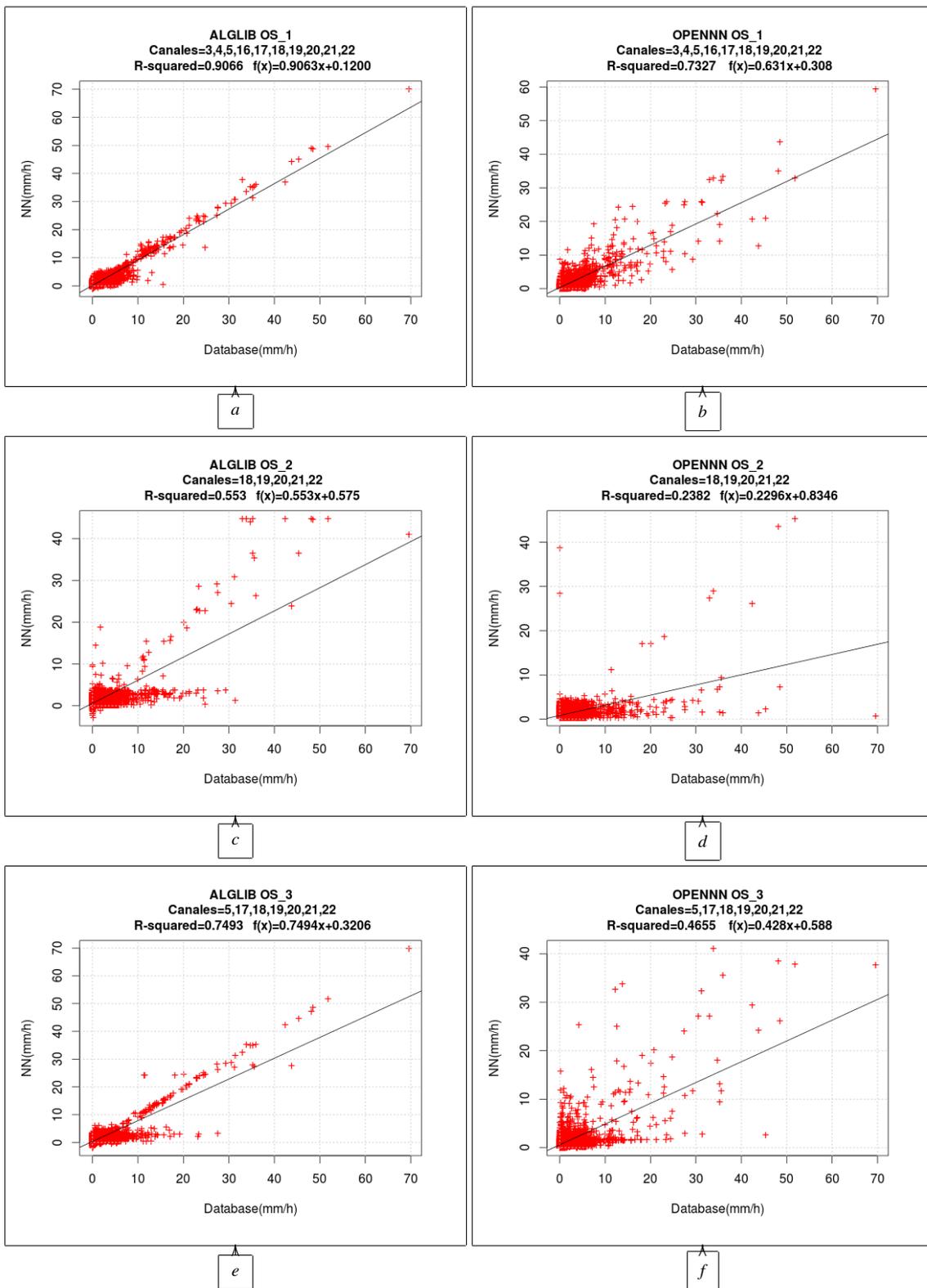


Figura 4.5: Comparación de la salida de la red neuronal con los datos de entrenamiento para la estimación de precipitación sobre pixeles de océano en época de verano.

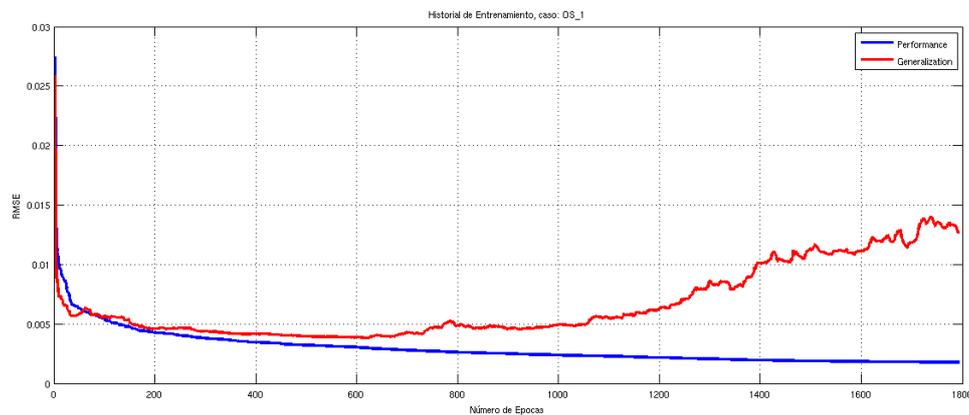


Figura 4.6: Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en *Open NN* del caso OS_1.

De igual modo, en las subfiguras *c* y *d* de la figura 4.5, se muestran los resultados para el experimento OS_2, en la cual la diferencia en coeficiente de correlación es del 56,92 % a favor de la red entrenada mediante *ALGLIB* con un coeficiente de correlación de 0.5530 delante de la red entrenada mediante *Open NN*, de la cual se obtuvo un coeficiente de correlación de 0.2382. En cuanto al error cuadrático medio se tiene que por parte de *ALGLIB* se obtuvo 3.34986, mientras que la figura 4.7 permite apreciar el error cuadrático medio durante la fase de entrenamiento de la red, de donde se deriva que el número de ciclos necesarios para el entrenamiento es 800.

Por otra parte, los resultados obtenidos para el experimento OS_3 se muestran en las subfiguras *e* y *f* de la figura 4.5, donde se aprecia que el coeficiente de correlación derivado a partir del entrenamiento de *ALGLIB* es de 0.7493, el cual es 37.87 % mayor que el obtenido por el entrenamiento con *Open NN* que resultó ser 0.46. La figura 4.8 muestra el error cuadrático medio del proceso de entrenamiento por *Open NN*, donde se deduce en base al comportamiento de el desempeño de entrenamiento y al error asociado a los datos de prueba, que el número de épocas óptimo para el entrenamiento de la red es 700, y por otro lado el error cuadrático medio reportado por *ALGLIB* es 2.98.

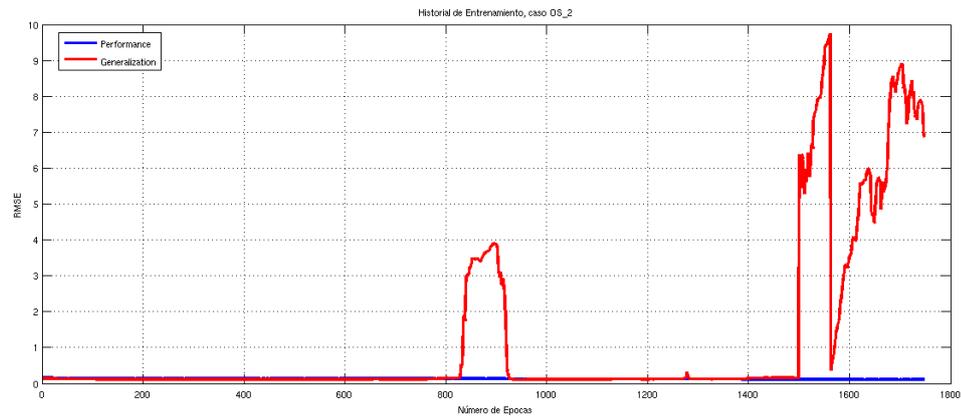


Figura 4.7: Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en *Open NN* del caso OS_2.

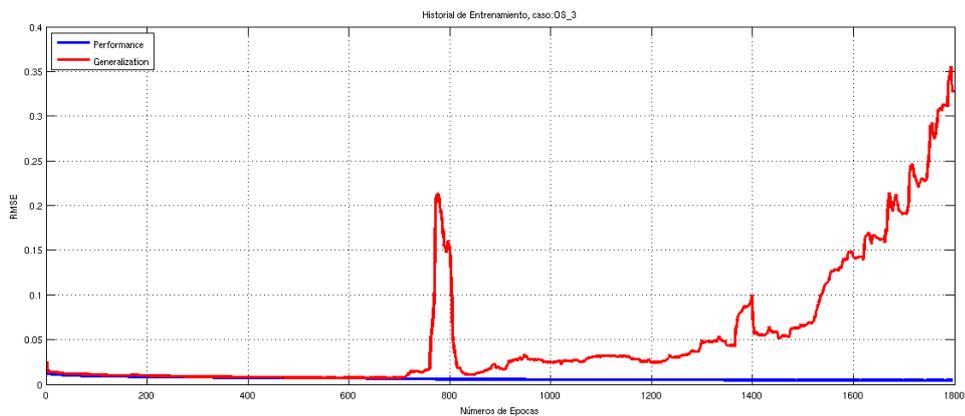


Figura 4.8: Error Cuadrático Medio del desempeño del proceso de entrenamiento (azul) y de los datos de prueba (rojo) durante el proceso de entrenamiento de la red en *Open NN* del caso OS_3.

En el cuadro 4.6 se presenta un resumen de los resultados obtenidos en esta sección, entre los cuales cabe destacar que nuevamente las redes neuronales entrenadas mediante la librería *ALGLIB* mostraron mayor ajuste con los datos de prueba que aquellas entrenadas por *Open NN*, también es necesario resaltar que el experimento *OS_1* mostró la correlación mas alta.

<i>Experimento</i>	R^2 con <i>ALGLIB</i>	R^2 con <i>Open NN</i>	<i>Diferencia Porcentual(%)</i>
OS_1	0.9066	0.7327	19.18
OS_2	0.5530	0.2382	56.92
OS_3	0.7493	0.4655	37.87

Cuadro 4.6: Resumen de resultados de coeficientes de correlación para las redes neuronales de estimación de tasa de precipitación sobre píxeles de océano en época de verano

En base a los resultados mostrados en la tabla 4.6 se concluyó que bajo la configuración diseñada para *Open NN*, y habiendo hecho las pruebas necesarias de entrenamiento, se selecciona a *ALGLIB* como la librería para el entrenamiento de las redes neuronales sobre las cuales tiene su fundamento el algoritmo de estimación de precipitación objeto de esta investigación.

4.2.3. Entrenamiento de la Red Neuronal para la estimación de precipitaciones sobre píxeles de tierra en época de invierno

El entrenamiento de redes neuronales para la estimación de precipitaciones sobre píxeles de tierra en época de invierno se llevó a cabo en base a un conjunto de datos simulados conformado por 5029 casos, y siguiendo los experimentos planteados en el cuadro 4.2.

En el experimento *LW_1* se obtuvo un error cuadrático medio de 0,46 y de la figura 4.9 se observa además que se alcanzó un coeficiente de correlación de 0,9136, el cual es notablemente mayor al obtenido para esta misma combinación en condiciones de verano, tal y como se muestra en la subfigura **a** de la figura 4.1. En este orden de ideas, el error cuadrático medio asociado al experimento *LW_2* es de 0,87, mientras que el coeficiente de correlación obtenido es de 0,4473, tal y como se muestra en la figura 4.10, que como se aprecia en la tabla 4.5 es bastante similar al análogo de esta combinación en condiciones de verano. Por otro lado, se aprecia en la figura 4.11 que el coeficiente de correlación obtenido para el experimento *LW_3* es de 0,6818, con un error cuadrático medio calculado en 0,39.

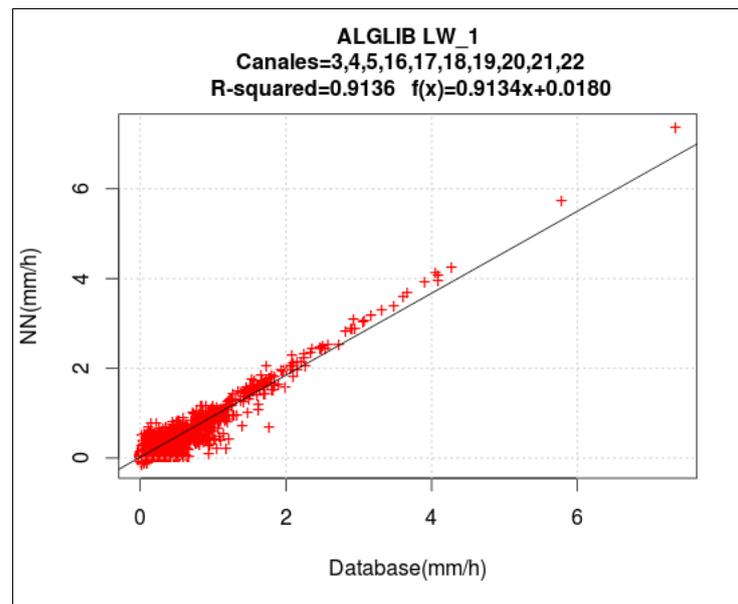


Figura 4.9: Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento *LW_1*

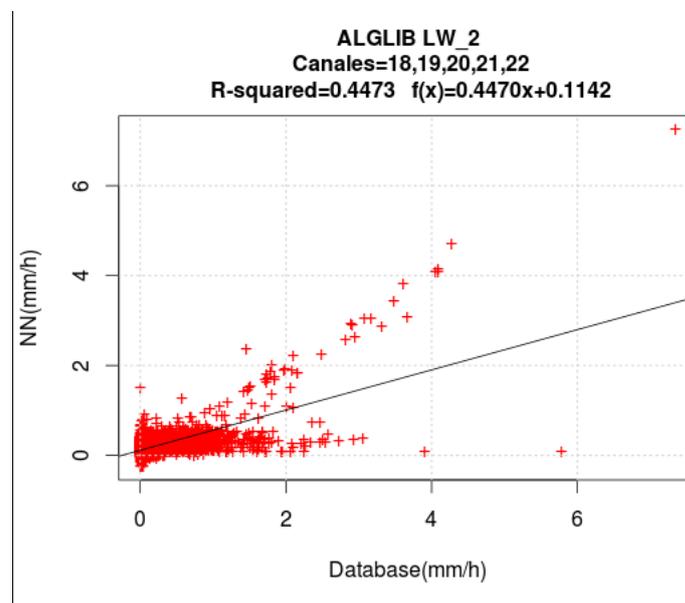


Figura 4.10: Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento *LW_2*

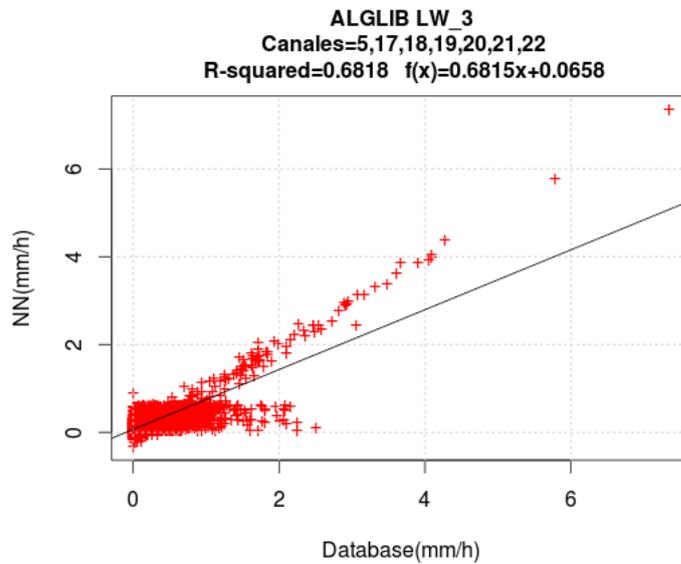


Figura 4.11: Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento *LW_3*

4.2.4. Entrenamiento de la Red Neuronal para la estimación de precipitaciones sobre pixeles de océano en época de invierno

El entrenamiento de redes neuronales para la estimación de precipitaciones sobre píxeles de océano en época de invierno se llevó a cabo en base a un conjunto de datos simulados conformado por 9310 casos, cumpliendo para esto el conjunto de experimentos presentados en el cuadro 4.4.

Entre los resultados obtenidos cabe destacar de la figura 4.12 que se obtuvo un coeficiente de correlación de 0,8986 con un error cuadrático medio de 1,2649 para el experimento *OW_1*, también de la figura 4.13 se muestra un coeficiente de correlación de 0,4060 para el experimento *OW_2*, el cual tiene asociado un error cuadrático medio de 1,6325, de esta figura cabe resaltar la alta dispersión observable, la cual puede atribuirse a la baja predictibilidad de la red neuronal entrenada sobre los datos de prueba simulados. Por otra parte, de la figura 4.11 se tiene el coeficiente de correlación para el experimento *OW_3*, el cual tiene el valor de 0,7451 con un error cuadrático medio asociado de 1,17738.

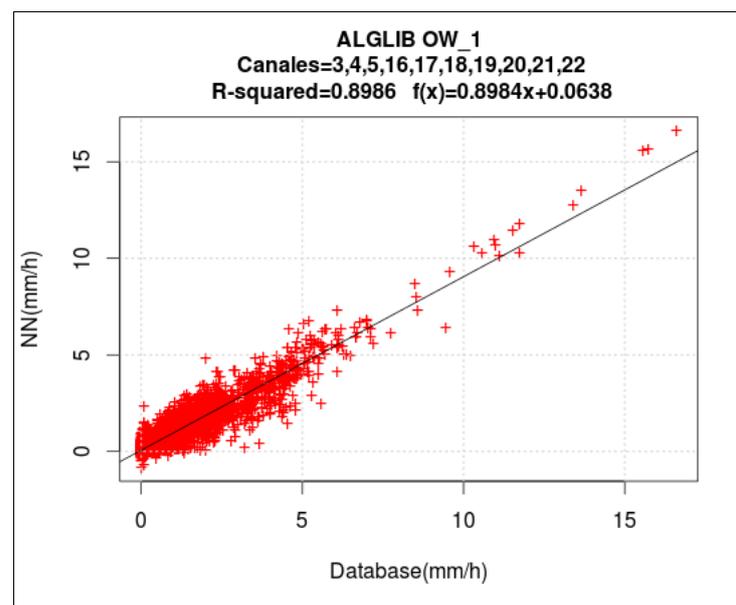


Figura 4.12: Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento *OW_1*

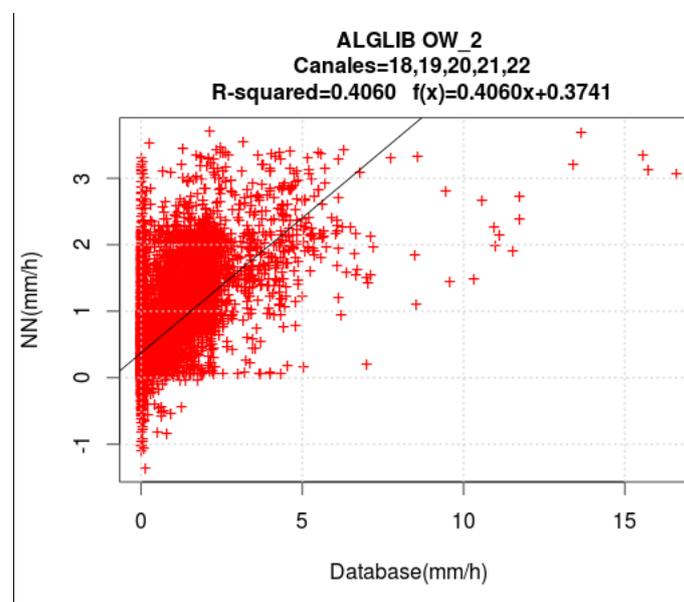


Figura 4.13: Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento *OW_2*

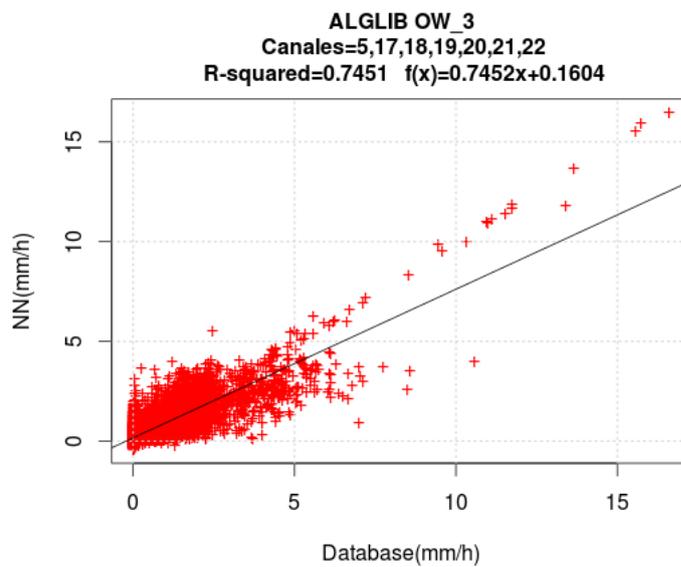


Figura 4.14: Diagrama de Dispersión entre salidas de la red para datos de prueba contra datos de entrenamiento para el experimento *OW_3*

4.2.5. Colofón

A partir de los resultados obtenidos desde los entrenamientos de redes neuronales para las 4 configuraciones planteadas al inicio de este capítulo, en la tabla 4.7 se muestra el conjunto de experimentos que mostraron mejor correlación y por ende predictibilidad de los datos de prueba simulados durante la fase de entrenamiento. Estos experimentos son la primera referencia para hacer los estudios de comparación de calidad de las salidas de tasa de precipitación por parte del algoritmo de estimación.

En este orden de ideas, es importante mencionar que en el caso hipotético de que estas primeras opciones generen mapas de precipitación con ruidos inesperados, se procede a tomar entonces la red neuronal con segunda mejor correlación, con la esperanza de que la nueva combinación de canales no contenga el hipotético ruido. Para mayor comprensión respecto a esta condicionante, puede añadirse que los posibles ruidos que pudieran tener presencia, se esperan para mapas de precipitación sobre píxeles de océano, considerando que los canales 3 y 4 mostraron ciertos niveles de ruido en la figura 3.5, sin embargo como

se acaba de mencionar esta es una posibilidad a ser considerada.

<i>Escenario</i>	<i>Mejor Experimento</i>	<i>Valor de Correlación</i>
LS	LS_1	0,8391
LW	LW_1	0,9136
OS	OS_1	0,9066
OW	OW_1	0,8986

Cuadro 4.7: Mejores resultados de coeficientes de correlación obtenidos por el proceso de entrenamiento de las redes neuronales

Es importante destacar, que el criterio para la selección del mejor experimento, ha sido la escogencia del experimento cuya producción genera mayor coeficiente de correlación lineal entre los valores de tasa de precipitación predichos por la red neuronal y los esperados. En este contexto, es necesario recalcar que la red fue entrenada usando valores de precipitación que han sido calibrados con datos simulados de los 22 canales del sensor **ATMS**, de tal forma que el mejor experimento es aquel para el cual ingresando los mismos datos con lo que fue entrenada la red (conformados por los datos de entrenamiento, de verificación y de prueba), reproduzca los valores de las tasas de precipitación con un coeficiente de correlación igual a 1.

Ahora bien, en ninguno de los casos se obtuvo una correlación perfecta, razón por la cual en cada caso se ha seleccionado el experimento que presente mayor coeficiente de correlación. Es de esperar, que estos valores de correlación disminuyan cuando se ingresen datos de entradas diferentes a los que se usaron para el proceso de entrenamiento, es por esto que se requieren pasos de evaluación del comportamiento de las salidas de las diferentes redes neuronales entrenadas, los cuales son llevados a cabo a partir de la selección de algunos casos de estudios relevantes como los que se muestran a continuación.

4.3. Casos de Estudio

4.3.1. CASO A.1: Precipitación sobre Kansas, Oklahoma, Misuri y noroeste de Texas del 29 de Abril de 2012

En un primer intento por evaluar las características de las salidas del algoritmo de estimación en relación a la ubicación y magnitud del cálculo de tasa de precipitación, se ha llevado a cabo un estudio sobre el caso de precipitación severa que se presentó en los estados de Kansas, Oklahoma, Misuri y Texas de los Estados Unidos el 29 de Abril de 2012. En la figura 4.15, se aprecia la predipitación acumulada para las 24 horas de la fecha, en la misma se puede observar la extensión de los eventos de precipitación observados, así como también el nivel de tasa de precipitación observado durante el día en los estados bajo estudio.

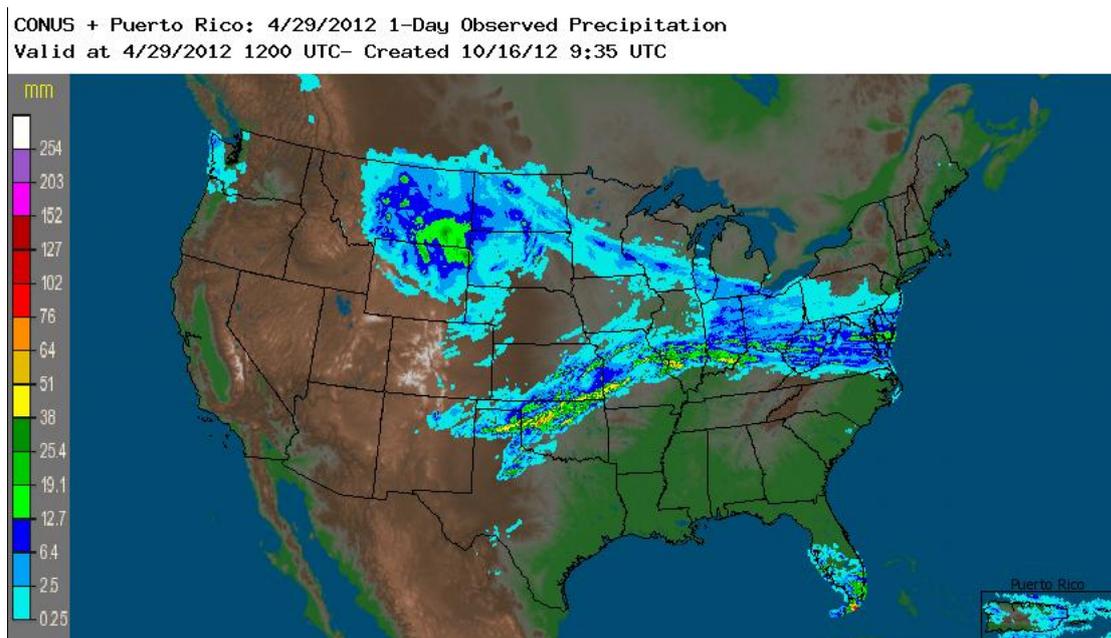


Figura 4.15: Precipitación acumulada para 24 horas del 29 de Abril de 2012 en Estados Unidos. (Fuente:<http://water.weather.gov/precip/>)

En este sentido, la estimación de precipitación del algoritmo es realizada a partir de una escena *ATMS* de la región de estudio capturada entre las 08:24 y 8:32 UTC de la fecha. De esta forma es importante destacar que el producto obtenido a partir del algoritmo es una tasa

de precipitación instantánea, con variaciones temporales en la escena debido a la diferencia de tiempo con que es barrido cada pixel.

En este contexto sería deseable contar con datos de tasas de precipitación instantánea para los efectos de la comparación y evaluación de la salida, motivo por el cual se ha empleado datos del sistema *NEXRAD* para los efectos de comparación de la salida a escala local, y por otra parte se ha hecho uso de datos de precipitación observada del Servicio Nacional de Meteorología (*NWS*) para hacer análisis de comparación sobre la tasa total de precipitación diaria y horaria, como también indagar sobre los patrones de precipitación a escala sinóptica proporcionados por el algoritmo.

En otro orden de ideas, es importante mencionar que si bien, ciertamente el estudio presentado es una evaluación cualitativa, el mismo no representa pérdida de generalidad o rigurosidad, esto considerando que los datos utilizados para dicha evaluación son datos estrictamente validados, y además de la evaluación a escala sinóptica, también son analizados algunos ejemplos a escala local para verificar la correspondencia de la salida del algoritmo con los datos reportados por el *NEXRAD* y el *NWS*.

Ahora bien, el primer parámetro a evaluar es a escala sinóptica, y esto es el patrón de precipitación observado. Con esta finalidad, en la figura 4.16 se muestra el mapa de precipitación observado por el *NWS* para las 09:00 horas UTC, en donde se han resaltado mediante las etiquetas 1, 2, 3 y 4 los focos más importantes de precipitación⁶. Es importante destacar que la tasa de precipitación máxima corresponde para los focos 2 y 3, al observarse valores de 1 pulgada, lo cual es 25,4 mm de precipitación acumulada para el horario comprendido entre las 08:00 y 09:00 UTC. En los focos 2, 3 y 4 hay grandes proporciones de pixeles a los que corresponden precipitaciones entre 0,50 (12,5mm) y 0,75 pulgadas (18,5mm).

⁶Es necesario resaltar que esta imagen es reportada por el *NWS* usando el sistema de medición inglés, por este motivo la precipitación se expresa en pulgadas (Inches), para expresar estos valores en milímetros solo es necesario multiplicar el valor reportado en la barra de colores por 25,4.

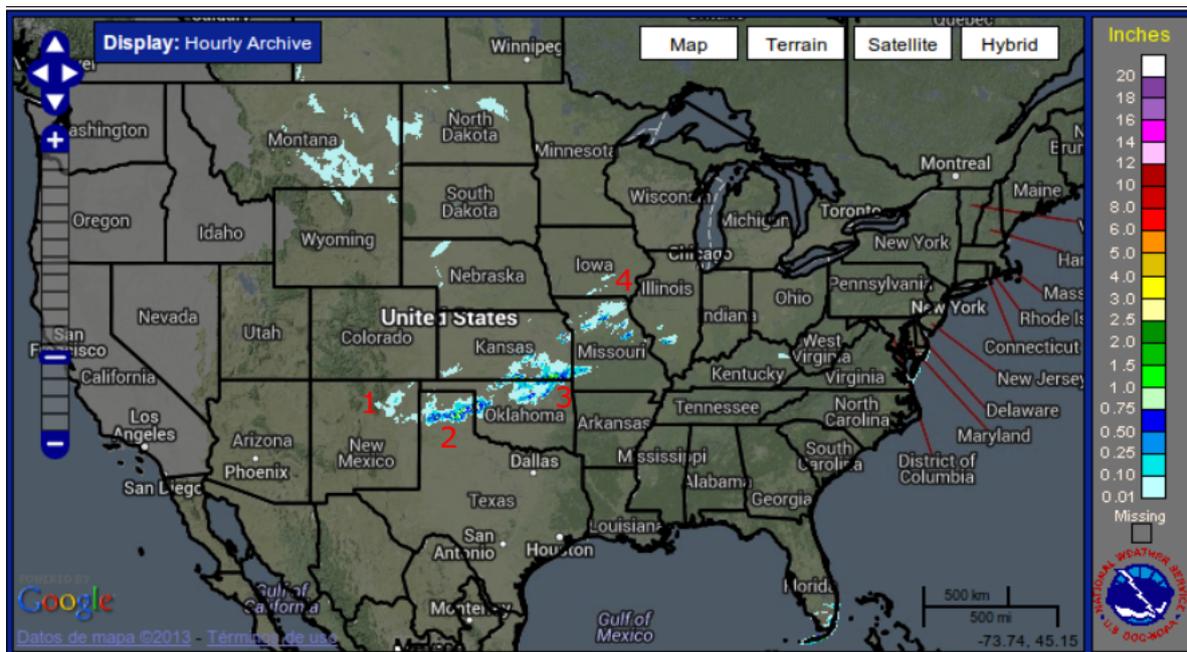


Figura 4.16: Precipitación horaria para las 09:00 horas UTC, del 29 de Abril de 2012 en los estados de Kansas, Oklahoma, Misuri y Texas de los Estados Unidos. (Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)

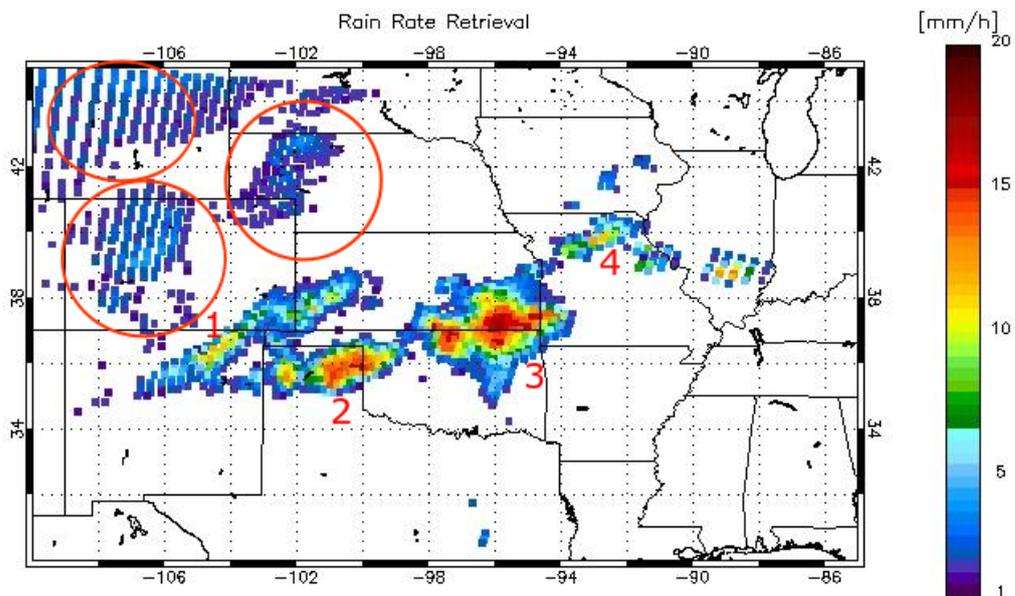


Figura 4.17: Tasa de Precipitación Instantánea, para la escena ATMS barrida entre las 08:24 y 8:32 UTC del 29 de Abril de 2012, estimada mediante la red LS_1

De igual forma la figura 4.17 muestra la tasa de precipitación instantánea estimada por el algoritmo, de donde puede observarse que el algoritmo ha reportado de forma coincidente a la figura 4.16 los 4 focos principales de precipitación con lo cual puede apreciarse que el patrón ha sido bien reproducido por el algoritmo al mostrar un alto nivel de correspondencia. Por otro lado, en la figura 4.17 se aprecian zonas de precipitación encerradas en círculos de perímetros rojos, lo cual corresponde a sobrestimaciones por parte del algoritmo.

Resulta interesante hacer un análisis a escala local de los niveles de precipitación para algunos casos, a partir de la comparación de la tasa de precipitación instantánea mostrada en la figura 4.17, y algunos radares del sistema *NEXRAD*. Por ejemplo La figura 4.18 corresponde al Radar *NEXRAD* denominado *KAMA*, el cual está ubicado al noroeste del estado de Texas. En la figura del radar se pueden identificar, dentro del foco de precipitación etiquetado como número 2, 3 subgrupos denominados en la imagen como A, B y C.

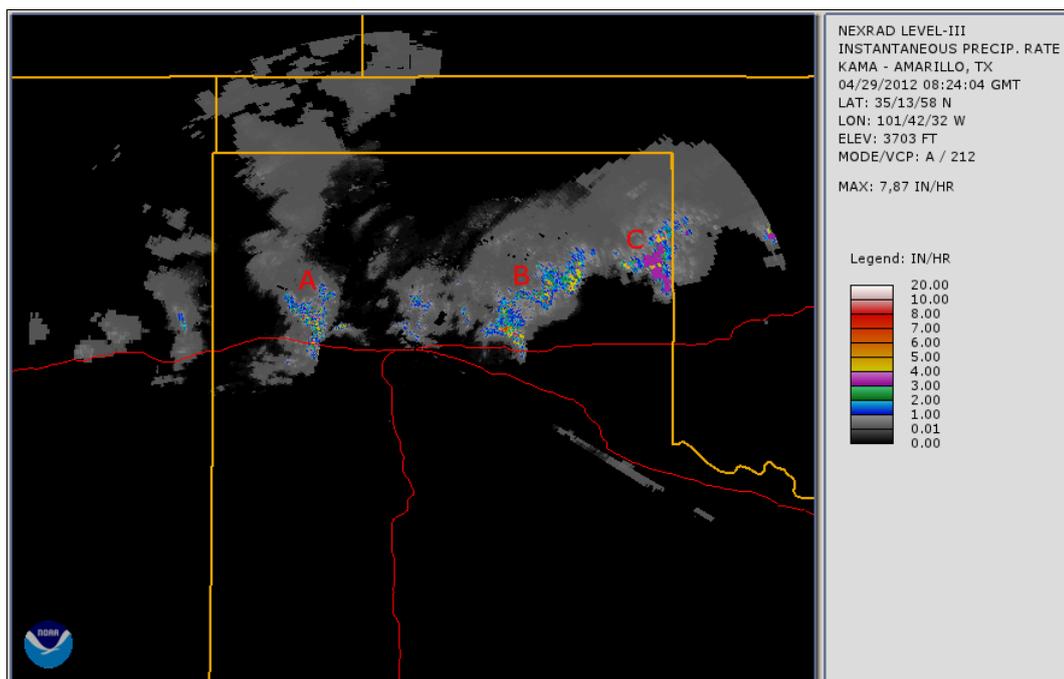


Figura 4.18: Tasa de Precipitación Instantánea sobre el noroeste del estado de Texas para las 08:24 UTC del 29 de Abril de 2012, estimada mediante la red *NEXRAD*

En el grupo A, se observa que el radar reporta para la mayoría de píxeles precipitaciones instantáneas de 1 pulgada (25,4mm) por hora, mientras que el algoritmo para esta región

reporta tasas de precipitación instantánea de aproximadamente 12mm/h. Para el grupo B, los resultados reportados por el radar son similares a los del grupo A, a diferencia de que existen dos concentraciones con tasas de precipitaciones por encima de 100mm/h lo cual es muy superior a los 14mm/h, que es el máximo de precipitación obtenido por el algoritmo. Por otra parte, en el caso del grupo C, el radar estima generalmente una tasa de precipitación superior a los 75mm/h, lo cual nuevamente es superior a la estimación realizada por el algoritmo.

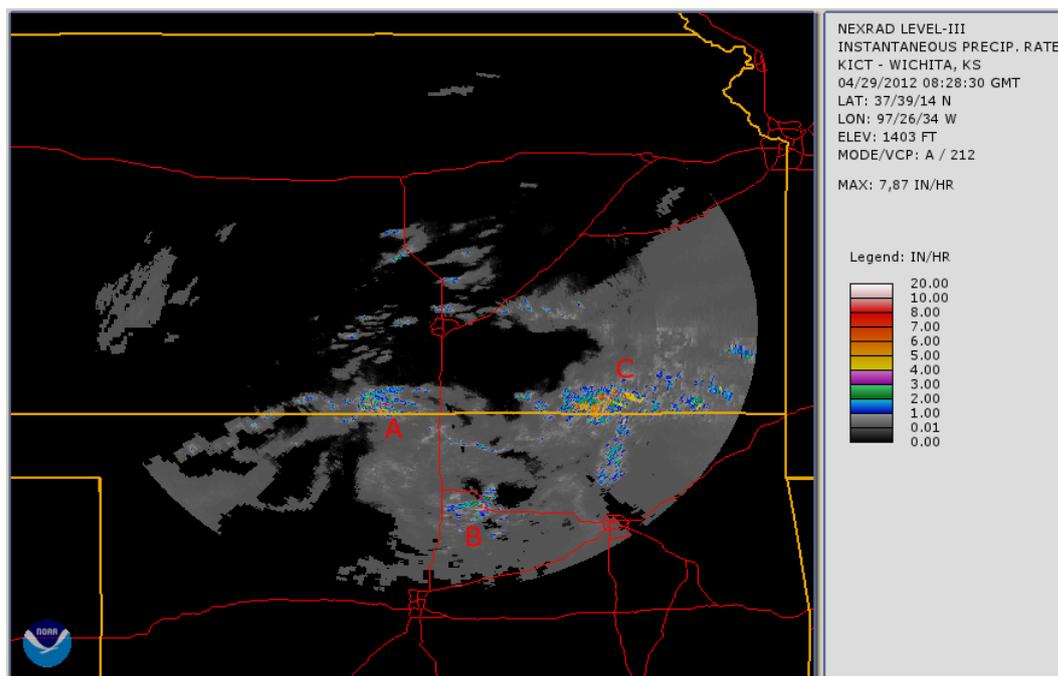


Figura 4.19: Tasa de Precipitación Instantánea sobre el norte del estado de Oklahoma y sur de Kansas para las 08:28 UTC del 29 de Abril de 2012, estimada mediante la red *NEXRAD*

Por otra parte, en la figura 4.19, se muestran las mediciones realizadas por el radar KICT del estado de Kansas, en el cual se han identificado tres grupos de precipitación A, B y C cada unos con tasa de precipitación instantánea superior a los 100mm/h respectivamente. Al observar la figura 4.17 se puede apreciar para esta zona que el algoritmo identifica en buena forma estos grupos de precipitación sin embargo subestima los niveles de tasa de precipitación instantánea.

En este orden de ideas, es importante destacar que los radares del *NEXRAD* proporcionan valores de tasa de precipitación instantánea cada 5 minutos, lo cual hace factible la

posibilidad de comparar operativamente con la salida de tasa de precipitación del algoritmo, sin embargo es necesario verificar la disponibilidad de datos, ya que para ciertas épocas algunos radares presentan ausencia de datos, debido principalmente a que se encuentran fuera de servicio por mantenimiento o inhabilitados para medir precipitación.

A modo de resumen, sobre el estudio de este caso, se puede resaltar que el algoritmo mostró una buena distinción entre píxeles de precipitación y píxeles de no precipitación, por otro lado, las salidas del algoritmo muestran cierta tendencia a subestimar tasa de precipitación en comparación con las mediciones llevadas a cabo por el radar, también el algoritmo dejó en evidencia la capacidad para discriminar entre zonas de altas, medias y bajas precipitaciones, esta afirmación se fundamenta en el hecho de que en el mapa de precipitación de la figura 4.17 se diferencia entre píxeles de alta tasa de precipitación instantánea, media y baja precipitación.

En base a los resultados reportados por el radar, se decide descartar el uso de los radares *NEXRAD* como herramienta de comparación, por la falta de homogeneidad de datos en relación a los casos de estudio presentados, también es importante mencionar que la diferencia entre los niveles de precipitación instantánea mostrados por radar y el algoritmo es tan grande, que dificulta su utilidad para establecer análisis concluyentes sobre la calidad del nivel de estimación reportado por el algoritmo bajo estudio.

4.3.2. CASO A.2: Precipitación sobre Oklahoma, Kansas, Misuri e Illinois del 30 de Abril de 2012

La precipitación del 30 de Abril de 2012, sobre los estados Oklahoma, Kansas, Misuri e Illinois dejó como resultado algunas situaciones de inundaciones ocasionadas por los altos niveles de precipitación registrados por el *NWS*, tal y como lo indica la figura 4.20, donde se muestran valores de precipitación acumulada superiores a 100mm en los estados de Kansas, Oklahoma y Misuri. Por otra parte, la figura 4.21 muestra la precipitación acumulada durante las 8:00 y 9:00 UTC, en la cual se han resaltado tres zonas de precipitación. La primera zona etiquetada con un 1, contiene valores de precipitación máximo de 12,5mm (0.50 pulgadas), y

geográficamente tuvo lugar en el oeste de la frontera entre los estados de Oklahoma y Texas.

La estimación obtenida mediante el algoritmo puede observarse en la figura 4.22, donde para la zona 1 se aprecia un tasa de precipitación instantánea máxima de 9 mm/h, la cual se corresponde con la zona del máximo horario presentado por el *NWS*. Respecto al patrón observado, el algoritmo identifica los píxeles de precipitación en correspondencia con los observado por el *NWS*. Por otra parte, en la figura 4.21 se muestra una gran zona de precipitación sobre los estados de Oklahoma, Kansas, Misuri e Illinois, etiquetada con el número 2. Los máximos de precipitación observados por el *NWS* son de 25,4 mm (1 pulgada) para el estado de Oklahoma, 12,5mm (0.50 pulgadas) sobre Kansas y 19,05mm (0.75pulgadas) sobre los estados de Misuri e Illinois.

En la figura 4.22 se puede apreciar que el patrón general estimado por el algoritmo para la zona 2, se corresponde en forma ajustada al observado por la *NWS*. En el estado de Oklahoma el algoritmo identifica 3 núcleos de precipitación con máximos de precipitación alrededor de 15 mm/h, por otro lado en el estado de Kansas el algoritmo sobrestima el nivel de precipitación al sur del estado, sin embargo en la zona central la precipitación estimada es bastante ajustada al reportarse un valor de 12 mm. En los estados de Misuri e Illinois se observan máximos de 17mm/h, lo cual en conjunto con los patrones observados definen satisfactoriamente el escenario de precipitación observado por la *NWS*

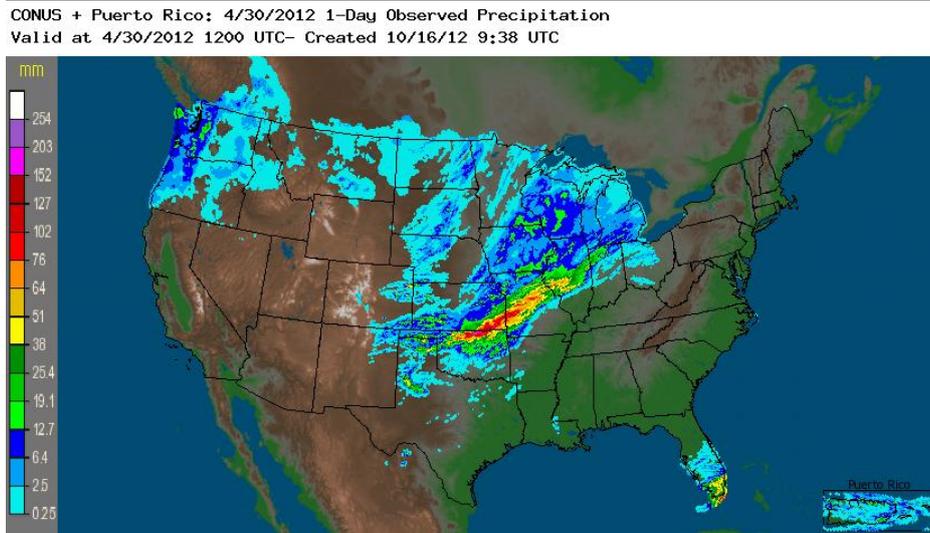


Figura 4.20: Precipitación acumulada para 24 horas del 30 de Abril de 2012 en los estados de Texas, Kansas, Oklahoma, Misuri y Illinois de los Estados Unidos. (Fuente:<http://water.weather.gov/precip/>)

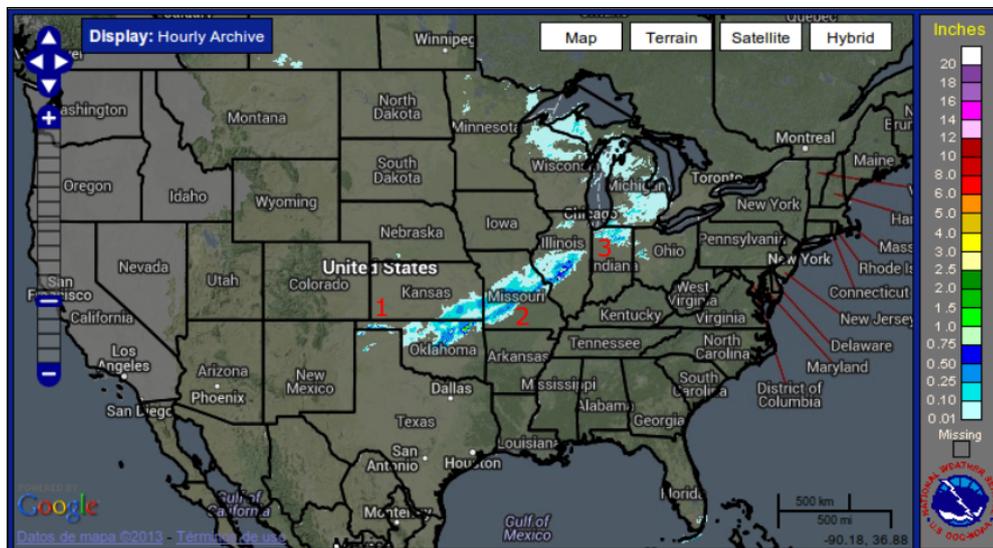


Figura 4.21: Precipitación horaria para las 09:00 horas UTC, del 30 de Abril de 2012 en los estados de Texas, Kansas, Oklahoma, Misuri y Illinois de los Estados Unidos. (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/)

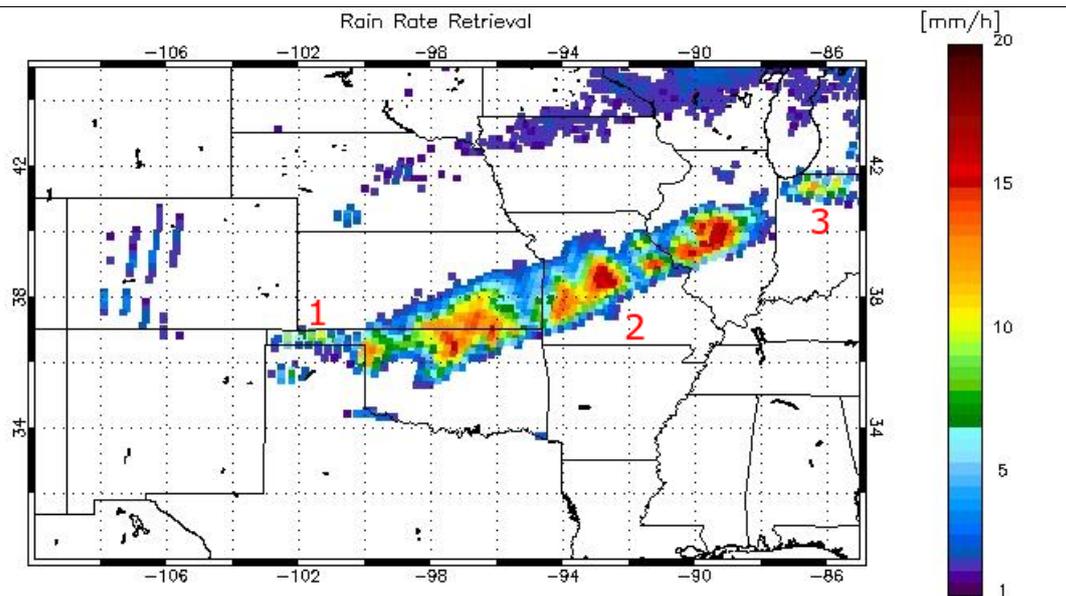


Figura 4.22: Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 08:00 y 8:08 UTC del 30 de Abril de 2012.

Otra zona de precipitación importante observada por el *NWS*, es la etiquetada con 3 en el estado de Indiana, la cual se muestra en la figura 4.21. La precipitación observada al norte de Indiana mostró valores máximos de precipitación de 12,5mm, mientras que para el instante de captura el algoritmo reporta una tasa de precipitación instantánea de 12 mm/h tal y como se aprecia en la figura 4.22. El patrón mostrado para la precipitación sobre Indiana es compatible con el observado por el *NWS*

En general, se puede mencionar para este caso de estudio, que el algoritmo reproduce de forma satisfactoria los patrones observados por el *NWS*, sin embargo sobre los estados de Colorado, Nebraska y Iowa el algoritmo identificó píxeles de precipitación que no fueron observados por el *NWS*, los cuales pudieran en este contexto identificarse como píxeles mal clasificados, en cuyos casos la tasa de precipitación máxima mostrada es en el orden de 3 mm/h.

Siguiendo la metodología mostrada en las comparaciones de los datos de precipitación observados por el *NWS* con las tasas de precipitación estimadas por el algoritmo para los casos A.1 y A.2 a continuación se muestran 4 casos mas de precipitación cuyos comen-

tarios generales se realizan considerando como variables de comparación a la reproducción del patrón de precipitación por parte del algoritmo en referencia al observado por el NWS reportado mediante la imagen de precipitación acumulada horaria. Por patrón de precipitación debe entenderse, la apreciación visual de la escena sobre los píxeles de precipitación que son correctamente definidos por el algoritmo respecto al patrón observado por el NWS para el evento. La otra variable a considerar es la tasa máxima de precipitación estimada por el algoritmo para las zonas de interés de estudio en cada caso, lo cual proporciona una referencia para conocer si el algoritmo estima en forma ajustada, subestima o sobrestima.

Finalmente la tercera variable a considerar en cada análisis es la presencia de píxeles de precipitación, definidos de forma incorrecta por el algoritmo en comparación con el mapa de precipitación horaria mostrada para cada caso.

4.3.3. CASO B.1: Evento de precipitación 9 de Marzo de 2013

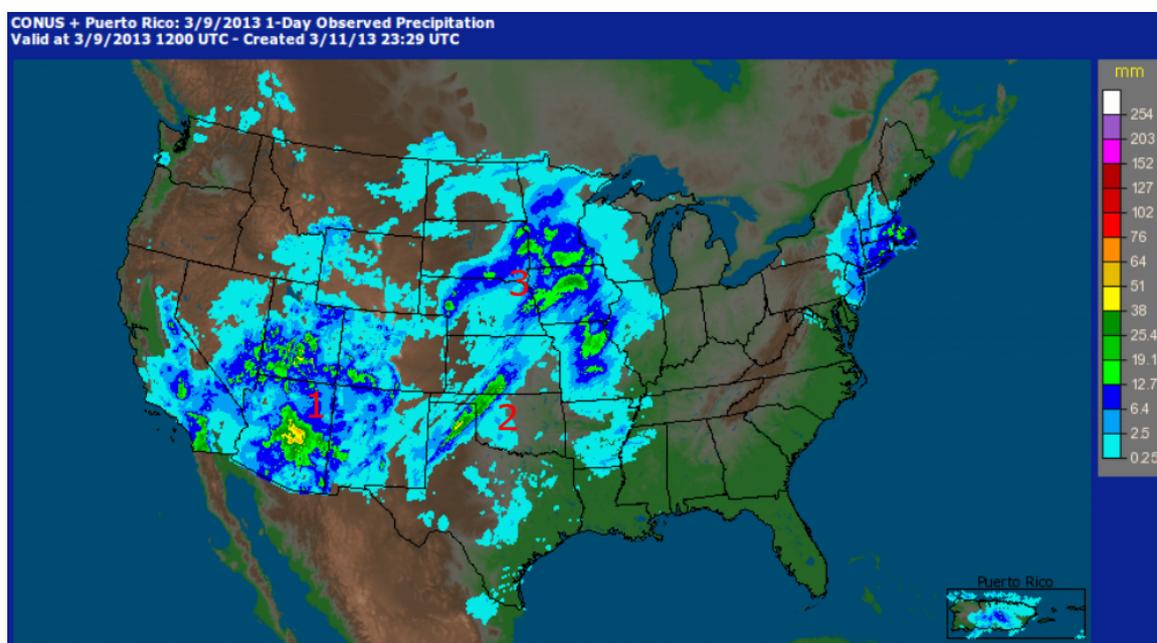


Figura 4.23: Precipitación acumulada para 24 horas del 09 de Marzo de 2013 en los Estados Unidos. Se han resaltado tres zonas de precipitación, en las cuales se aprecian precipitaciones máximas superiores a los 25mm (Fuente:<http://water.weather.gov/precip/>)

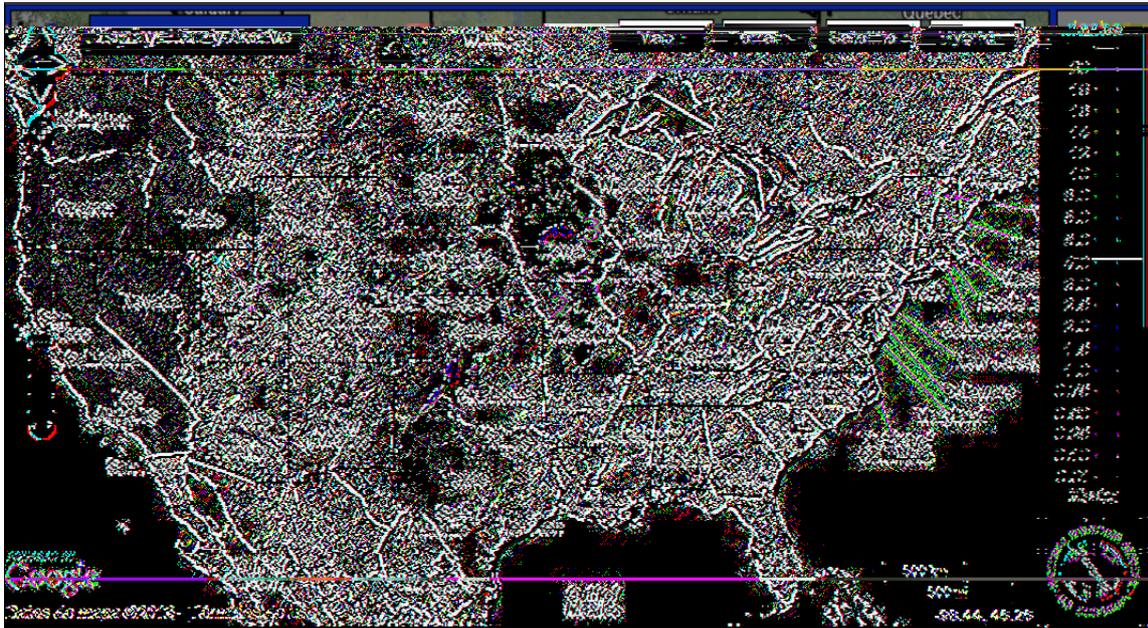


Figura 4.24: Precipitación horaria para las 09:00 horas UTC, del 9 de Marzo de 2013 en los Estados Unidos. Se muestran 3 zonas de interés de precipitación las cuales son comparadas con las estimaciones del algoritmo (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/)

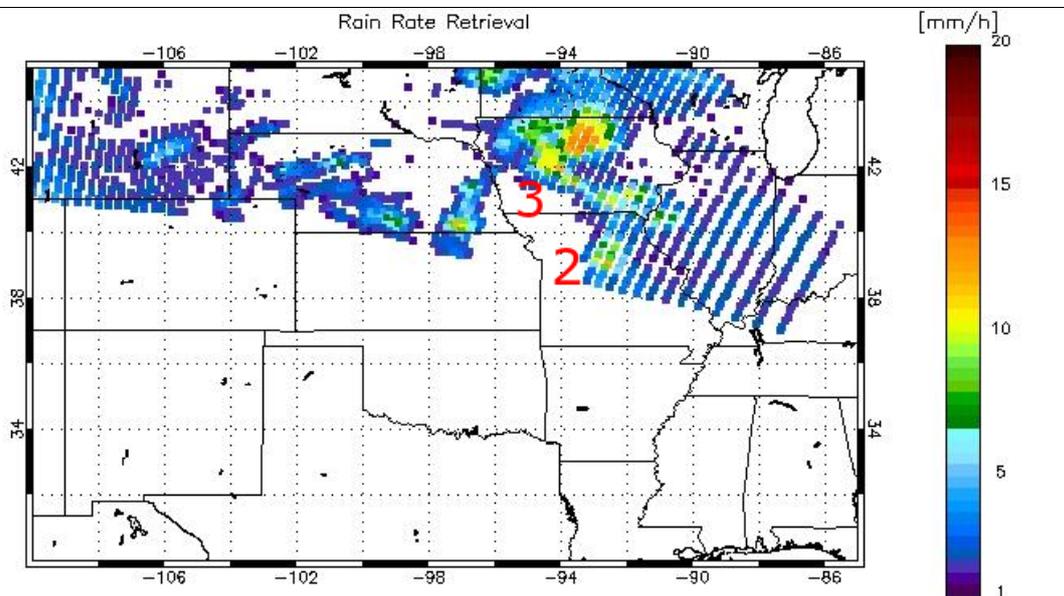


Figura 4.25: Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 08:30 y 8:38 UTC del 29 de Abril de 2012. Lo zona 3 muestra precipitaciones sobre el norte y sureste de Iowa con tasas de precipitaciones máximas de 12mm/h y 11mm/h respectivamente, mientras que la zona 2 muestra una parte de la precipitación sobre Misuri con un máimo de precipitación de 8mm/h

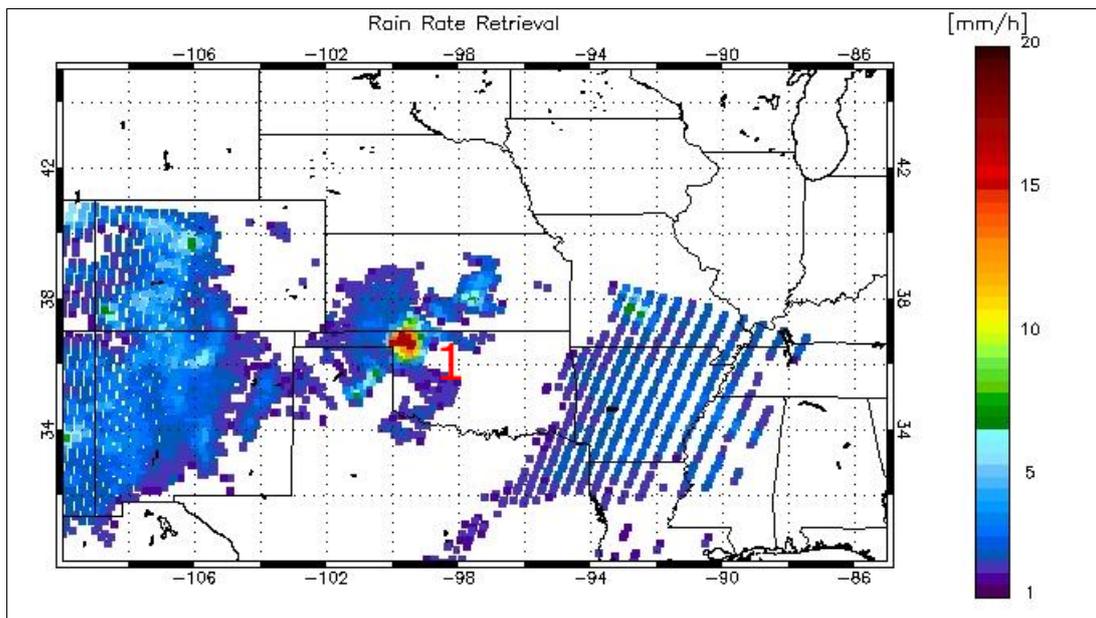


Figura 4.26: Tasa de Precipitación Instantánea, para la escena ATMS barrida entre las 08:38 y 8:46 UTC del 29 de Abril de 2012. Se aprecia la segunda parte del patrón mostrado en el mapa de precipitación horaria por la NWS, en donde en el foco de precipitación que se aprecia en Oklahoma tiene una tasa de precipitación superior a 15mm/

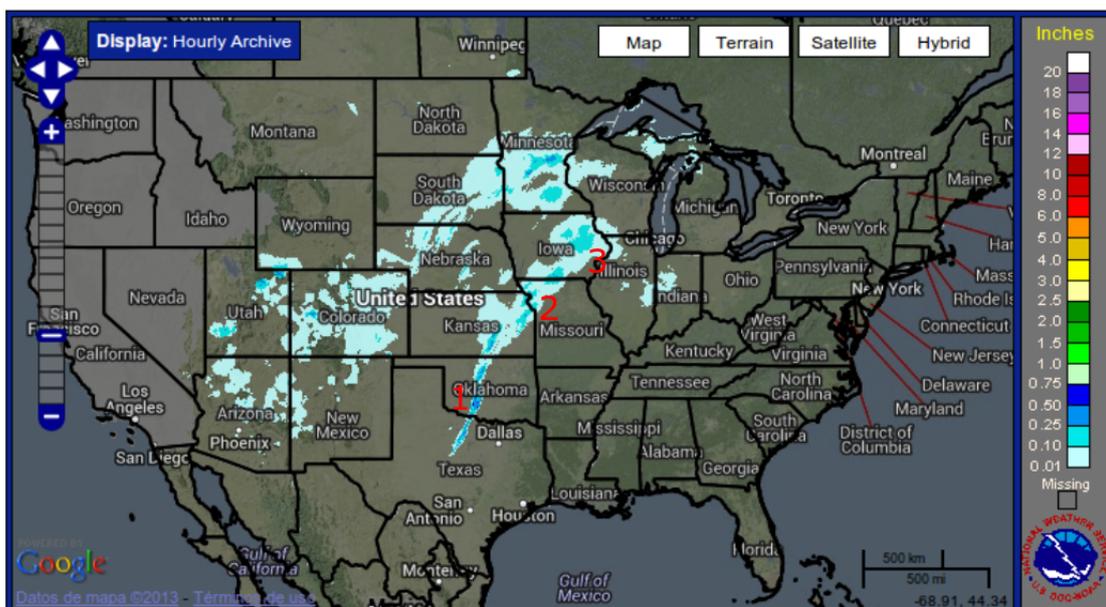


Figura 4.27: Precipitación horaria para las 20:00 horas UTC, del 9 de Marzo de 2013 en los Estados Unidos. Para este instante la precipitación sobre los estados de Iowa, Misuri, Oklahoma y Texas continua con un ligero corrimiento hacia el sureste (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/)

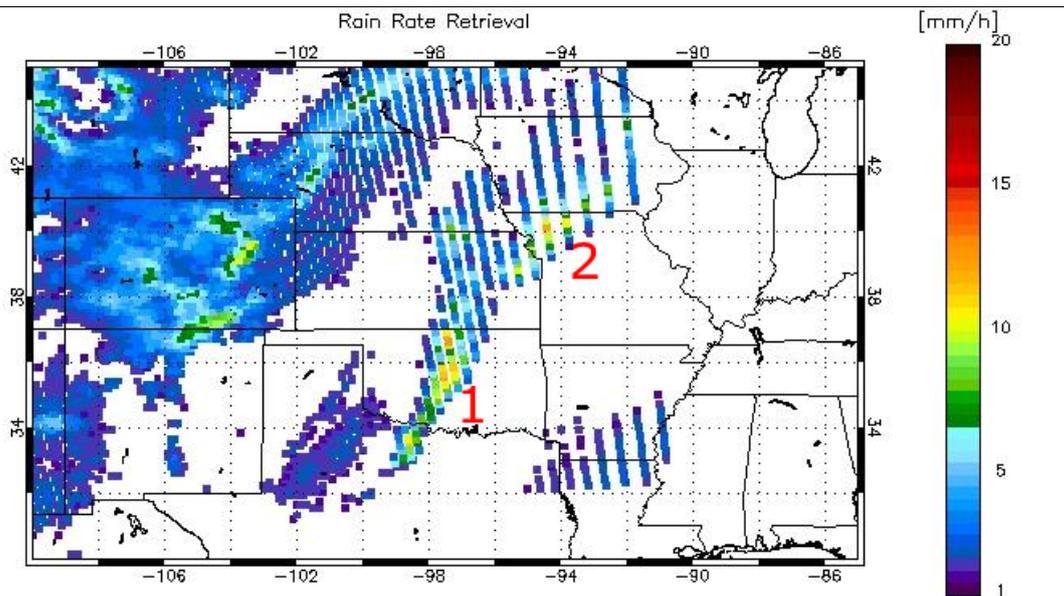


Figura 4.28: Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 19:58 y 20:06 UTC del 9 de Marzo de 2013. Se aprecia que el patrón de precipitación sobre Texas, Oklahoma y Misuri es bien reproducido por el algoritmo, con tasas máximas de precipitación de 11mm/h, 8mm/h y 11mm/h respectivamente.

En la comparación llevada a cabo mediante este caso de estudio se aprecia buena tendencia en el patrón de precipitación obtenido a partir del algoritmo, además también se obtuvo valores satisfactorios en términos de la magnitud de la tasa de precipitación obtenida mediante el algoritmo. Sin embargo, se aprecia una tendencia a clasificar de forma incorrecta, como píxeles de precipitación en algunos casos. Por ejemplo en la figura 4.25 se aprecia esta estimación incorrecta de píxeles de precipitación en ambos extremos del swath, al comparar este patrón con el mostrado en la figura 4.24, este mismo comportamiento puede apreciarse en las figura 4.26.

Esta clasificación incorrecta puede atribuirse al hecho de la necesidad de corrección geométrica de algunos canales como por ejemplo el 3, 4 y 5. Sin embargo, de la figura 4.28 se aprecia que para las zonas de precipitación 1 y 2, las tasas fueron estimadas mediante el algoritmo en zonas donde se espera mayor error debido a la falta de corrección de los datos *ATMS*, sin embargo tanto el patrón como los niveles de precipitación estimados son satisfactorios, con lo cual se deduce que ante la presencia de altos niveles de precipitación disminuye la afectación debido a la presencia de ruido introducido al algoritmo.

4.3.4. CASO B.2: Evento de precipitación 10 de Marzo de 2013

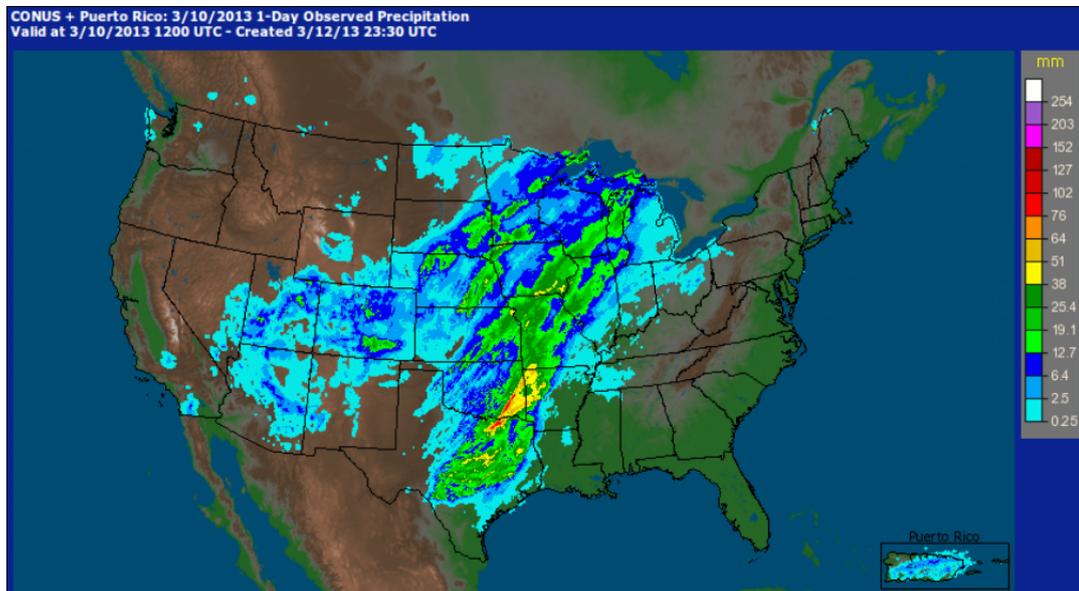


Figura 4.29: Precipitación acumulada para 24 horas del 10 de Marzo de 2013. Se aprecian altas precipitaciones sobre gran parte de la región central del país, de las cuales resaltan las máximas tasas de precipitación sobre los estados de Oklahoma y Arkansas, las cuales superan los 50mm (Fuente:<http://water.weather.gov/precip/>)

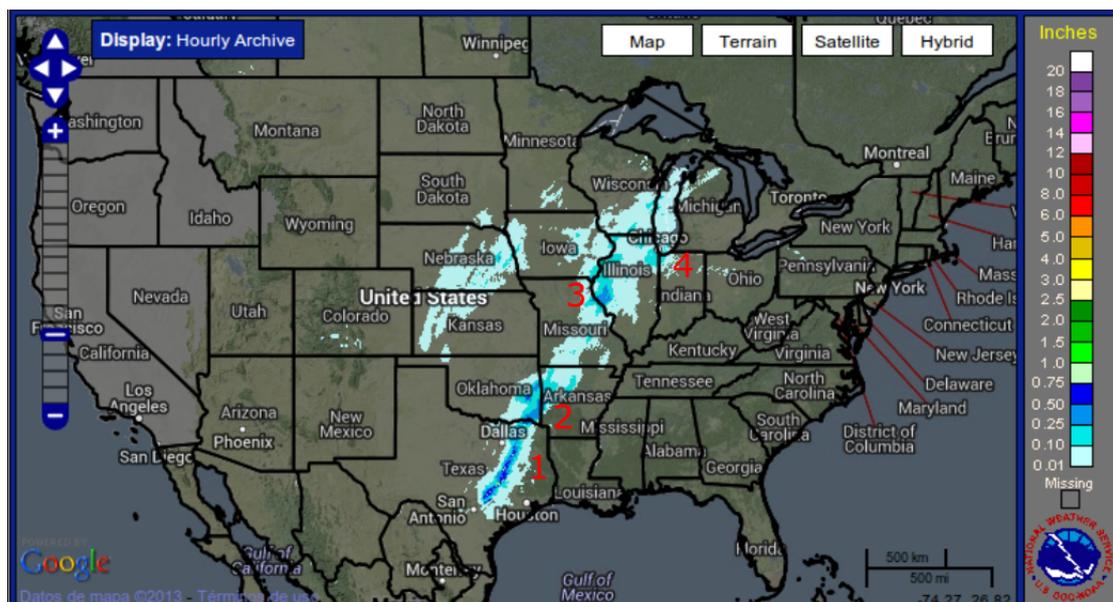


Figura 4.30: Precipitación horaria para las 09:00 horas UTC, del 10 de Marzo de 2013 en los Estados Unidos. Se aprecian 4 zonas de precipitación con máximos de precipitación que varían desde 12,5 mm (0.50 pulgadas) hasta 19 mm (0.75 pulgadas) (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/)

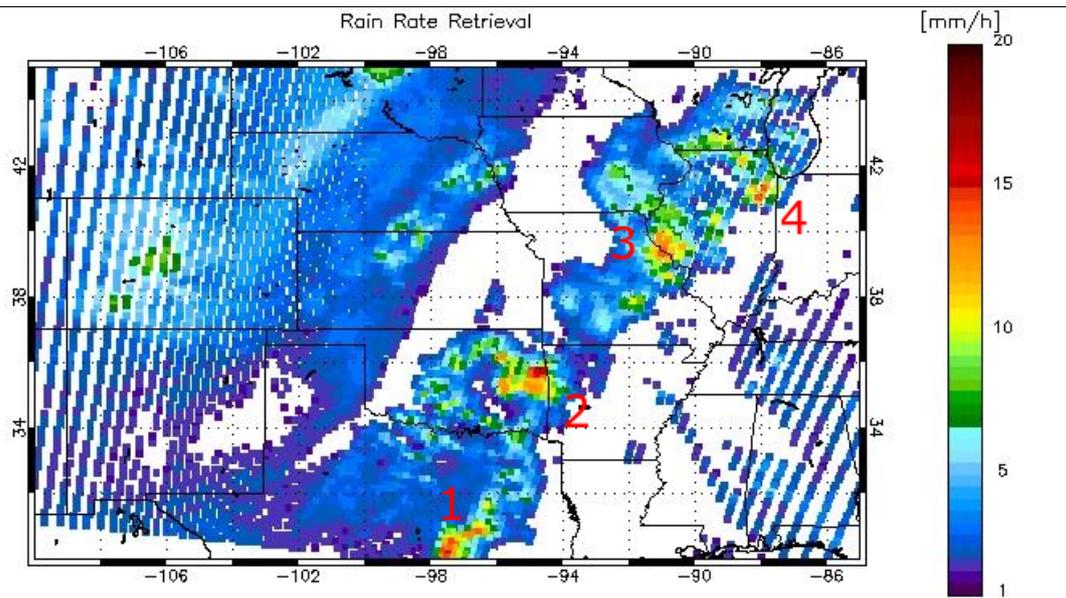


Figura 4.31: Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 08:14 y 08:22 UTC del 9 de Marzo de 2013. Se observa buena tendencia en la estimación de tasa de precipitación para las zonas de interés resaltadas, sin embargo también se aprecian una cantidad considerable de píxeles mal clasificados como precipitantes en comparación con el mapa de precipitación horario presentado por la *NWS*.

4.3.5. CASO C.1: Evento de precipitación 8 de Mayo de 2013

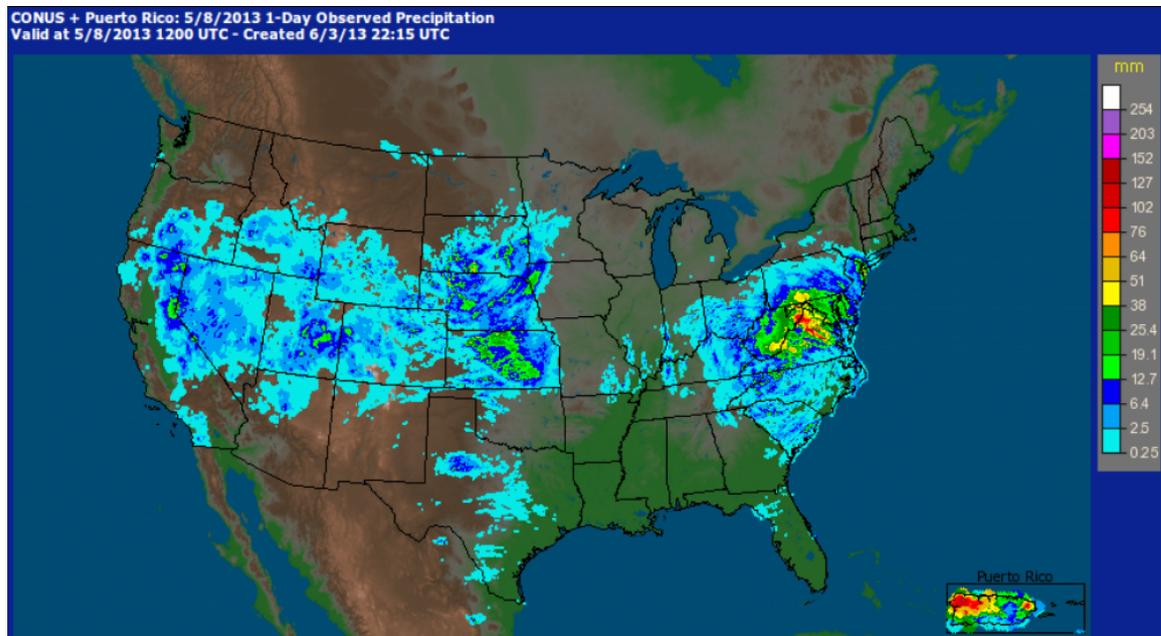


Figura 4.32: Precipitación acumulada para 24 horas del 8 de Mayo de 2013. Para este escenario la adquisición de datos ATMS estuvo orientada a la región central del país, lo cual comprende los estados de Kansas, Nebraska, Dakota del Sur e Iowa (Fuente:<http://water.weather.gov/precip/>)

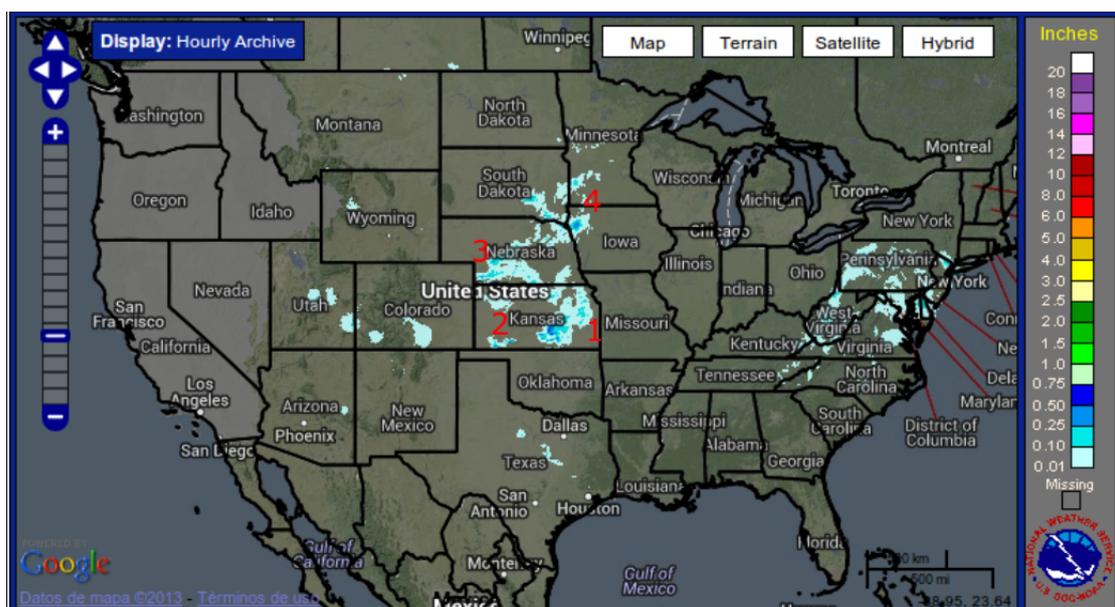


Figura 4.33: Precipitación horaria para las 09:00 horas UTC, del 8 de Mayo de 2013 en los Estados Unidos. Se han etiquetado 4 zonas de precipitación para ser comparados con la salida del algoritmo (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/)

4.3.6. CASO C.2: Evento de precipitación 9 de Mayo de 2013

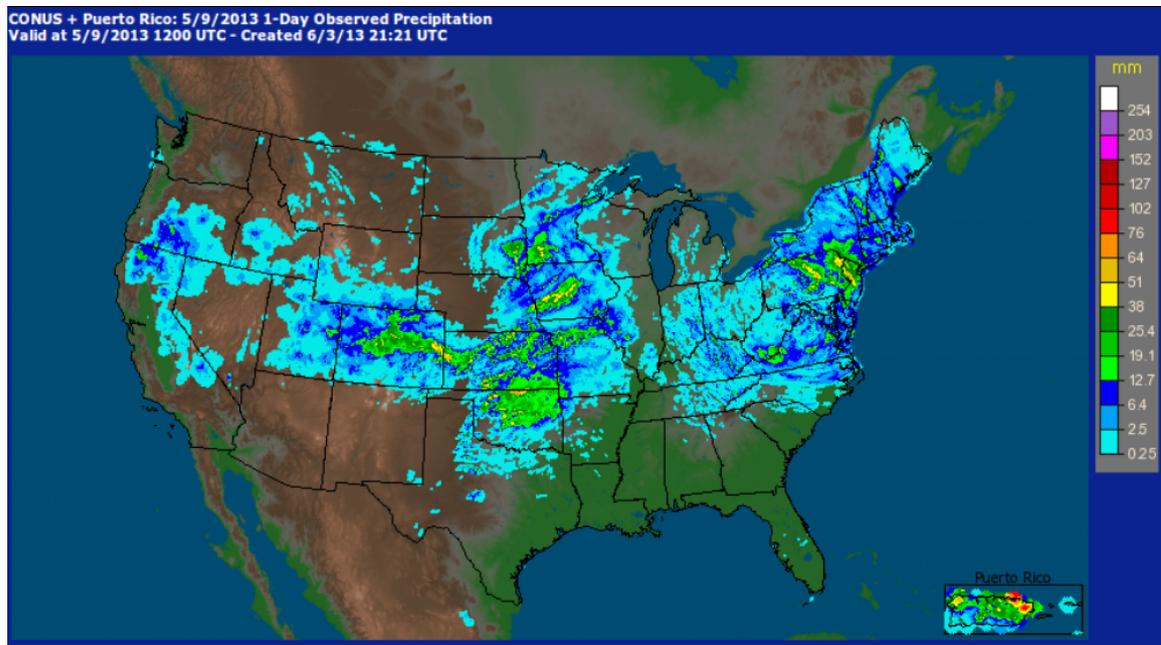


Figura 4.36: Precipitación acumulada para 24 horas del 9 de Mayo de 2013 en los Estados Unidos. (Fuente: <http://water.weather.gov/precip/>)

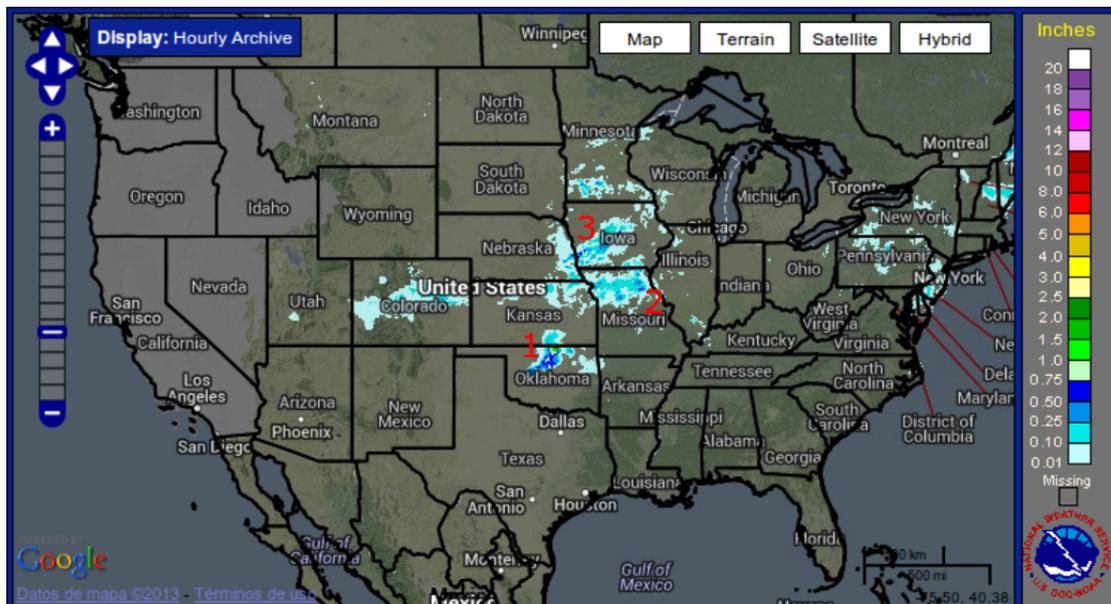


Figura 4.37: Precipitación horaria para las 09:00 horas UTC, del 9 de Mayo de 2013 en los Estados Unidos. Se seleccionan 3 zonas de precipitación para la comparación sobre los estados de Oklahoma, Misuri y Iowa (Fuente: http://www.srh.noaa.gov/ridge2/RFC_Precip/)

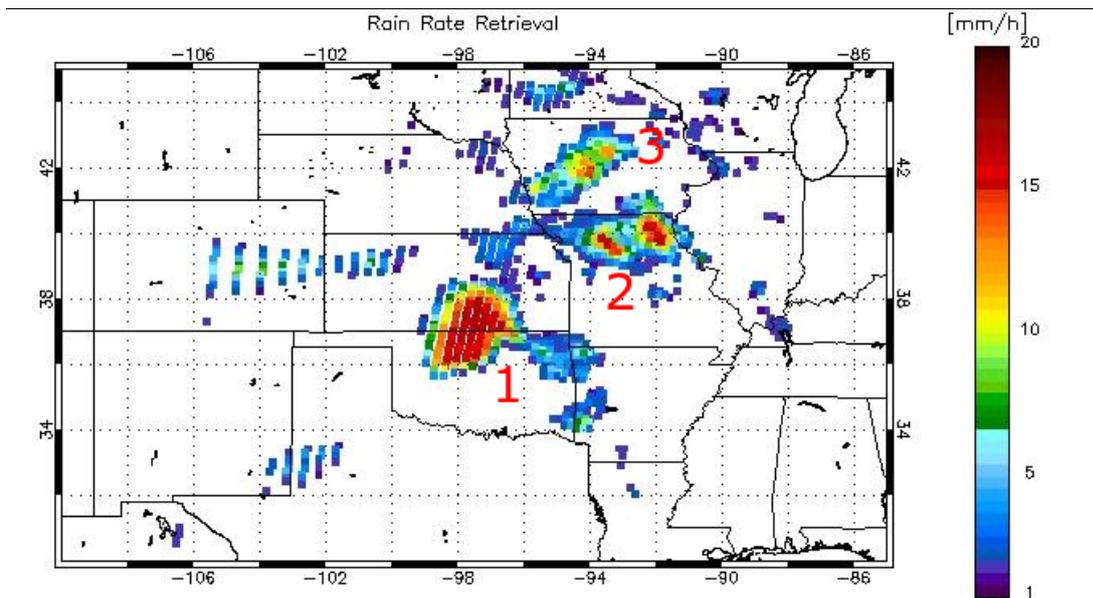


Figura 4.38: Tasa de Precipitación Instantánea, para la escena ATMS barrida entre las 07:49 y 07:57 UTC del 9 de Mayo de 2013. Se aprecia una clara tendencia a reproducir en buena forma los patrones de precipitación, de igual forma los niveles de tasa de precipitación estimados son satisfactorios. Sin embargo se aprecia una sobrestimación de precipitación al sur de Kansas.

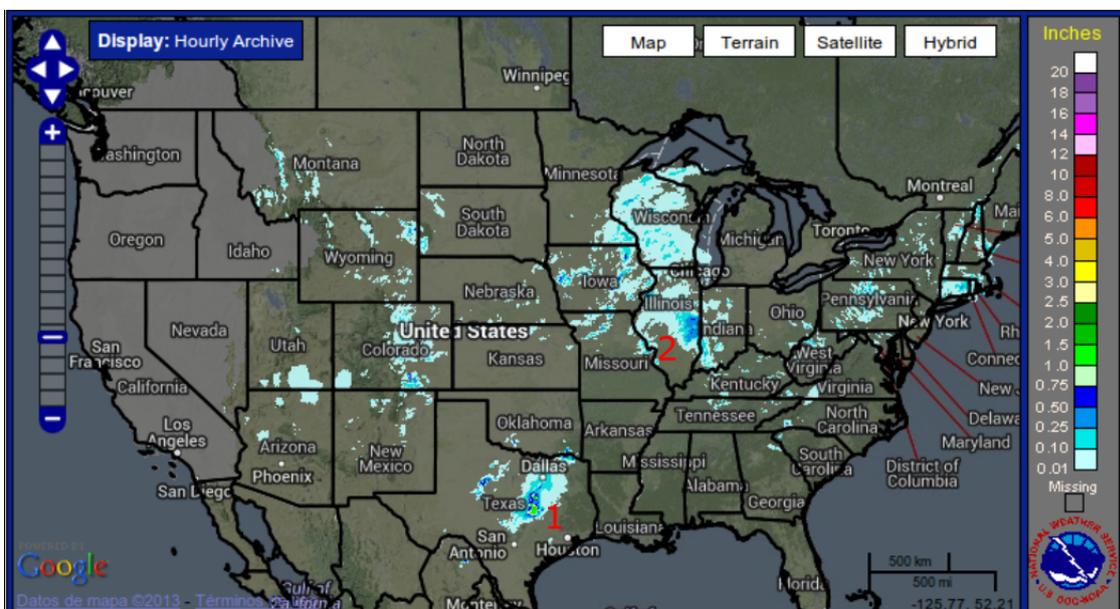


Figura 4.39: Precipitación horaria para las 20:00 horas UTC, del 9 de Mayo de 2013 en los Estados Unidos. De la cual se seleccionan dos focos de precipitación centrados sobre los estados de Texas e Illinois respectivamente (Fuente:http://www.srh.noaa.gov/ridge2/RFC_Precip/)

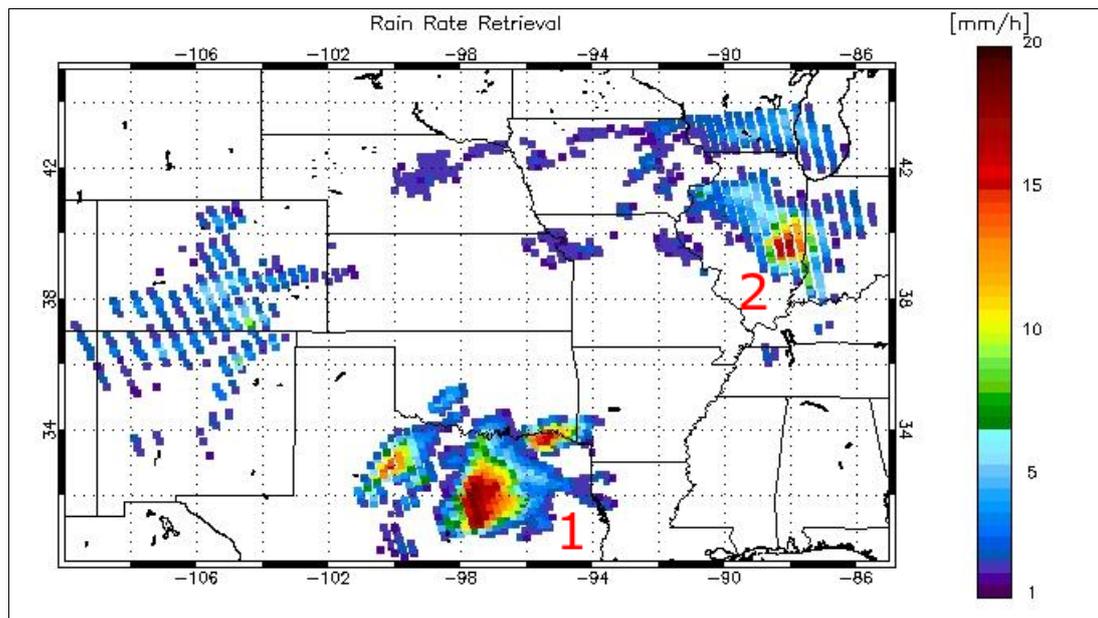


Figura 4.40: Tasa de Precipitación Instantánea, para la escena *ATMS* barrida entre las 19:09 y 19:17 UTC del 9 de Mayo de 2013. Se aprecia buena definición de los patrones de precipitación niveles de tasa de precipitación satisfactorios

CAPÍTULO 5

CONCLUSIONES

5.1. Conclusiones

5.1.1. Sobre la selección de canales *ATMS* para el entrenamiento de redes neuronales para la elaboración del algoritmo de estimación de tasa de precipitación

La selección de canales *ATMS* para la elaboración del algoritmo es determinante sobre las capacidades de estimación de precipitación, por este motivo es importante establecer metodologías basadas en principios de la teledetección de la atmósfera para microondas pasivas con la finalidad de optimizar el aprovechamiento de los datos. En este contexto, ha sido necesario descartar algunos canales debido a la falta de relación con la variable precipitación, otros fueron descartados debido a que requerían procesos de corrección que, por la complejidad y el tiempo necesario para su elaboración, escapan de los objetivos de

esta investigación.

En este orden de ideas, se realizaron pruebas con los canales que mostraron menor nivel de ruidos, los cuales resultaron ser los canales 3, 4, 5, 16, 17, 18, 19, 20, 21 y 22. Estos canales, y algunas combinaciones conformadas por estos canales, fueron evaluados para el entrenamiento de redes neuronales obteniendo los resultados mas altos para la combinación que comprende los 10 canales.

5.1.2. Sobre el uso de Redes Neuronales para el diseño de Algoritmos de Estimación de Tasa de Precipitación

Las redes neuronales entrenadas para el desarrollo de este algoritmo, demostraron capacidad para reproducir los fenómenos que tienen lugar en el proceso de precipitación, lo cual es afirmado a partir de la comparación cualitativa los mapas de precipitación obtenidos por el algoritmo. El diseño de redes utilizado en este algoritmo involucró entrenamiento de redes neuronales para regresiones lineales, por lo cual fue requerido un conjunto de datos para el proceso de entrenamiento que esencialmente estuvo compuesto por datos simulados para los canales *ATMS*, que además estaban altamente correlacionados con diferentes valores de tasas de precipitación en tierra, esto fue logrado mediante la implementación de la ecuación de transferencia radiativa.

Es destacable el hecho de que las capacidades de predicción de una red neuronal, depende entre otras cosas, de la calidad de los datos usados para el proceso de entrenamiento, y esta calidad se refiere a la capacidad intrínseca de los datos de entrenamiento para explicar, reproducir y describir la variable precipitación. En este contexto, es importante mencionar que la inclusión de otro tipo de información, adicional a las temperaturas de brillo de los 22 canales del *ATMS*, aumentarían las capacidades de estimación de precipitación por parte de la red neuronal.

En este sentido, se debe resaltar el hecho de que solo fueron usados 10 canales del sensor *ATMS* para el entrenamiento de las redes implementadas en el algoritmo, así que una

alternativa para mejorar la capacidad de estimación de las redes neuronales, sería incluir los canales con alta correlación con la variable precipitación, llevando a cabo un proceso previo de corrección no solo para estos canales, sino además para todos los que ya han sido incorporados al algoritmo y requieran de corrección.

Entre las principales ventajas que presenta el uso de redes neuronales para la elaboración de algoritmos de estimación de precipitación basados en datos provenientes de instrumentos satelitales de microondas pasivas resalta el hecho de que no se considera de forma explícita los principios físicos que tienen lugar en los procesos de generación de precipitación, en su lugar, las interrelaciones entre las variables atmosféricas y la precipitación son asimiladas por la red neuronal de forma indirecta a partir de los datos de entrenamiento, los cuales pueden ser simulados como en este trabajo, o también pueden ser generados a partir de sistemas de observación y medición cuidadosamente supervisados. Esto sin lugar a dudas, reduce considerablemente la complejidad asociada a la formulación del algoritmo, lo cual se traduce en una minimización en esfuerzo y horas hombre necesaria para el desarrollo del mismo.

Las herramientas computacionales y numéricas, también influyen de forma importante en los resultados obtenidos mediante las redes neuronales, lo cual es comprensible a partir de la comparación que se llevó a cabo para seleccionar entre las librerías *Open NN* y *ALGLIB*, la que resultaría definitivamente implementada para entrenar las redes neuronales del algoritmo. En este contexto, también es importante para la obtención de resultados óptimos, una correcta selección del algoritmo de entrenamiento, en este caso, se utilizó el algoritmo **Quasi Newton**, el cual por ser un algoritmo de segundo orden resultó ser una buena combinación entre el tiempo exigido para el proceso de entrenamiento y la calidad de la salida de la red neuronal. Es importante destacar que las diferencias observadas entre ambas librerías, pueden estar relacionadas a las técnicas de parada del proceso de entrenamiento para evitar el sobre o sub-entrenamiento, así como también a los métodos de optimización implementados en cada caso.

5.1.3. Sobre las Salidas del Algoritmo

El objetivo de esta investigación, en lo que refiere a la salida del algoritmo, comprendía establecer una metodología de comparación con alguna fuente de datos validada. Dicha comparación será utilizada para identificar deficiencias y fortalezas que permitan en posteriores investigaciones tomar en cuenta las medidas correctivas para lograr un producto con capacidades operativas para la estimación de tasa de precipitación.

En este contexto, se debe mencionar que el algoritmo fue evaluado en tres aspectos para 6 casos de precipitación sobre los Estados Unidos, considerando como variables de referencia el patrón espacial de precipitación, las máximos de tasa de precipitación estimados y la definición incorrecta de zonas de precipitación. En este sentido se puede mencionar que generalmente el algoritmo mostró una propiedad positiva en reproducir y estimar los patrones espaciales de precipitación observados por el *NWS*, sobre todo en los casos de tasas de precipitación altas (aproximadamente entre 10 y 20mm).

De igual modo, los niveles de tasa de precipitaciones estimadas tanto para píxeles de precipitación estratiforme o convectiva fueron en general satisfactorios, sin embargo no pudo ser llevado a cabo un proceso de comparación cuantitativa debido a la carencia de datos que cumplieran con los requisitos necesarios para llevar a cabo esta tarea, tales como coincidencia en resolución temporal y espacial, así como el formato nativo en el cual algunos datos son distribuidos. Por ejemplo, se intentó realizar comparación cuantitativa con datos de tasa de precipitación instantánea del sistema de radares *NEXRAD* de los Estados Unidos debido principalmente a la resolución temporal, sin embargo la diferencia en resolución espacial y el formato en el que los datos están dispuestos dificultaron el proceso de evaluación y a su vez estas diferencias aumentan la pérdida de correspondencia y generalidad de cualquier comparación cuantitativa que pudiera derivar de estos datos, principalmente por la diferencias de las escalas de los fenómenos físicos que son apreciados a partir del radar y del *ATMS*.

En cuanto a la incorrecta designación de píxeles de precipitación se tiene que los casos mas notorios son explicados a partir del hecho que algunos canales del *ATMS*, como el 3, 4 y 5 requieren de corrección geométrica por diferencias en el ángulo zenith de captura

entre la dirección nadir del sensor y los extremos de swath. Por tal motivo, es importante la implementación de un proceso de corrección previo al proceso de entrenamiento de las redes neuronales.

En base a las características observadas a partir de la comparación de las salidas estimadas por el algoritmo con las observaciones del *NWS*, se concluye que en general el algoritmo obtenido es capaz de discriminar precipitaciones y estimar tasas de precipitaciones de forma satisfactoria ante escenarios de precipitación convectiva. Por otro lado, en caso de precipitaciones estratiformes el algoritmo tiende a clasificar de forma incorrecta como píxeles de precipitación en los contornos próximos a los píxeles de verdadera precipitación.

Finalmente, en regiones de los datos *ATMS* cercanas a los extremos del swath, donde el ruido incluido por los canales 3, 4 y 5 se hace considerable, se observó que la mayoría de las veces, las tasa de precipitaciones reportadas con valores inferiores a 3mm son falsas precipitaciones.

5.2. Recomendaciones

A continuación se presenta una lista de recomendaciones que al juicio del autor contribuirían a mejorar los resultados presentados en este trabajo.

1. Realizar la corrección por diferencia en ángulo zenith de los canales *ATMS*.
2. Probar otras arquitecturas de redes neuronales, incorporando los canales previamente corregidos y que estén correlacionados con la variable precipitación.
3. Comparación de resultados obtenidos mediante el algoritmo con datos validados para verano e invierno, para píxeles sobre tierra y océano respectivamente.

5.3. Próximas Investigaciones

1. Realizar la corrección por diferencia en ángulo zenith de los canales *ATMS*.
2. Evaluar la inclusión de variables o índices atmosféricos altamente relacionadas con la precipitación como datos asimilados al proceso de entrenamiento de las redes neuronales.
3. Validación y Calibración de las salidas del algoritmo en escenarios de tierra y océano, bajo condiciones de verano e invierno.
4. Evaluación de las potencialidades del algoritmo para el análisis de precipitaciones severas como herramienta que facilite la toma de decisiones ante escenarios de riesgos y emergencias.

REFERENCIAS

- Andrews, D., 2010. An introduction to atmospheric physics. CAMBRIDGE UNIVERSITY PRESS. ([document](#)), [2.4](#), [2.5](#), [2.2.3](#), [2.2.3](#), [2.7](#), [2.2.3](#)
- Arias, F., 1999. El proyecto de investigación: guía para su elaboración. Episteme, Caracas, Venezuela. [3](#)
- Arkin, P., 1979. The relationship between the fractional coverage of high cloud and rainfall accumulations during gate over the b-scale array. Mon. Wea. Rev., 1382–1387. [2.1.3](#)
- Arkin, P. A., Meisner, B. N., 1987. The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. Mon. Wea. Rev., 51–74. [2.1.3](#)
- Barrett, E., 1970. The estimation of monthly rainfall from satellite data. Mon. Weath. Rev., 322–327. [2.1.1](#)
- Barrett, E., 1997. Satellite rainfall monitoring: recent progress and remaining problems. Proceedings of Rabat Symposium S3, 141–148. [2.1.1](#), [2.1.1](#)
- Barrett, E., Beaumont, M., 1994. Satellite rainfall monitoring: an overview. Remote Sensing Rev., 23–48. [2.1.1](#)

- Barrett, E., Martin, D., 1981. The use of satellites in rainfall monitoring. *Mon. Weath. Rev.* [2.1.1](#)
- Bellerby, T., Todd, M., Kniveton, D., Kidd, C., 2000. Rainfall estimation from a combination of trmm precipitation radar and goes multispectral satellite imagery through the use of an artificial neural network. *American Meteorological Society*, 2115–2128. ([document](#)), [2.1.3](#), [2.4](#), [2.1.3](#), [2.1.3](#), [3.1.1](#)
- Bishop, C., 2005. Neural network for pattern recognition. [2.4.2](#)
- Blackwell, W., Chen, F., 2009. *Neural Networks in Atmospheric Remote Sensing*. Massachusetts Institute of Technology. [2.4](#)
- Blackwell, W. J., Chen, F., 2005. Neural network applications in high-resolution atmospheric remote sensing. *Lincoln Laboratory Journal*, 299–322. [2.1.3](#), [3.1.1](#)
- Brasseur, G., Salomon, S., 2005. *Aeronomy of the middle atmosphere: Chemistry and physics of the stratosphere and mesosphere*. *ATMOSPHERIC AND OCEANOGRAPHIC SCIENCES LIBRARY*. ([document](#)), [2.6](#)
- Chen, F., Bickmeier, L., Blackwell, W., Leslie, R., Staelin, D., Surussavadee, C., 2006. Satellite-based estimation of precipitation using passive opaque microwave radiometry. Workshop of the International Precipitation Work-ing Group, Melbourne, Australia. ([document](#)), [2.2](#), [2.3](#), [2](#), [3.1.1](#)
- Chen, F., Staelin, D., 2003. Airs/amsu/hsb precipitation estimates. *IEEE Transactions on Geoscience and Remote Sensing*, 410–417. [2](#)
- Di Paola, F., Casella, D., Dietrich, S., Mugnai, A., Ricciardelli, E. Romano, F., Sanó, P., 2012. Combined mw-ir precipitation evolving technique (pet) of convective rain fields. *Nat. Hazards Earth Syst. Sci.*, 3557–3570. [1.1.3](#)
- Di Tomaso, E., Romano, F., Cuomo, V., 2009. Rainfall estimation from satellite passive microwave observations in the range 89ghz to 190ghz. *Journal of Geophysical Research*, 114. [1.1.3](#)
- Ferraro, R., Weng, F., Grody, N., Zhao, L., 2000. Precipitation characteristics over land from the noaa-15 amsu sensor. *Geophysical Research Letters*, 2669–2672. [2](#)

- Freeman, J., Skapura, D. M., 1991. Neural networks: Algorithms, applications, and programming techniques. Addison-Wesley, 401. [2.1.3](#)
- Grimes, D., Coppola, E., Verdecchia, M., Visconti, G., 2003. A neural network approach to real-time rainfall estimation for africa usind satellite data. american Meteorological Society, 1119–1133. [2.1.3](#)
- Haykin, S., 1999. Neural networks. a comprehensive foundation. Upper Saddle River. [2.3](#), [2.3.1](#)
- Hertz, J., Krogh, A., Palmer, R., 1991. Introduction to the theory of neural computation. [2.3.2](#)
- Hocking, J., Rayer, P., Saunders, R., Matricardi, M., Geer, A., Brunel, P., 2012. Rttov v10 users guide. ECMWF, KNMI and MÃ©tÃ©o France. [3.2](#)
- Hong, Y., Hsu, K., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. Journal of Applied Meterology, 1834–1852. [2.1.3](#)
- Houl, A., Skofronick-Jackson, G., Kummerow, C. D., Shepherd, J. M., 2008. Global precipitation measurement. In: Precipitation: Advances in Measurement, Estimation and Prediction. [1.1.3](#)
- Hsieh, W. W., 2009. Machine learning methods in the environmental sciences. neural networks and kernels. ([document](#)), [2.9](#), [2.11](#), [2.4](#)
- Hsu, K., Gao, X., Sorooshian, S., Gupta, V., 1997. Precipitation estimation from remotely sensed information using artificial neural networks. J. Appl. Meteor., 1176–1190. [2.1.3](#), [2.1.3](#)
- Iguchi, T., Kozu, T., Meneghini, R., Awaka, J., Okamoto, K., 2000. Rain-profiling algorithm for the trmm precipitation radar. J. Appl. Meteorol., 2038–2050. [2.1.1](#)
- Kidd, C., 2001. Satellite rainfall climatology: A review. Int J Climatol, 1041–1066. [1.1.3](#), [1.1.3](#), [2.1.1](#), [2.1.1](#)

- Kidd, C., Levizzani, V., 2011. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.*, 1109–1116. [2.1.1](#)
- Kummerow, C., Barnes, W., Toshiaki, T., Shiue, J., Simpson, J., 1998. The tropical rainfall measuring mission (trmm) sensor package. *Journal of Atmospheric and Oceanic Technology*, 809–817. [2.1.1](#)
- Kummerow, G., Giglio, L., 1995. A method for combining passive microwave and infrared rainfall observations. *J. Atmos. Oceanic Technol.*, 33–45. [2.1.3](#)
- Kuo-Nan, L., 1980. An introduction to atmospheric radiation. ACADEMIC PRESS, INC. [1](#), [2.2.3](#), [2.2.4](#), [2.2.4](#), [2.2.6](#)
- Laviola, S., Levizzani, V., 2011. The 183-wsl fast rain rate retrieval algorithm. part 1: Retrieval design. *Atmospheric Research*, 443–461. [1.1.3](#), [1.2](#), [2.1.1](#)
- LeCun, Y., Kanter, I., Solla, S. A., 1991. Second order properties of error surfaces: Learning time and generalization. *Neural Information Processing Systems*, 918–24. [2.3.3](#)
- Leslie, R. V., Blackwell, W. J., Bickmeier, L. J., Jairam, L. G., 2008. Neural network microwave precipitation retrievals and modeling results. In: *Microwave Remote Sensing of the Atmosphere and Environment VI*. ([document](#)), [1.2](#), [2.1.3](#), [2.5](#), [2.1.3](#)
- Lethbridge, M., 1967. Precipitation probability and satellite radiation data. *Mon. Weath. Rev.*, 487–490. [2.1.1](#)
- Levizzani, V., 2003. Satellite rainfall estimates: new perspectives for meteorology and climate from the eurainsat project. *ANNALS OF GEOPHYSICS*, 363–372. [2.1.1](#)
- Levizzani, V., Bauer, P., Turk, F. J., 2007. Measuring precipitation from space: Eurainsat and the future. *Springer*, 722. [1.1.3](#)
- Levizzani, V., Porcú, F., Marzano, F., Mugnai, A., Smith, E., Prodi, F., 1996. Investigating a ssm/i microwave algorithm to calibrate meteosat infrared instantaneous rain-rate estimates. *Meteorol. Appl.*, 5–17. [2.1.3](#)

- Levizzani, V., Schmetz, J., Lutz, H., Kerkmann, J., Alberoni, P., Cervino, M., 2001. Precipitation estimations from geostationary orbit and prospects for meteosat second generation. *Meteorol. Appl.*, 23–41. [1.1.3](#)
- Liu, G., Holton, J., Pyle, J., Curry, E., 2002. Satellite remote sensing: Precipitation. *Encyclopedia of Atmospheric Sciences*. [2.1.1](#)
- Lovejot, S., Austin, G., 1979. The delineation of rain areas from visible and infrared satellite data for gate and mid-Ålatitudes. *Atmos. Ocean.*, 1048–1054. [2.1.3](#)
- Macquart, J., 2010. Radiative process in astrophysics. Cambridge University Press. [2.2.5](#)
- Marzano, F., Cimini, D., Coppola, E., Verdecchia, M., Levizzani, V., Tapiador, F., Turk, J., 2005. Satellite radiometric remote sensing of rainfall fields: multi-sensor retrieval techniques at geostationary scale. *Advances in Geosciences*, 267–272. [2.1.3](#)
- Mas, J., Flores, 2008. The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, 617–663. ([document](#)), [2.3](#), [2.3.2](#), [2.10](#)
- Menzel, W. P., Purdom, J. F. W., 1994. Introducing goes-i: the first of new generation of geostationary operational environmental satellites. *Bull. Am. Meteorol. Soc.*, 757–781. [1.1.3](#)
- Minsky, M., Papert, S., 1969. Perceptrons. [2.3.3](#)
- Muñoz, E., Falcón, N., 2008. Meso-scale hydrostatic atmospheric model. In: VI Congreso Venezolano de Física. pp. 1–12. [1.1.1](#), [2.2.1](#), [2.2.1](#), [2.2.2](#)
- New, M., Todd, M., Hulme, M., Jones, P., 2001. Precipitation measurements and trends in the twentieth century, 1899–1922. [2.1.1](#)
- Prigent, C., 2010. Precipitation retrieval from space: An overview. *Comptes Rendus Geoscience*, 380–389. [2.1.1](#)
- Quartly, G. D., Guymer, T., Srokosz, M., 2002. Measuring rainfall at sea: Part 2 - Spaceborne sensors. Southampton Oceanography Centre. [2.1.1](#)

- Rivolta, G., Marzano, E., Coppola, E., Verdecchia, M., 2006. Artificial neural network technique for precipitation nowcasting from satellite imagery. *Advances in Geosciences*, 97–103. [2.1.3](#)
- Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 386–408. [2.3.3](#)
- Rosenblatt, F., 1962. *Principles of neurodynamics*. New York: Spartan. [2.3.3](#)
- Rumelhart, D. E., Hinton, G. E., Williams, R. J., 1986. Learning internal representations by error propagation. *Parallel Distributed Processing*, 318–62. [2.3.4](#), [2.4.2](#)
- Saunders, R., Hocking, J., Rayer, P., Matricardi, M., Geer, A., Bormann, N., Brunel, P., Karbou, F., Aires, F., 2006. Rttov-10 science and validation report. EUMETSAT. [3.2](#)
- Schmetz, J., Woick, H., Tjemkes, S., Rattenborg, M., 1998. From meteosat to meteosat second generation (msg). *Preproceedings of the 9th Conference on Satellite Meteorology and Oceanography*, 335–338. [1.1.3](#)
- Schonhuber, M., Lammer, G., Randeu, W. L., 2008. The 2d-video-distrometer. In: *Precipitation: Advances in Measurement, Estimation and Prediction*. [1.1.2](#)
- Scorer, R., 1987. Cloud reflectance variations in avhrr channel-3. *Int. J. Remote Sens.*, 675–686. [2.1.3](#)
- Seinfeld, J., Pandis, S., 2006. *Atmospheric chemistry and physics: From air pollution to climate change*. JOHN WILEY and SONS, INC. [2.2.1](#)
- Skapura, D., 1996. *Building neural networks*. ACM Press, 282. [2.1.3](#)
- Smith, E., Lamm, J., Adler, R., Alishouse, J., Onashi, K., Barrett, E., Bauer, P., Berg, W., Chang, A., Ferraro, R., Ferriday, J., Goodman, S., Grody, N., Kidd, C., Kniveton, D., Kummerow, C., Liu, G., Marzano, F., Mugnai, A., Olson, W., Petty, G., Shibata, A., Spencer, R., Wentz, F., Wilheit, T., Zipser, E., 1998. Results of the wetnet pip-2 project. *J. Atmos. Sci.*, 1483–1536. [2.1.1](#)

- Sorooshian, S., Hsu, K., Imam, B., Hong, Y., 2008. Global precipitation estimation from satellite image using artificial neural networks. In: *Hydrological Modeling in Arid and Semi-Arid Areas*. pp. 21–28. [1.2](#)
- Staelin, D. H., Chen, F., 2000. precipitation observations near 54 and 183 ghz using the noaa-15 satellite. *IEEE Transactions on Geoscience and Remote Sensing*, 2322–2332. ([document](#)), [2.1](#), [2.1.1](#), [2.2](#), [2](#), [2.1.3](#), [2.1.3](#)
- Strangeways, I., 2007. *Precipitation: Theory, measurement and distribution*. Cambridge University Press, New York. ([document](#)), [2.1](#), [2.1.1](#)
- Surussavadee, C., Staelin, D., 2008a. Global millimeter-wave precipitation retrievals trained with a cloud-resolving numerical weather prediction model, part i: Retrieval design. *IEEE Transactions on Geoscience and Remote Sensing*, 99–108. [2.1.3](#)
- Surussavadee, C., Staelin, D., 2008b. Global millimeter-wave precipitation retrievals trained with a cloud-resolving numerical weather prediction model, part ii: Performance evaluation. *IEEE Transactions on Geoscience and Remote Sensing*, 99–108. [2.1.3](#)
- Surussavadee, C., Staelin, D., 2010. Npoess precipitation retrievals using the atms passive microwave spectrometer. *IEEE Geoscience and Remote Sensing Letters*, 440–444. ([document](#)), [2.1.2](#), [2.3](#)
- Tapiador, F. J., Kidd, C., Hsu, K., Marzano, F., 2004. Neural networks in satellite rainfall estimation. *Meteorol. Appl.*, 83–91. [1.2](#), [2.1.3](#), [2.1.3](#)
- Tomassetti, B., Marzano, F., Montopoli, M., Verdecchia, M., 2008. Rainfall radar nowcasting using a neural-network cascade approach. *The fifth European Conference on Radar in Meteorology and Hydrology*. [7](#)
- Tsintikidis, D., Haferman, J., Anagnostou, N., 1997. A neural network approach to estimating rainfall from spaceborne microwave data. *IEEE Trans. Geosci. Remote Sens.*, 1079–1092. [2.1.3](#)
- Vicente, G., 1996. Algorithm for rainfall rate estimation using a combination of goes-8, 11 and 3.9 measurements. Preprints, Eighth Conf. on Satellite Meteorology and Oceanography, 274–278. [2.1.3](#)

- Weigend, A., Huberman, B., Rumelhart, D., 1991. Prediction sun-spots and exchange rates with connectionist networks. Addison-Wesley, 1–36. [2.1.3](#)
- Werbos, P. J., 1974. Beyond regression: new tools for prediction and analysis in the behavioural sciences. [2.3.4](#)
- Widrow, B., Hoff, M. E., 1960. Adaptive switching circuits. IRE WESCON Convention Record, 96–104. [2.3.3](#)
- Wilheit, T., Adler, R., Avery, S., Barrett, E., Bauer, P., Berg, W., Chang, A., Ferriday, J., Grody, N., Goodman, S., Kidd, C., Kniveton, D., Kummerow, C., Mugnai, A., Olson, W., Petty, G., Shibata, A., Smith, E., SPENCER, R., 1994. Algorithms for the retrieval of rainfall from passive microwave measurements. Remote Sensing Rev., 163–194. [2.1.1](#)
- Xingui, H., Shaohua, X., 2010. Process neural network: Theory and applications. ([document](#)), [2.8](#), [2.3.2](#), [2.4.2](#)

Lista de Acrónimos

- ATMS Sonda de Tecnología Avanzada para Microondas
- AMSU Unidad Avanzada de Sondeo de Microondas
- PMW Microondas Pasivas
- ESMR Radiómetro de Microondas de Escaneo Eléctrico
- SMMR Radiómetro de Microondas de Escaneo Multicanal
- DMSP Programa de Satélites Meteorológicos de Defensa
- TRMM Misión para Mediciones de Precipitaciones Tropicales
- NASDA Agencia Nacional para el Desarrollo Espacial de Japón
- JAXA Agencia Aeroespacial y de Exploración de Japón
- TMI Sensor Imager de Microondas

-
- SSM/I Sensor Especial de Microondas Imager
 - PR Precipitación
 - VIS Región Visual del Espectro Electromagnético
 - IR Región Infrarojo del Espectro Electromagnético
 - BT Temperatura de Brillo
 - MI Imager de Microondas
 - MS Sonda de Microondas
 - NOAA Administración Nacional de Estados Unidos de América para la Atmósfera y el Océano
 - FFMLP Feed Forward Multilayer Perceptron
 - NEXRAD Sistema de Radars Meteorológicos de alta resolución Doppler de Nueva Generación
 - NWS Servicio Nacional Meteorológico de Estados Unidos de América
 - PC Componentes Principales
 - SEPS Sistemas Satelitales de Estimación de Precipitación

APÉNDICE A

COMPILACIÓN Y EJECUCIÓN DEL ALGORITMO DE ESTIMACIÓN DE PRECIPITACIÓN BAJO AMBIENTE LINUX

Este apéndice contiene un breve manual para el uso del algoritmo desarrollado, lo cual tiene por objeto ser un tutorial que contemple las instrucciones y secuencias necesarias a seguir desde el momento en que se descargan los datos del sensor *ATMS* hasta obtener la salida de los mapas de precipitación.

Es importante mencionar que el algoritmo se basa en 2 librerías de C++¹, cuyo fichero raíz es */home/user/*, con lo cual esta ruta será considerada como absoluta en todos los casos. Por otro lado, los módulos en C++ y en Bash/UNIX que conforman el algoritmo están ubicados en la ruta absoluta */home/user/Documentos/Pruebas_Teg_IG/SRC/*, desde la cual se referencian todos los módulos del algoritmo².

¹*ALGLIB* y *HDF5*

²El directorio **user**, corresponde al nombre de sesión de su usuario en su computador

A.1. Descarga de Datos ATMS

La primera etapa para la ejecución del algoritmo esta relacionada con la descarga de los datos de los 22 canales del sensor ATMS, lo cuales son adquiridos a partir del portal de internet http://www.nsof.class.noaa.gov/saa/products/search?sub_id=0&datatype_family=ATMS&submit.x=32&submit.y=7. En la figura A.1, se muestra la visión general que se tiene al ingresar a la interfaz para descargar datos TDR (*Temperature Data Record, por sus siglas en inglés*) provenientes del sensor ATMS.

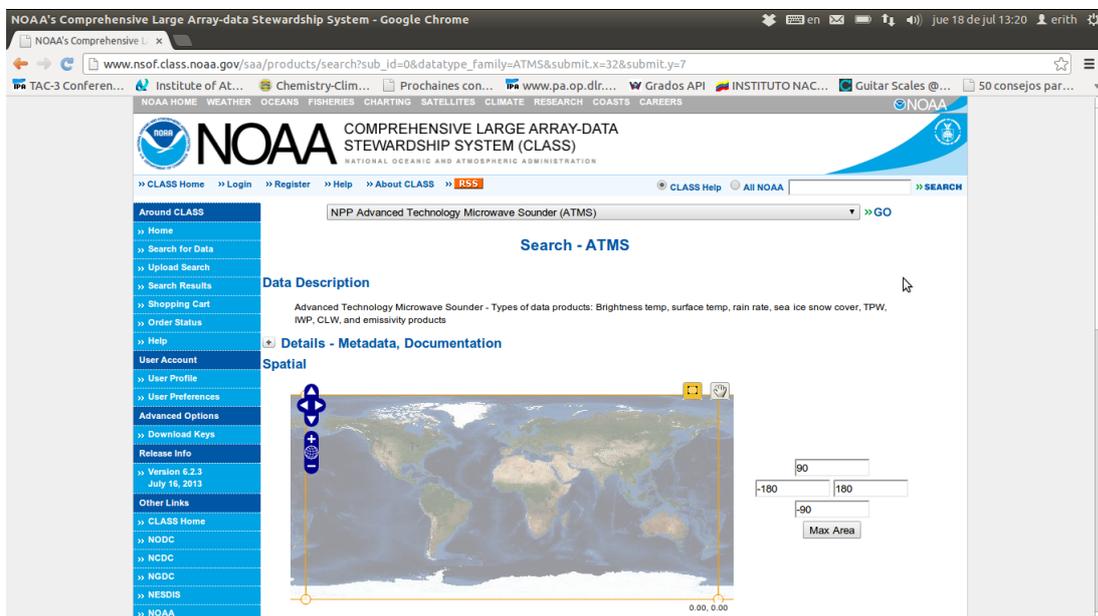
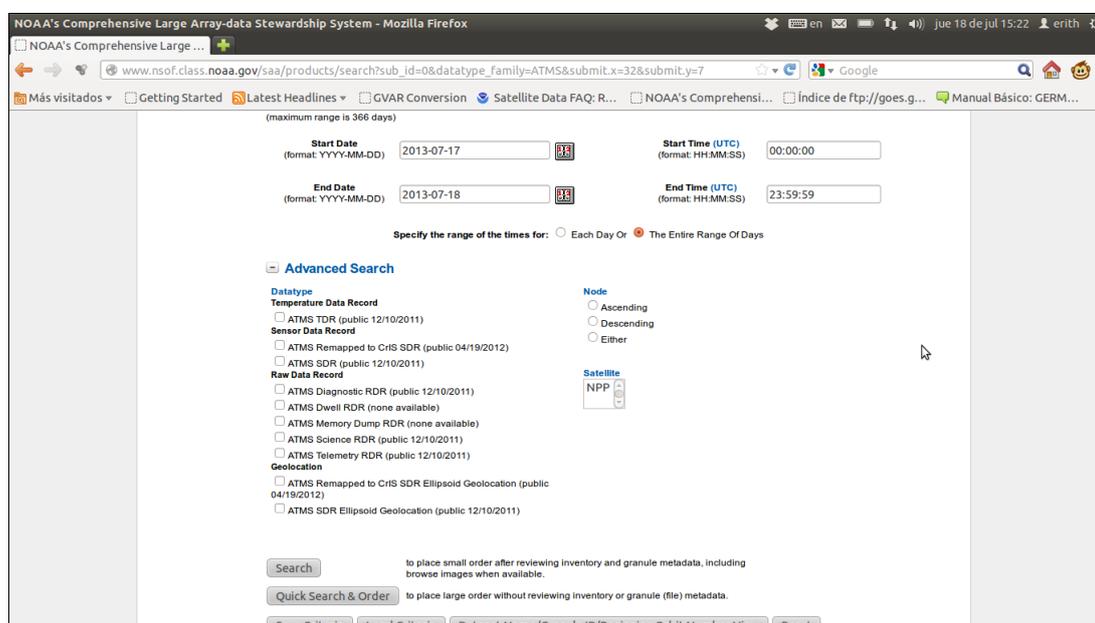


Figura A.1: Interfaz proporcionada por el NOAA-CLASS para la adquisición de datos ATMS, parte superior

A partir de la figura A.1, se aprecia un mapa a través del cual puede seleccionarse una cobertura geográfica de interés para la adquisición de los datos, o también del lado derecha se muestran campos donde al ingresar el rango en longitud y latitud de la zona de interés se puede lograr con mayor precisión la tarea de seleccionar la región de interés para los datos. Luego en la parte inferior del mismo portal se debe insertar el rango de fecha para los datos, si se desean capturas ascendentes o descendentes, el producto que en este caso es el TDR y las opciones para iniciar la búsqueda tal y como lo muestra la figura A.2.

Ahora bien, una vez ingresada la región de interés y el rango de fechas, la aplicación proporciona una nueva interfaz donde se puede apreciar en mapas las coberturas de cada conjunto de datos y a partir de aquí se tiene la opción para descartar datos que no resulten útiles si es el caso. Por otro lado, es importante mencionar que para poder descargar los datos es necesario registrarse al sistema CLASS para recibir un usuario y contraseña para poder adquirir los datos. Una vez que se confirma el pedido de datos, se reciben dos correos por parte del sistema CLASS, el primero confirmando que ha sido recibido el pedido, y el segundo presentando un resumen de los datos disponibles y una dirección ftp para la descarga de los datos.



The screenshot shows the NOAA's Comprehensive Large Array-data Stewardship System (CLASS) search interface. The browser address bar shows the URL: `www.nsof.class.noaa.gov/saa/products/search?sub_id=0&datatype_family=ATMS&submit.x=32&submit.y=7`. The page title is "NOAA's Comprehensive Large Array-data Stewardship System - Mozilla Firefox".

The search form includes the following fields and options:

- Start Date** (format: YYYY-MM-DD): 2013-07-17
- End Date** (format: YYYY-MM-DD): 2013-07-18
- Start Time (UTC)** (format: HH:MM:SS): 00:00:00
- End Time (UTC)** (format: HH:MM:SS): 23:59:59

Below the date and time fields, there is a radio button selection: "Specify the range of the times for: Each Day Or The Entire Range Of Days".

The **Advanced Search** section includes several categories of data records:

- Datatype**
 - Temperature Data Record
 - Sensor Data Record
 - ATMS TDR (public 12/10/2011)
 - ATMS Remapped to CrIS SDR (public 04/19/2012)
 - ATMS SDR (public 12/10/2011)
 - Raw Data Record
 - ATMS Diagnostic RDR (public 12/10/2011)
 - ATMS Dwell RDR (none available)
 - ATMS Memory Dump RDR (none available)
 - ATMS Science RDR (public 12/10/2011)
 - ATMS Telemetry RDR (public 12/10/2011)
 - Geolocation
 - ATMS Remapped to CrIS SDR Ellipsoid Geolocation (public 04/19/2012)
 - ATMS SDR Ellipsoid Geolocation (public 12/10/2011)
- Node**
 - Ascending
 - Descending
 - Either
- Satellite**
 - NPP

At the bottom of the search form, there are buttons for "Search", "Quick Search & Order", "Save Criteria", "Load Criteria", "Dataset Name/Granule ID/Beginning Orbit Number View", and "Reset".

Figura A.2: Interfaz proporcionada por el NOAA-CLASS para la adquisición de datos ATMS, parte inferior

A.2. Lectura de Ficheros HDF5, conversión a Temperatura de Brillo y Creación de archivos de datos de entrada para el algoritmo

La siguiente fase a realizar es la lectura de cada uno de los ficheros en formato HDF5 descargados desde el NOAA-CLASS, donde cada archivo corresponde a un instante de tiempo en el cual el satélite tiene pasada dentro del rangos de horas que se ha solicitado descargar, la figura A.3 se muestra una vista general del tipo de archivos que se obtiene, cuyos nombres indican entre otras características el sensor de procedencia, la fecha de captura de los datos, hora de captura, etc.

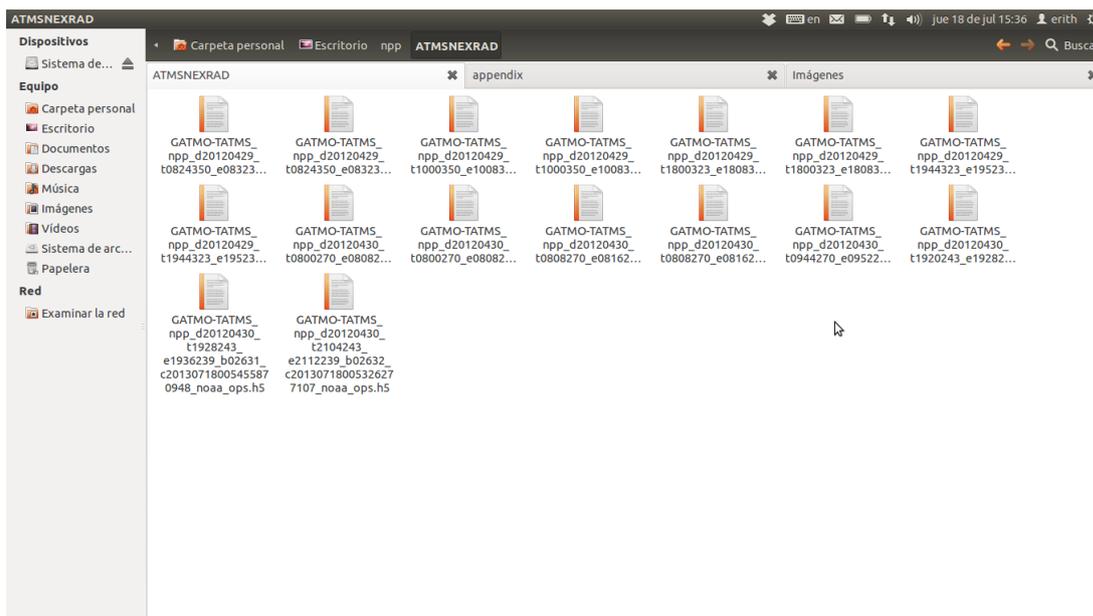


Figura A.3: Visión general de una carpeta contentiva de archivos de datos ATMS descargados via ftp a partir del sistema CLASS de la NOAA

En está etapa, la conversión de los números digitales (*DN*, por sus siglas en inglés) de datos a Temperatura de Brillo (*BT*) puede ser llevada a cabo de dos formas. Las mismas se explican a continuación:

A.2.1. Conversión a BT de los 22 Canales

Mediante este módulo, es posible generar un archivo .dat conformado por 26 columnas, donde las primeras 2 son índices correlativos para facilitar la graficación en lenguaje *IDL*, las siguientes dos son la latitud y longitud respectivamente y 22 columnas más, donde cada una de estas columnas corresponde a la *BT* para cada uno de los 22 canales que conforman el sensor *ATMS* y ordenados en forma ascendente. Esta modalidad es útil si se requiere graficar los datos en *BT* correspondientes a cada canal para la región de interés y obtener gráficas como las presentadas en la figura 3.5. Es importante que esta operación sea llevada a cabo para cada archivo que conforme el conjunto de datos *ATMS* descargado, de tal forma que el número de ficheros .dat generados sea igual al número de archivos .h5 descargados.

Para la ejecución de este módulo es necesario primeramente ubicarnos con una terminal en la ruta contentiva del archivo `readBTn_channels.cpp` escribiendo en la terminal la orden:

```
cd /Home/user/Documentos/Pruebas_Teg_IG/SRC 3
```

La compilación del archivo `readBTn_channels.cpp` se realiza mediante el comando:

```
h5c++ readBTn_channels.cpp -o readBTn_channels.exe
```

Donde *h5c++*, es un archivo .sh proporcionado por la librería *HDF5* para facilitar la compilación de archivos que incluyan esta librería, y `readBTn_channels.exe` es el nombre que se le asigna al ejecutable derivado del proceso de compilación. Terminada la fase de compilación, para llevar a cabo el proceso de ejecución es necesario editar el archivo `Lanzador_HDF5_DAT_Open_NN_channel.sh`, el cual ejecuta de forma automática al ejecutable `readBTn_channels.exe`, para todos los archivos que conformen el conjunto de datos *ATMS*. A continuación se muestra el script contenido en el archivo de ejecución.

```
#!/bin/csh

set INDIR = /home/erith/Documentos/Pruebas_Teg_IG/Julio2013/ATMS4292012/
set OUTDIR = /home/erith/Documentos/Pruebas_Teg_IG/Julio2013/ATMS4292012/outputs_all_channels/
```

³Desde esta ruta se ejecutan todos los módulos del algoritmo, de tal forma que cada vez que se compile o ejecuten códigos se sobreentiende que la terminal UNIX esta posicionada sobre esta ruta.

```
if (! -d $OUTDIR) mkdir -p $OUTDIR
cd $INDIR
set inplist = `ls -1 *ATMS*.h5`
foreach inpfiler ( $inplist )
    echo $inpfiler

    set file = `echo $inpfiler | awk '{print substr($0,0,100)}'`
    echo $file
    set outfile = $file'.dat'
    /home/erith/Documentos/Pruebas_Teg_IG/SRC/readBTn_channels.exe $inpfiler $OUTDIR$outfile

end
exit
```

En este código lo que normalmente debe editarse es, la ruta hacia los archivos de entrada *ATMS*, mediante la variable *INDIR* y la ruta donde se guardarán los archivos *.dat* de salida mediante la variable *OUTDIR*. La ejecución de este script shell se realiza mediante la orden:

`./Lanzador_HDF5_DAT_Open_NN_channel.sh`

El resultado de la ejecución puede apreciarse en la figura [A.4](#), en donde además puede notarse que el número de archivos *.dat* creados se corresponde que el número de archivos *ATMS* descargados y también que fueron guardados en la ruta que se asignó en la variable *OUTDIR*.

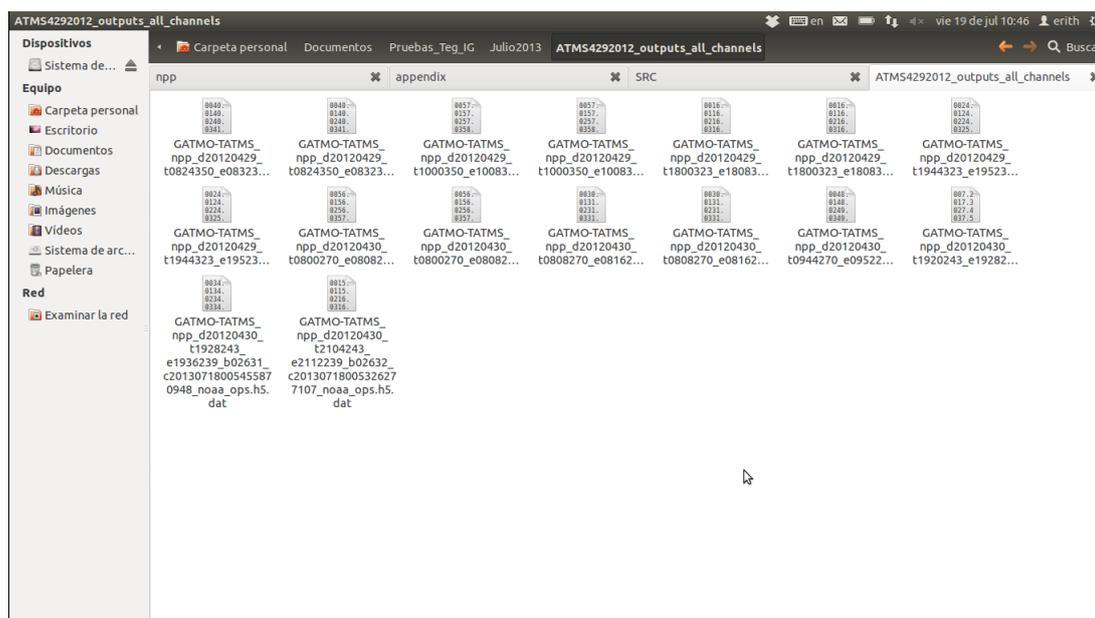


Figura A.4: Visión general de la carpeta contentiva de los archivos .dat generados a partir de la ejecución del script Lanzador_HDF5_DAT_Open NN_channel.sh

A.2.2. Conversión a BT de canales pre-seleccionados

En este caso a diferencia de la generación de los 22 canales, los archivos .dat son creados considerando únicamente los canales que sirven de entrada a la red neuronal que tenga lugar. En este sentido cabe destacar la disponibilidad de 4 archivos, uno para cada condición de corrida del algoritmo, esto es:

1. readBTn_channel_NNinput_LS.cpp (Tierra-Verano)
2. readBTn_channel_NNinput_LW.cpp (Tierra-Invierno)
3. readBTn_channel_NNinput_OS.cpp (Océano-Verano)
4. readBTn_channel_NNinput_OW.cpp (Océano-Invierno)

A modo ilustrativo a continuación se compila y ejecuta como ejemplo el código para el caso de Tierra-Invierno. La secuencia de comando es:

```
h5c++ readBTn_channel_NNinput_LW.cpp -o readBTn_channel_NNinput_LW.exe
```

```
./Lanzador_HDF5_DAT_Open_NN_channel_NNinput.sh
```

Donde se han configurado las variables *INDIR* y *OUTDIR* para la ejecución del script `Lanzador_HDF5_DAT_Open_NN_channel_NNinput.sh`. El resultado de la ejecución es un conjunto de archivos `.dat` cuyo contenido contiene igual número de columnas como canales de entrada a la red neuronal y al final una columna de *unos*, que guarda espacio en memoria para algunas tareas llevadas a cabo durante el proceso de evaluación en la red neuronal. Cada columna en estos archivos expresa en temperatura de brillo los datos *ATMS* para los canales de entrada a la red neuronal.

A.3. Cálculo de la tasa de precipitación mediante la red neuronal

El código de producción de la red neuronal para el caso de un escenario simple se encuentra en la ruta `/home/user/Open_NN/examples/Rainrate` y su nombre es **process.cpp**. Los archivos de compilación de **process.cpp**, fueron realizados en *CMAKE* ya que en esta se encuentran integrada bajo una misma compilación las librerías *ALGLIB*, *Open NN* y *HDF5*, si bien es cierto que la librería *Open NN* fué descartada para el entrenamiento de las redes neuronales usadas en el algoritmo, es importante descartar que esta librería proporciona herramientas útiles para la disposición y manejo de los datos en relación al proceso de entrenamiento y producción de la red neuronal, por esta razón esta librería ha sido usada y compilada con el código de producción **process.cpp**.

Para mayor comodidad este código ha sido previamente compilado⁴ para los 4 escenarios de estudio, y sus ejecutables son respectivamente:

1. `process_LS.exe`

⁴Para mayor información sobre la compilación de `process.cpp`, solicitar información a erith7@gmail.com

2. process_LW.exe
3. process_OS.exe
4. process_OW.exe

Ahora bien, la ejecución de *process_LW.exe* es llevada a cabo de forma automática mediante el script Lanzador_OPENN_RR_process_LW.sh, donde es necesario configurar las variables *INDIR* y *OUTDIR* para la obtención del conjunto de archivos .dat, cuyo unico contenido es la tasa de precipitación del escenario asociado a cada archivo *ATMS* procesado.

A.4. Elaboración de los archivos ascii para la generación de mapas de precipitación

La siguiente fase consiste en tomar los archivos .dat que contienen la tasa de precipitación como salida de la red neuronal y escribir un archivo ascii para cada caso, el cual debe contener 2 columnas con índices de referencia para la graficación en lenguaje *IDL*, una columna para la latitud, 1 columna para la longitud y 1 columna con la tasa de precipitación.

El módulo creado con este fin tiene por nombre **imagerRR.cpp**, y es necesario un paso de edición como paso previo a su compilación. En este sentido en la línea 89 del código aparece el siguiente fragmento:

```
to=strcpy(to, "/home/erith/Documentos/Pruebas_Teg_IG/Julio2013/MODIFICAR/");  
to=strcat(to, argv[1]);  
to=strcat(to, ".dat.dat");
```

Donde aparece la palabra **MODIFICAR**, debe sustituirse por el nombre de la carpeta que contiene los archivos con tasa de precipitación provenientes de las salidas de las redes neuronales. Con lo cual, luego del proceso de edición quedaría en la forma:

```
to=strcpy(to, "/home/erith/Documentos/Pruebas_Teg_IG/Julio2013/NNprocessATMS4292012/");  
to=strcat(to, argv[1]);  
to=strcat(to, ".dat.dat");
```

Finalizada la fase de edición se procede a la compilación, y posteriormente a la ejecución mediante el script *Lanzador_RR_Imager.sh* con las variables *INDIR* y *OUTDIR* redefinidas, la secuencia de comandos se muestra a continuación, a partir de la cual los archivos *ascii*, necesarios para generar los mapas de precipitación, son guardados en el directorio definido en *OUTDIR*:

```
h5c++ imagerRR.cpp -o imagerRR.exe
```

```
./Lanzador_RR_Imager.sh
```

A.5. Generación de los mapas de precipitación

Los mapas de precipitación son realizados mediante *IDL* a partir del código **mapeoRR.pro**, en el cual es necesario editar la ruta donde se encuentran almacenados los archivos de entrada y el número de archivos *ascii* de entrada para la elaboración de los mapas. Este archivo puede ser ejecutado en Linux mediante el comando:

```
idl -e “.r ./mapeoRR.pro”
```

A.6. Tiempo Estimado de Ejecución del Algoritmo

El proceso de ejecución del algoritmo de estimación de precipitación es realizado completamente en un tiempo inferior a 10 minutos, contando con la pericia necesaria para seguir los pasos antes descritos. A modo de referencia es importante mencionar que para generar 92 mapas de precipitación, lo que implica el procesamiento de 92 archivos H5 descargados de la herramienta NOAA-CLASS, fue requerido un tiempo inferior a 10 minutos, de los cuales aproximadamente 3 minutos fueron empleados exclusivamente a la etapa de graficación de mapas mediante *IDL*.