

**MODELO CON RETRASO Y CORRELACIÓN CRUZADA PARA
PREDECIR LA INTENSIDAD DE LA LLUVIA A CORTO PLAZO**

Por

Roberto Carlos Trespalacios Alias

Tesis sometida en cumplimiento parcial de los requerimientos para el grado de

MAESTRÍA EN CIENCIAS

en

MATEMÁTICA ESTADÍSTICA

UNIVERSIDAD DE PUERTO RICO
RECINTO UNIVERSITARIO DE MAYAGÜEZ

2012

Aprobada por:

Edgar Acuña Fernández, Ph.D
Miembro, Comité Graduado

Fecha

Dámaris Santana Morant, Ph.D
Miembro, Comité Graduado

Fecha

Nazario Ramírez Beltrán, Ph.D
Presidente, Comité Graduado

Fecha

Yolanda Ruiz Vargas, Ph.D
Representante de Estudios Graduados

Fecha

Omar Colón Reyes, Ph.D
Director del Departamento

Fecha

Resumen de Disertación Presentado a Escuela Graduada
de la Universidad de Puerto Rico como requisito parcial de los
Requerimientos para el grado de Maestría en Ciencias

**MODELO CON RETRASO Y CORRELACIÓN CRUZADA PARA
PREDECIR LA INTENSIDAD DE LA LLUVIA A CORTO PLAZO**

Por

Roberto Carlos Trespalacios Aliés

2012

Consejero: Nazario Ramírez Beltrán Ph.D

Departamento: Ciencias Matemáticas

El uso de los sistemas de computo, permitido el mejor acondicionamiento de las predicciones que se realizan de los fenómenos atmosféricos. Sin embargo, aún este tipo de predicciones es un reto que se les presenta a la mayoría de los científicos de las ciencias atmosféricas. Este trabajo busca encontrar un modelo que sea óptimo y construir una algoritmo rápido y eficaz. La correcta adecuación de los factores que intervienen en la predicción es otro problema; ya que por su accionar, éste se dispersa de manera súbita en periodos cortos y por tal razón, no es fácil conseguir una secuencia de datos que sean aplicables a distintos modelos de series temporales. En el trabajo se explora la posibilidad de utilizar modelos que son funcionales en fenómenos parecidos. Los vectores de desplazamiento, es otro problema que se aborda en este trabajo. Para esto, se utilizan métodos estadísticos robustos y de fácil aplicación, para encontrar solución a este problema.

Abstract of Dissertation Presented to the Graduate School
of the University of Puerto Rico in Partial Fulfillment of the
Requirements for the Degree of Master of Sciences

**MODEL FOR DELAY AND CROSS CORRELATION TO PREDICT
RAIN INTENSITY AT SHORT TERM**

By

Roberto Carlos Trespalacios Alias

2012

Chair: Nazario Ramírez Beltrán Ph.D
Major Department: Mathematical Sciences

The use of computer systems has permitted a better conditioning of the predictions made on atmospheric phenomena. However, these types of predictions still present themselves as a risk to the majority of the scientists in the atmospheric sciences. This work seeks to find an optimal model and construct an algorithm that is fast and efficient. The proper adjustment of the factors that intervene in the prediction is another problem; because of its actions, it disperses suddenly and in short periods of time. For this reason, it is not easy to find a sequence of data that will be applicable to different models of temporal series. This work will explore the possibility of using models that are functional in similar phenomena. The displacement vectors are another problem that abroad this work. For this, we utilize statistic robust methods and of fast application, to find the solution to this problem.

Copyright © 2012

por

Roberto Carlos Trespalacios Alies

A mis padres, hermanos y sobrinos...

Cimiento de mi vida.

AGRADECIMIENTOS

A mis padres María Cósima y Guillermo por brindarme siempre su apoyo y mostrarme los verdaderos valores del ser humano.

A mis hermanos Guillermo, Erik, Douglas, Carlos, Jorge y Daniela; por las alegrías que me brindan y por que nunca me abandonan.

Al Dr. Nazario Ramírez Beltrán, por sus aportes, sugerencias y ayuda en la ejecución de la tesis y por la paciencia e interés que tuvo en todo el tiempo en el cual se desarrollo este trabajo.

A Johan Castro, estudiante doctoral de Ingeniería Civil, por su trabajo en el algoritmo de estimación y sus aportes siempre útiles.

Al departamento de Ciencias Matemáticas del Recinto de Mayagüez y su director el Dr. Omar Colón.

A Madeline Ramos, Oficial administrativa del departamento de Ciencias Matemáticas; por su cordialidad y por su ayuda siempre oportuna.

A los profesores del recinto por brindarme su conocimiento; en especial a la Dra. Damaris Santana, de quien aprendí mucho.

A Karen por su amor y comprensión.

A mis amigos y compañeros del programa: Charlie Aragonés, Julio León, Fabián Seoanes, Luis Fuentes, Sindy Díaz, John Villavicencio, Velcy Palomino, Einstein Morales, Daiver Velez, Walter Quispe, Ysela Ochoa, Hector Moyano, Ricela Feliciano, Neysha Matos, Andrés Chamorro ; en especial Shirley Huanca, quien me ayudó en la parte final del trabajo. A todos muchas gracias.

Índice general

	<u>página</u>
RESUMEN	II
ABSTRACT	III
AGRADECIMIENTOS	VI
Índice de tablas	IX
Índice de figuras	X
LISTA DE ABREVIATURAS	XII
LISTA DE SÍMBOLOS	XIII
1. INTRODUCCIÓN	1
1.1. Motivación	2
1.2. Objetivos	3
1.2.1. Objetivo General	3
1.2.2. Objetivos Específicos	3
1.3. Limitaciones	3
2. ASPECTOS TEÓRICOS	5
2.1. Procesos Estocásticos	6
2.2. Procesos Estacionarios	6
2.3. Función de Autocovarianza	7
2.4. Correlación Cruzada	8
2.4.1. Función de Correlación Cruzada Muestral	10
2.5. Procesos Lineales Estacionarios	10
2.5.1. Secuencia Aleatoria y Ruido Blanco	11
2.5.2. Proceso Autorregresivo AR(p)	12
2.5.3. Procesos de Medias Móviles MA(q)	14
2.5.4. Procesos Autorregresivos de Medias Móviles ARMA(p,q)	15
2.5.5. Sistemas Lineales Discretos	17
2.5.6. Modelo de Función Estocástica de Transferencia	19
2.5.7. Modelo de Retraso en el Tiempo	24
2.5.8. Variograma	25
2.5.9. Método de Interpolación de Kriging	26
2.5.10. “Forward Selection” (Selección Hacia Adelante)	34
2.6. Técnicas de Procesamiento de Imágenes (Morfología)	35

2.7.	Medición y Cobertura del Radar Doppler	37
3.	METODOLOGIA	40
3.1.	Procesamiento de la Imagen	43
3.1.1.	Imágenes de Mayor Pixelado	43
3.1.2.	Eliminación de Regiones	44
3.2.	Vector de Movimiento de la Tormenta	44
3.2.1.	Correlación Cruzada e Imágenes	46
3.3.	Estimación de la Intensidad de Lluvia	50
4.	ANÁLISIS DE RESULTADOS DEL MODELO RT Y VECTOR DE DESPLAZAMIENTO	57
5.	CONCLUSIONES Y TRABAJOS FUTUROS	72
5.1.	Conclusiones	72
5.2.	Trabajos Futuros	73
	APÉNDICES	74
A.	ALGORITMOS	75
A.1.	Algoritmos para encontrar vectores de desplazamiento	75
A.1.1.	Función que interpola una matriz usando el método PMA.	75
A.1.2.	Función que encuentra la máxima reflectividad.	76
A.1.3.	Función que encuentra los elementos del vector velocidad.	76
A.1.4.	Función que genera el vector asociados a cada pixel.	77
A.1.5.	Función que genera los vectores de desplazamiento.	77
A.1.6.	Función que genera los gráficos de vectores desplazamiento de la nube	79
A.2.	Algoritmos para estimar la reflectividad	81
A.2.1.	Función que genera los gráficos de reflectividad estimada	81
A.2.2.	Función que genera la interpolación en las regiones de en- trada.	82
A.2.3.	Función para interpolación en las regiones.	84
A.2.4.	Función interpolación de Kriging.	85
A.2.5.	Función multicolumna.	86
A.2.6.	Función factor de la varianza.	87
A.2.7.	Función que selecciona las mejores variables para el modelo.	87
A.3.	Eventos de lluvias y tormentas	89

Índice de tablas

<u>Tabla</u>		<u>página</u>
3-1.	Construcción de las variables predictoras.	53
4-1.	Eventos usados para validación del modelo RT.	57
4-2.	Coefficiente de determinación R^2 para cada modelo RT	58
A-1.	Muestra de los eventos de lluvias y tormentas usadas en el trabajo. . .	89

Índice de figuras

<u>Figura</u>	<u>página</u>
2-1. Esquema sistema lineal simple	18
2-2. Esquema sistema con filtro lineal dinámico	22
2-3. Interpolación de Kriging	27
2-4. Kriging ordinario.	30
2-5. Medición del radar Doppler.	38
2-6. Radar Doppler (WSR-88D).	39
3-1. Diagrama de flujo para los procesos.	43
3-2. Movimiento de las partículas de lluvia.	45
3-3. Desplazamiento de la nube.	48
3-4. Correlación cruzada del rastreador.	48
3-5. Construcción de las variables del modelo RT	52
4-1. Reflectividad observada, para el evento del 2003/04/17 a las 00:45 horas en PR.	59
4-2. Reflectividad observada y estimada 15 min hacia adelante, para el evento del 2003/04/17 a las 01:00 horas en PR.	60
4-3. Reflectividad observada y estimada 30 min hacia adelante, para el evento del 2003/04/17 a las 01:15 horas en PR.	60
4-4. Reflectividad observada y estimada 60 min hacia adelante, para el evento del 2003/04/17 a las 01:45 horas en PR.	61
4-5. Reflectividad observada y estimada 120 min hacia adelante, para el evento del 2003/04/17 a las 02:45 horas en PR.	61
4-6. Contraste de reflectividad y validación cruzada para los valores ob- servados y estimados 15 min hacia adelante, para el evento del 2003/04/17 a las 01:00 horas en PR.	62

4-7. Contraste de reflectividad y validación cruzada para los valores observados y estimados 30 min hacia adelante, para el evento del 2003/04/17 a las 01:15 horas en PR.	63
4-8. Contraste de reflectividad y validación cruzada para los valores observados y estimados 60 min hacia adelante, para el evento del 2003/04/17 a las 01:45 horas en PR.	63
4-9. Contraste de reflectividad y validación cruzada para los valores observados y estimados 120 min hacia adelante, para el evento del 2003/04/17 a las 02:45 horas en PR.	64
4-10. Vectores de desplazamiento de la nube, para el evento del 2003/04/17 a las 00:45 horas en PR.	65
4-11. Reflectividad observada, para el evento del 2007/10/28 a las 00:00 horas en PR.	66
4-12. Reflectividad observada y estimada 15 min hacia adelante, para el evento del 2007/10/28 a las 12:15 horas en PR.	66
4-13. Reflectividad observada y estimada 30 min hacia adelante, para el evento del 2007/10/28 a las 12:30 horas en PR.	67
4-14. Reflectividad observada y estimada 60 min hacia adelante, para el evento del 2007/10/28 a las 00:00 horas en PR.	67
4-15. Reflectividad observada y estimada 120 min hacia adelante, para el evento del 2007/10/28 a las 02:00 horas en PR.	68
4-16. Contraste de reflectividad y validación cruzada para los valores observados y estimados 15 min hacia adelante, para el evento del 2003/12/7 a las 12:15 horas en PR.	68
4-17. Contraste de reflectividad y validación cruzada para los valores observados y estimados 30 min hacia adelante, para el evento del 2003/12/7 a las 12:30 horas en PR.	69
4-18. Contraste de reflectividad y validación cruzada para los valores observados y estimados 60 min hacia adelante, para el evento del 2003/12/7 a las 1:00 horas en PR.	69
4-19. Contraste de reflectividad y validación cruzada para los valores observados y estimados 120 min hacia adelante, para el evento del 2003/12/7 a las 2:00 horas en PR.	70
4-20. Vectores de desplazamiento de la nube, para el evento del 2007/10/28 a las 00:00 horas en PR.	71

LISTA DE ABREVIATURAS

dBz	Decibeles.
DE	Decremento Exponencial.
FET	Función Estocástica de Transferencia.
RT	Modelo de Retraso en el Tiempo.
AR(p)	Proceso Autorregresivo de Orden p.
MA(q)	Proceso de Medias Móviles de Orden q.
ARMA(p,q)	Proceso Autorregresivo de Media Móvil de Orden (p,q).

LISTA DE SÍMBOLOS

\mathbb{Z}	Conjunto de los números enteros.
\mathbb{R}	Conjunto de los números reales.
$N(\mu, \sigma^2)$	Distribución normal con media μ y varianza σ^2 .
Ω	Espacio muestral.
T_0	Espacio paramétrico.
$\rho_X(\cdot)$	Función de autocorrelación.
$\gamma_X(\cdot)$	Función de autocovarianza.
$\hat{\rho}_{yx}(\cdot)$	Función de correlación cruzada muestral.
$\mathcal{Y}(h)$	Función variograma.
$M_t^{i,j}$	Matriz de dimensión cuadrada, con centro en (i, j) y en el tiempo t .
$B(\cdot)$	Operador de retraso.
Z_t	Reflectividad en el tiempo t .
\mathcal{F}	Sigma álgebra.
A'	Transpuesta de la matriz A .
$z_t^{i,j}$	Valor de reflectividad para z , con centro en (i, j) en el tiempo t .

Capítulo 1

INTRODUCCIÓN

La definición de *pronóstico inmediato* o *nowcasting* ha evolucionado a través de los años. Browning, 1982. [1] dice que es una combinación de descripción del tiempo presente más extrapolación de sus tendencias hasta 2 horas hacia delante. En Browning & Conway, 1987. [2], se interpreta como el pronóstico detallado hasta 6 horas, con refinamientos de hasta 12 horas, haciéndose cada vez más borrosa su distinción del pronóstico a muy corto plazo.

A pesar de los grandes avances de las dos últimas décadas en la predicción numérica operacional del tiempo, es muy probable que el pronóstico inmediato mantenga un importante rol en el futuro inmediato de la vigilancia meteorológica, ver Li & Lai, 2004. [3]. Esto se debe a la poca habilidad de la modelación numérica en las primeras horas, producto principalmente del tiempo requerido para el procesamiento, análisis y asimilación de gran cantidad de datos, más el tiempo consumido propiamente en la simulación, podemos ver esto en Germann & Zawadzki, 2002 y 2004. [4], [5]. En Wilson et al, 1998. [6], la tendencia de las técnicas de nowcasting es la de incorporar, además del pronóstico del movimiento y evolución de las tormentas, la predicción cuantitativa de la cantidad de precipitación.

Las herramientas primarias para detectar precipitación son los radares del tiempo, los detectores de rayos y los satélites meteorológicos. Los primeros son los más empleados históricamente para estos fines y entre sus principales ventajas se encuentra la alta resolución espacial y temporal de sus datos, además de permitir la

observación tridimensional directa de su objetivo.

La estrategia que se pretende utilizar en este trabajo va encaminada hacia la obtención de un modelo estadístico que describa el movimiento y evolución de la lluvia y que a su vez sea robusto y eficiente para poder implementarlo de manera eficaz en algunos centros de predicción meteorológica; esto representaría una gran ventaja en términos de velocidad de cálculos y minimización de tiempo en la predicción, algo que es de suma importancia para los centros de prevención de desastres. Típicamente los modelos estadísticos son más precisos que los modelos numéricos para predicciones a corto plazo (Wilson et al, 1998. [6]). Sin embargo, cuando la predicción es a largo plazo los modelos numéricos son más eficientes.

1.1. Motivación

El estudio de los fenómenos atmosféricos ha sido siempre un tema lleno de constante trabajo y objeto de investigación por muchas entidades estatales y privadas como: universidades y centros de predicción meteorológica. Es bien conocido por todos, que la predicción de los fenómenos atmosféricos tales como las precipitaciones en forma de lluvia, puede ayudar de manera ostensible a la prevención de desastres tales como: inundaciones, avalanchas, deslizamientos; causando daños severos a la población y a la infraestructura y además, daños a los cultivos y en las telecomunicaciones (Tabla A-1). Existen en la literatura muchos resultados y avances sobre la predicción de las lluvias; la gran mayoría, siempre enfocados hacia la implementación de ecuaciones diferenciales que describan el comportamiento de la lluvia como si ésta tuviera las mismas propiedades de un fluido.

Algunos ejemplos de este tipo de resultados, para modelar el desplazamiento de la tormenta, que utilizan las ecuaciones de Lagrange y Laplace pueden verse en Germann & Zawadzki, 2002 y 2004. [4], [5]. El análisis estadístico de las mediciones de

radar y de satélite es de mucha utilidad, debido a que nos puede brindar información sobre la posibilidad de que llueva en una región muy extensa; la cual a su vez, puede ser útil para predecir el comportamiento de la lluvia a corto plazo y así poder mitigar el impacto de tormentas severas.

1.2. Objetivos

1.2.1. Objetivo General

Predecir la intensidad de la lluvia a intervalos de tiempo cortos y describir el desplazamiento de la lluvia.

1.2.2. Objetivos Específicos

- Aplicar la correlación cruzada como método para la obtención de un modelo para describir la dirección y magnitud del desplazamiento de la lluvia.
- Reconocer las variables adecuadas, que sean útiles en la elaboración de un modelo de predicción de la lluvia a corto plazo.
- Desarrollar un modelo estadístico que permita la predicción de la intensidad de la lluvia para intervalos de tiempo cortos y conocer el desplazamiento de la nube.
- Diseñar un algoritmo de predicción óptimo y eficaz, que permita la implementación en sistemas de cómputo; cuya lectura y comprensión de resultados sea clara, rápida y concisa.

1.3. Limitaciones

El modelo desarrollado en este trabajo, tiene algunas limitaciones las cuales se detallan a continuación.

El radar proporciona información muy detallada relativa al campo de lluvia a intervalos de tiempo pequeños (entre 5 y 15 minutos) y sobre grandes superficies; si bien es cierto esto permite utilizar esa información para realizar una previsión

de la evolución temporal del campo de lluvia y por lo tanto activar los sistemas de prevención de lluvias intensas; pero sólo en estimaciones a corto plazo.

Otro punto importante a tomar en cuenta es que en muchos casos el campo de lluvia no presenta un movimiento homogéneo, sino diferentes velocidades en zonas distintas de precipitación. Además los procesos de intensificación o decrecimiento de la precipitación son difíciles de modelar, por lo que, aunque que se conozca la dirección del campo, es complejo determinar la intensidad de la lluvia en el futuro, sobre todo en situaciones de lluvias intensas con presencia de núcleos convectivos.

Finalmente cabe resaltar que el problema principal en el estudio es que no se cuenta con mediciones relativamente continuas para periodos prolongados; lo cual hace más complicada la adecuación de un modelo con series temporales.

Capítulo 2

ASPECTOS TEÓRICOS

En esta primera parte, se hablará de aspectos mínimos, pertinentes y necesarios para la comprensión del uso de los modelos que serán discutidos y luego acondicionados o ajustados para luego ser utilizados en el desarrollo del trabajo.

En los últimos años, el análisis de series de tiempo, se ha convertido en un método poderoso para analizar conjuntos de datos, cuya característica fundamental, es que se hayan observado a través del tiempo y a intervalos iguales, así los datos usualmente son dependientes entre sí, y cuyos principales objetivos es describir el comportamiento típico de una serie de tiempo $X_t = P_t + T_t + E_t$, donde P_t es el componente periódico, T_t es el componente de tendencia y E_t es el componente estocástico; y realizar predicciones o interpolaciones lo más confiable posibles, sin importar el área de investigación en que se encuentra. Típicamente, P_t y T_t , se consideran componentes determinísticos en el modelo.

De otro lado, las características ineludibles de una serie temporal, obligaron a los académicos, a desarrollar nuevos y apropiados métodos para la inferencia estadística. La econometría fue la más privilegiada con el tratamiento de las series de tiempo, esto se vio reflejado cuando Engle, R.F. y Granger C.W., recibieron conjuntamente en el 2003, el premio Nobel en economía por avances que fueron dados en el área de la econometría, por haber desarrollado métodos de análisis temporales con tendencias comunes; esto se puede ver en sus trabajos: "*Autorregressive Conditional Heteroskedasticity with Estimates of United Kingdom Inflation (ARCH)*" (Engle, 1982. [7]) y "*Co-integration and Error Correction: Representation, Estimation and*

Testing” (Engle & Granger, 1987. [8]). En la década del setenta, surge un gran texto, cuyos autores, Box & Jenkins, 1976. [9] establecen los principios fundamentales sobre el tratamiento de las series de tiempo.

2.1. Procesos Estocásticos

En el análisis de series de tiempo es necesario seleccionar un modelo matemático adecuado para la secuencia observada, que nos permita conocer la naturaleza de los valores futuros ó predicciones; de esta manera suponemos que cada observación x_t es un valor realizado de cierta variable aleatoria X_t , por lo tanto, una serie de tiempo será el conjunto $\{x_t, t \in T_0\}$ es entonces una realización de la familia de variables aleatorias $\{X_t, t \in T_0\}$, donde T_0 es un espacio de parámetros.

Definición 2.1. *Un proceso estocástico es una familia de variables aleatorias $\{X_t, t \in T_0\}$ definido sobre un espacio de probabilidad (Ω, \mathcal{F}, P) , donde Ω es el espacio muestral, \mathcal{F} es un sigma álgebra y P es la medida de probabilidad.*

Definición 2.2. *Sea $\mathcal{F} = \{(t_1, \dots, t_n) \in T^n : t_1 < t_2 < \dots < t_n, n = 1, 2, \dots\}$, entonces, la función de distribución finita dimensional de $\{X_t, t \in T_0\}$ son las funciones $\{F_t(\cdot), \mathbf{t} \in \mathcal{F}\}$ definidas por $F_t(\mathbf{x}) = P(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n)$, donde $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$.*

2.2. Procesos Estacionarios

Definición 2.3. *Un proceso estocástico $\{X_t, t \in T\}$, se dice estrictamente estacionario, si la distribución conjunta de probabilidad de los vectores aleatorios $(X_{t_1}, X_{t_2}, \dots, X_{t_k})'$ y $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})'$ son iguales para todo $k \in \mathbb{Z}^+$ y todo $h, t_i \in T$, con $i \in \mathbb{Z}^+$. Esto es, $F_t(X_{t_1}, X_{t_2}, \dots, X_{t_k}) = F_t(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_k+h})$.*

Definición 2.4. *Un proceso estocástico $\{X_t, t \in T_0\}$ es estacionario en sentido débil si se cumplen las tres condiciones siguientes:*

- I. $E(X_t) = \mu$, para todo $t \in T$.
- II. $\text{var}(X_t) = \sigma^2$, para todo $t \in T$.
- III. $\text{Cov}(X_{t_i}, X_{t_j}) = \text{Cov}(X_{t_i+k}, X_{t_j+k})$ para todo $i \neq j$; donde μ y σ^2 son los parámetros de una distribución normal y k un número entero.

Definición 2.5. *El proceso $\{X_t, t \in T_0\}$ es débilmente estacionario de orden m , si todos sus momentos hasta orden m existen y no dependen del tiempo. También son llamados procesos estacionarios en covarianza. Lo anterior, implica que el proceso estocástico $\{X_t, t \in T_0\}$ tiene media constante, varianza constante, además la covarianza no depende del tiempo; sólo depende de la distancia a la que se toman las observaciones.*

Definición 2.6. *El proceso $\{X_t, t \in T_0\}$ es Gaussiano si y solamente si las funciones de distribución de $\{X_t, t \in T_0\}$, son todas normales multivariadas. Si un proceso Gaussiano es estacionario, entonces el proceso será estrictamente estacionario. En la parte restante de este documento se usará la siguiente notación: el parámetro t (tiempo) discreto, esto es, $t \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ y el proceso será escrito como $\{X_t, t \in \mathbb{Z}\}$.*

2.3. Función de Autocovarianza

La función de autocorrelación se define como la correlación de la señal consigo misma. La función de autocorrelación resulta de gran utilidad para encontrar patrones repetitivos dentro de una señal, como por ejemplo, la periodicidad de una señal enmascarada bajo el ruido o para identificar la frecuencia fundamental de una señal que no contiene dicha componente, pero aparecen numerosas frecuencias armónicas

de esta. De manera mas formal tenemos la siguiente:

Definición 2.7. Sea $\{X_t, t \in \mathbb{Z}^+\}$ un proceso estacionario y $h \in \mathbb{Z}^+$, entonces, definimos la función de autocovarianza como:

$$\gamma_X(h) = Cov(X_{t+h}, X_t). \quad (2.1)$$

Llamaremos a $\gamma_X(h)$, la función de autocovarianza con retardo h .

2.4. Correlación Cruzada

La función de correlación cruzada es una medida muy útil de la fuerza y la dirección de la relación entre dos variables aleatorias. Dados dos procesos estocásticos $\{X_t, t \in \mathbb{Z}\}$ y $\{Y_t, t \in \mathbb{Z}\}$, decimos que son conjuntamente estacionarios si ambos procesos invariantes estacionarios y la covarianza cruzada entre X_t y Y_t es una función de la diferencia temporal. La covarianza cruzada se define como:

$$\gamma_{YX}(h) = Cov(Y_t, X_{t-h}), \text{ para toda } h = 0, \pm 1, \pm 2, \dots \quad (2.2)$$

ó

$$\gamma_{XY}(h) = Cov(X_t, Y_{t-h}), \text{ para toda } h = 0, \pm 1, \pm 2, \dots \quad (2.3)$$

Las covarianzas cruzadas necesariamente son simétricas alrededor de 0. Por el contrario, es muy posible que X_t , cause Y_t y por lo tanto, esté fuertemente correlacionada con valores futuros de Y_t , de manera que $\gamma_{YX}(h) \neq 0$, $h > 0$ y que sin embargo no esté correlacionada con valores retardados de Y_t , es decir, $\gamma_{YX}(-h) = 0$, $h > 0$. Por lo tanto, en principio,

$$\gamma_{YX}(h) \neq \gamma_{YX}(-h), \quad (2.4)$$

sin embargo, se cumple que:

$$\gamma_{YX}(h) = \gamma_{XY}(-h). \quad (2.5)$$

Se deben usar series estacionarias ya que las covarianzas cruzadas sólo podrán interpretarse cuando se efectúen sobre series estacionarias. La siguiente es la interpretación de $\gamma_{YX}(h) = Cov(Y_t, X_{t-h})$.

- a. Un valor positivo de h en $\gamma_{YX}(h)$, denota retardo mientras que un valor negativo de h denota adelanto.
- b. En $\gamma_{YX}(h)$ el h escrito a la derecha de X , se refiere a los desfases (si h es positiva) o adelantos (si h es negativa) de la variable X respecto al período presente t de la variable Y . Así pues, si el valor presente de Y viene influenciado por dos valores retardados de la variable X , deberá cumplirse que $\gamma_{YX}(2) \neq 0$. Esto obviamente también implica que $\gamma_{YX}(-2) = 0$.

Dado que suponemos que la variable X puede influir en Y pero que la inversa no puede darse; esto es, existe causalidad unidireccional de X hacia Y , pero no a la inversa (no hay *feedback*). Estamos suponiendo de partida que solamente nos interesarán valores no negativos de h en $\gamma_{YX}(h)$. La función de correlación cruzada se define estandarizando la función de covarianzas cruzada, esto es,

$$\rho_{YX}(h) = \frac{\gamma_{YX}(h)}{\sigma_Y \sigma_X} \text{ para toda } h = 0, \pm 1, \pm 2, \dots \quad (2.6)$$

Siendo σ_X y σ_Y las desviaciones estándar de las variables aleatorias X y Y respectivamente. Dando valores a h , se obtiene la función de correlaciones cruzadas. Estas no son simétricas alrededor de 0, es decir, $\rho_{YX}(h) \neq \rho_{XY}(h)$. La función de correlación cruzada es útil para determinar los siguientes elementos:

- I. Si las series X_t y Y_t son estacionarias.
- II. Si X_t y Y_t muestran alguna relación entre ellas.

III. El tipo de relación que existe entre ambas series.

2.4.1. Función de Correlación Cruzada Muestral

Sea x_t y y_t una realización de series temporales X_t y Y_t , $1 \leq t \leq n$, entonces, la función de correlación cruzada:

$$\rho_{YX}(h) = \frac{\gamma_{YX}(h)}{\sigma_Y \sigma_X} \text{ para toda } h = 0, \pm 1, \pm 2, \dots, \quad (2.7)$$

se estima por la función de correlación cruzada muestral

$$\hat{\rho}_{yx}(h) = \frac{\hat{\gamma}_{yx}(h)}{S_y S_x} \text{ para toda } h = 0, \pm 1, \pm 2, \dots, \quad (2.8)$$

donde

$$\hat{\gamma}_{yx}(h) = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-h} (x_t - \bar{x})(y_{t+h} - \bar{y}), & \text{si } h \geq 0, \\ \frac{1}{n} \sum_{t=1-h}^n (x_t - \bar{x})(y_{t+h} - \bar{y}), & \text{si } h < 0, \end{cases} \quad (2.9)$$

y

$$S_x = \sqrt{\hat{\gamma}_{xx}(0)}, \quad S_y = \sqrt{\hat{\gamma}_{yy}(0)}. \quad (2.10)$$

Además, \bar{x} y \bar{y} , son las medias muestrales de las variables aleatorias x_t y y_t .

2.5. Procesos Lineales Estacionarios

Una gran familia de procesos lineales estacionarios paramétricos son utilizados con gran frecuencia en el estudio de las series de tiempo, los más usados son los procesos autoregresivos AR, medias móviles MA, y la combinación de estos dos, llamados procesos ARMA. Los cuales tienen un espacio importante dentro de la teoría de predicción lineal

A continuación se hará un repaso de algunos aspectos importantes concernientes a este tipo de procesos.

2.5.1. Secuencia Aleatoria y Ruido Blanco

Consideremos $\{X_t, t = 1, 2, \dots\}$ una secuencia de variables aleatorias definidas en el mismo espacio muestral Ω ; así, tenemos una secuencia aleatoria de parámetro discreto o simplemente un proceso estocástico discreto. Para todo $t \geq 1$ podemos escribir:

$$P(X_1 = a_1, \dots, X_t = a_t) = P(X_1 = a_1) \times P(X_2 = a_2 | X_1 = a_1) \quad (2.11) \\ \times \dots \times P(X_n = a_n | X_1 = a_1, \dots, X_{t-1} = a_{t-1}),$$

donde los a_i , representan el espacio de estados; que puede ser tomado como el conjunto de los números reales. El caso más simple es cuando tenemos una secuencia $\{X_t, t = 1, 2, \dots\}$ de variables aleatorias mutuamente independientes, entonces (2.12) se puede escribir como:

$$P(X_1 = a_1, \dots, X_t = a_t) = P(X_1 = a_1) \times P(X_2 = a_2) \times \dots \times P(X_t = a_t). \quad (2.12)$$

Si las variables aleatorias X_1, X_2, \dots , tiene todas la misma distribución, entonces tenemos una secuencia de variables aleatorias independientes e igualmente distribuidas (iid). En este caso, el proceso X_t , se llama *debilmente estacionario*, si $E(X_t) = \mu$ y $var(X_t) = \sigma^2$, para toda $t \geq 1$, entonces:

$$\gamma(h) = Cov(X_t, X_{t+h}) = \begin{cases} \sigma^2, & \text{si } h = 0, \\ 0, & \text{si } h \neq 0. \end{cases} \quad (2.13)$$

Definición 2.8. *El proceso $\{\varepsilon_t; t \in \mathbb{Z}^+\}$ se llama ruido blanco discreto, si las variables aleatorias ε_t son tales que $Cov(\varepsilon_{t_i}, \varepsilon_{t_j}) = 0$, para todo $t_i \neq t_j$.*

Si un proceso $\{\varepsilon_t; t \in \mathbb{Z}^+\}$, es tal que: $E(\varepsilon_t) = \mu$ y $var(\varepsilon_t) = \sigma^2$, para todo t y además su función de autocovarianza cumple la propiedad (2.13), entonces se llama estacionario. Queda claro que si ε_t son variables aleatorias independientes, entonces

también serán no correlacionadas. Una sucesión de variables aleatorias *iid* es llamada *proceso puramente aleatorio*. En adelante nos referiremos al proceso $\{\varepsilon_t; t \in \mathbb{Z}^+\}$ como ruido blanco; supondremos que $E(\varepsilon_t) = 0$, y se escribirá $\varepsilon_t \sim RB(0, \sigma^2)$. En el caso de que el proceso sea puramente aleatorio, entonces se escribirá $\varepsilon_t \sim iid(0, \sigma^2)$.

2.5.2. Proceso Autorregresivo AR(p)

Definición 2.9. *Un proceso estocástico $\{X_t; t = 1, 2, \dots\}$, se llamara autoregresivo de orden p , si satisface la ecuación de diferencias:*

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \phi_2(X_{t-2} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \varepsilon_t, \quad (2.14)$$

donde $\mu, \phi_1, \phi_2, \dots, \phi_p$, ($\phi_i \neq 0$ son los parámetros del modelo y $\varepsilon_t \sim RB(0, \sigma^2)$). De manera abreviada se escribe: $X_t \sim AR(p)$.

Si tomamos de la anterior definición la hipótesis de que $E(X_t) = \mu$, y además escribimos el proceso AR(p) de la forma:

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t, \quad (2.15)$$

entonces se tiene que $E(X_t) = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$. Definamos el operador de retardo por $B^r X_t = X_{t-r}$, para $r \geq 1$, $r \in \mathbb{Z}$; ahora, si se resta μ , en la ecuación (2.15), entonces, se puede escribir de la siguiente manera:

$$\phi_t(B)X_t = \varepsilon_t, \quad (2.16)$$

donde $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ es el operador autoregresivo de orden p . Al multiplicar por $\phi_t^{-1}(B)$ la ecuación (2.16), entonces la solución de (2.16) viene dada por:

$$X_t = \phi_t^{-1}(B)\varepsilon_t = \psi(B)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \quad (2.17)$$

donde $\psi(B) = 1 + \psi_1(B) + \psi_2(B) + \dots$. Note que, una de las condiciones necesarias para que la solución (2.17), sea estacionaria es que $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ y además puesto que $\phi(B)\psi(B) = 1$, entonces los coeficientes de ψ_j pueden ser encontrados en función de los ϕ_j por lo tanto, debemos suponer también que las raíces de $\phi(B) = 0$ deben encontrarse por fuera del círculo unidad; de manera equivalente, esto significa que $|\phi_j| < 1$, para $j = 1, 2, \dots, p$; lo que significa que los valores ϕ_j están dentro del círculo unitario. Podemos ver esto con mas detalle en Box & Jenkins, 1994. [10]. De otro lado, si en la ecuación (2.14) se multiplica por X_{t-h} y suponiendo estacionalidad, al tomar esperanza se tiene que:

$$\sigma_X^2 = \gamma(0) = \frac{\sigma^2}{1 - \phi_1\rho_1 - \phi_2\rho_2 - \dots - \phi_p\rho_p}, \text{ para } h = 0, \quad (2.18)$$

$$\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \dots + \phi_p\gamma(h-p), \text{ para } h > 0. \quad (2.19)$$

Las ecuaciones (2.18) y (2.19) son respectivamente: varianza y función de autocovarianza del procesos AR(p). De igual manera, para obtener la función de autocorrelación del proceso AR(p), se debe dividir la ecuación (2.19) por $\gamma(0)$; el resultado será:

$$\rho(h) = \phi_1\rho(h-1) + \phi_2\rho(h-2) + \dots + \phi_p\rho(h-p), \text{ para } h \geq p. \quad (2.20)$$

Sean $\rho_0, \rho_1, \rho_2, \dots, \rho_p$ las condiciones iniciales de (2.20), las cuales fueron determinadas a partir de los coeficientes $\phi_1, \phi_2, \dots, \phi_p$. La solución de la ecuación (2.20) permite calcular los valores de $\rho(h)$, para $h \geq p$. Tomando $h = 1, 2, \dots, p$ en (2.20), obtenemos el siguiente sistema de ecuaciones:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} \\ \rho_2 &= \phi_1\rho_1 + \phi_2 + \dots + \phi_p\rho_{p-2} \\ &\vdots \\ \rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \dots + \phi_p. \end{aligned} \quad (2.21)$$

El sistema (2.21) se puede escribir en forma matricial de la siguiente forma:

$$\boldsymbol{\phi}_p \mathbf{P}_p = \boldsymbol{\rho}_p. \quad (2.22)$$

Cuya solución viene dada por:

$$\boldsymbol{\phi}_p = \mathbf{P}_p^{-1} \boldsymbol{\rho}_p, \quad (2.23)$$

donde $\mathbf{P}_p = [\rho_{ij}]$ con $\rho_{ij} = \rho_{|i-j|}$, $i, j = 1, 2, \dots, p$; $\boldsymbol{\phi}_p = (\phi_1, \phi_2, \dots, \phi_p)'$ y $\boldsymbol{\rho}_p = (\rho_1, \rho_2, \dots, \rho_p)'$. La ecuación (2.23) es de mucha utilidad, ya que con ella se puede obtener estimadores de los parámetros ϕ_j ; esto se puede lograr sustituyendo las funciones de autocorrelación por sus estimados.

2.5.3. Procesos de Medias Móviles MA(q)

Definición 2.10. *Un proceso estocástico $\{X_t; t = 1, 2, \dots\}$, se llamara Procesos de Medias Móviles de orden q , si satisface la ecuación de diferencias:*

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (2.24)$$

donde $\mu, \theta_1, \theta_2, \dots, \theta_q$, son los parámetros del modelo y $\varepsilon_t \sim RB(0, \sigma^2)$. De manera abreviada, se escribe: $X_t \sim MA(q)$.

Se puede probar que el el proceso $X_t \sim MA(q)$, es estacionario y $E(X_t) = \mu$; además, puesto que los ε_t son no correlacionados, entonces, se tiene que:

$$var(X_t) = var(\mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}), \quad (2.25)$$

$$\sigma_X^2 = \sigma^2(1 + \theta_1 + \theta_2 + \dots + \theta_q). \quad (2.26)$$

Supongase ahora que $\mu = 0$, entonces la función de autocovarianza del proceso es:

$$\gamma(h) = \begin{cases} \sigma^2 \sum_{j=0}^{q-|h|} \theta_j \theta_{j+|h|}, & \text{si } |h| \leq q, \\ 0, & \text{si } |h| > q. \end{cases} \quad (2.27)$$

En la anterior ecuación $\theta_0 = 1$. Ahora, si dividimos $\gamma(0)$ entre σ_X^2 , obtenemos

$$\rho(h) = \frac{\gamma(0)}{\sigma_X^2}. \quad (2.28)$$

En particular, si tomamos $|h| > q$, entonces, se obtiene $\rho(h) = 0$. De otro lado, el proceso (2.24), puede ser escrito como:

$$X_t = \theta(B)\varepsilon_t, \quad (2.29)$$

donde $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ es el operador de media móvil o polinomio de media móvil de orden q . La solución para (2.29) sería:

$$\pi(B)X_t = \varepsilon_t, \quad (2.30)$$

en la cual, $\theta^{-1}(B) = \pi(B)$; por lo tanto,

$$\sum_{j=0}^{\infty} \pi_j X_{t-j} = \varepsilon_t, \text{ donde } \pi_0 = 1, \quad (2.31)$$

donde $\pi(B) = 1 + \pi_1 B + \pi_2 B^2 + \dots$, tan que $\sum_{j=0}^{\infty} |\pi_j| < \infty$, y puesto que $\theta^{-1}(B) = \pi(B)$, entonces, $\theta(B)\pi(B) = 1$; por lo tanto, para que el proceso sea invertible, las raíces de $\theta(B) = 0$ deberían estar por fuera del círculo unitario.

2.5.4. Procesos Autorregresivos de Medias Móviles ARMA(p,q)

Definición 2.11. Decimos que $\{X_t; t = 1, 2, \dots\}$, es un proceso autorregresivo de medias móviles de orden (p, q) , si X_t satisface la ecuación de diferencias:

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \phi_2(X_{t-2} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}, \quad (2.32)$$

donde $\mu, \phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$, son los parámetros del modelo y $\varepsilon_t \sim RB(0, \sigma^2)$.

De manera abreviada, este proceso se escribe: $X_t \sim ARMA(p, q)$.

Usando los operadores autorregresivo y de media móvil, definidos anteriormente, podemos escribir (2.32) de la siguiente forma:

$$\phi(B)\tilde{X}_t = \theta(B)\varepsilon_t, \quad (2.33)$$

donde $\tilde{X}_t = X_t - \mu$, de ahora en adelante se asumirá que $\mu = 0$; para un proceso ARMA(p,q), una condición de estacionariedad es igual a la necesaria para los procesos AR(p); esto es, que las raíces de $\phi(B) = 0$ deben estar fuera del círculo unitario; del mismo modo, una condición de invertibilidad es la misma que para los procesos MA(q); esto es, que las raíces de $\theta(B) = 0$, deben estar fuera del círculo unitario. la función de autocovarianza: Multiplicando (2.32) por X_{t-h} y si se considera que $\mu = 0$, entonces, al tomar esperanza se optiene la siguiente función de autocovarianza:

$$\gamma(h) = \phi_1\gamma(h-1) + \dots + \phi_p\gamma(h-p) + \gamma_{X_\varepsilon}(h) + \theta_1\gamma_{X_\varepsilon}(h-1) + \dots + \theta_q\gamma_{X_\varepsilon}(h-q), \quad (2.34)$$

donde $\gamma_{X_\varepsilon}(h)$ es la correlación cruzada de X_t y ε_t , la cual está definida por $\gamma_{X_\varepsilon}(h) = E(X_{t-h}\varepsilon_t)$. Como X_{t-h} depende únicamente de los ε_t para los instantes en los que $t_k > h - t$ para algún $k \in \mathbb{N}$, tenemos que esta covarianza cruzada sólo es diferente de cero para $h \leq 0$, luego:

$$\gamma(h) = \phi_1\gamma(h-1) + \phi_2\gamma(h-2) + \dots + \phi_p\gamma(h-p), \text{ para } h > q. \quad (2.35)$$

Lo anterior significa que las autocovarianzas y por tanto las autocorrelaciones, con sus respectivos retardos, los cuales son afectados por los parámetros de medias móviles; pero para $h > q$ estas tienen el mismo comportamiento que los modelos autorregresivos. Para complementar más los detalles de la función de autocorrelación para un proceso ARMA(p,q), puede ver Ramírez y Sastri, 1997. [11]. En esta publicación

se muestra que la función de autocorrelación para un proceso ARMA(p,q) es la función de los parámetros $\theta's$, $\phi's$ y σ^2 .

2.5.5. Sistemas Lineales Discretos

En muchas ocasiones, las investigaciones sobre la relación causal entre dos variables es a menudo el enfoque de los estudios físicos. En este sentido, los sistemas lineales han sido ampliamente utilizados en modelos que abordan problemas físico-atmosféricos. Básicamente, un sistema contiene un conjunto de elementos que están relacionados entre sí para cumplir una función particular. Cuando el sistema recibe un estímulo de entrada, este responderá a la señal de entrada de acuerdo con la estructura interna del sistema, es decir, la red de relaciones entre los elementos del sistema. La estructura interna del sistema, que regula la entrada y la salida, filtra la señal de entrada y posteriormente arroja la respuesta de salida. En términos de tiempo, la entrada en el tiempo $t - b$ actúa en el sistema, que pasa a través del filtro, y luego se transforma en la respuesta de salida en el tiempo t ; el número entero b se denomina respuesta de retraso; que es el intervalo de tiempo durante el cual el proceso de filtrado se lleva a cabo. El esquema de este concepto se presenta en la Figura 2-1. Si conocemos el mecanismo interno del sistema completamente, el filtro puede ser representado de manera determinista por ecuaciones diferenciales. Estas ecuaciones se pueden resolver bajo ciertos supuestos de condiciones ideales tales como la homogeneidad. Hay muchos modelos geográficos de este tipo, sobre todo en la geografía física. Entre las aplicaciones más desarrolladas destacan: clasificación de la cubierta superficial del suelo (i.e. Ahern et al., 1993; Lemoine et al., 1994), delimitación de la extensión y dinámica de los humedales (i.e. Morrisey et al., 1994), determinación de la variabilidad espacio-temporal del hielo en altas latitudes (i.e. Meade et al., 1999) o la estimación de la humedad del suelo (i.e. Wagner et al., 1999; Quesney et al., 2000). Sin embargo, las condiciones ideales utilizadas en los modelos

deterministas rara vez se aplican en la práctica. Los sistemas están afectados por perturbaciones aleatorias o no aleatorias que pueden ser causados por variaciones temporales y espaciales en el entorno subyacente en el que opera el sistema. Este problema nos lleva a un modelo de la relación en un estocástico o probabilístico con enfoque realista. La entrada y la salida del sistema son considerados como procesos estocásticos determinadas por leyes probabilísticas, que atañen a problemas de tipo espacial y/o temporales.



Figura 2–1: Relación entrada-salida de un sistema lineal simple.

En muchos de los problemas ambientales, en particular atmosféricos, se trabajan modelos que requieren sistemas en los cuales existen variables de entrada y de salida que se encuentran relacionadas con el tiempo. En el caso continuo, este tipo de problemas se abordan con ecuaciones diferenciales; y para el caso discreto, entonces, se aborda su estudio por medio de ecuaciones de diferencias. En este trabajo, se considerará el tiempo discreto únicamente; es decir, $t = 1, 2, \dots$. La forma del modelo de un sistema lineal discreto es:

a. Para una sola entrada.

$$Z_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots + \varepsilon_t, \quad (2.36)$$

b. Para varias entradas.

$$\begin{aligned} Z_t = v_{1,0} X_{1,t} + v_{1,1} X_{1,t-1} + \dots + v_{2,0} X_{2,t} + v_{2,1} X_{2,t-1} + \\ \dots + v_{k,0} X_{k,t} + v_{k,1} X_{k,t-1} + \dots + \varepsilon_t, \end{aligned} \quad (2.37)$$

donde Z_t , X_t y $X_{i,t}$; $i = 1, 2, \dots$, $t = 1, 2, \dots$, son series estacionarias con media cero y ε_t , es un proceso ARMA(p,q). La ecuación de respuesta de los mismos parámetros

v en las ecuaciones (2.36) y (2.37) son llamados los pesos de la respuesta impulso; estos parámetros, son análogas a los coeficientes de la regresión lineal. Si estos valores son 0 para $t > 1$, entonces la ecuación (2.37) se convierte en un modelo de regresión.

2.5.6. Modelo de Función Estocástica de Transferencia

El modelo de **Función Estocástica de Transferencia** (FET) puede definirse como una ecuación en diferencias estocástica, que se caracteriza por las funciones de respuesta al impulso que se relacionan con la dinámica de una variable de respuesta con las variables de entrada de un sistema dado. Se ha demostrado que una ecuación diferencial de primer orden se podría aproximar como una ecuación en diferencias de primer orden. Una discusión más clara sobre este tema, puede verse en el libro de Box y Jenkins, 1976. [9]. En Ramírez et al, 2008. [12], se desarrolla Una ecuación diferencial y luego se aproxima por un modelo de FET para encontrar, por ejemplo, una solución a un problema hidro-meteorológico. A continuación se discutirán los aspectos fundamentales del modelo FET. Consideremos una serie estocástica $Z(t)$ que se encuentra determinada por otra serie estocástica $X(t)$, ambas variables son continuas. El modelo de transferencia puede ser deducido de un sistema lineal dinámico; en el cual $X(t)$ es la variable de entrada y $Z(t)$ es la variable de salida. El modelo en tiempo discreto tiene la forma:

$$Z_t = \frac{w(B)}{\delta(B)} X_{t-r} + \varepsilon_t, \quad (2.38)$$

donde,

- Z_t es la variable respuesta en el tiempo t .
- X_{t-r} es la variable de entrada en el tiempo t y con retraso de r unidades de tiempo
- $w(B)$ es un polinomio de grado s .
- $\delta(B)$ es un polinomio de grado r .

- B es la variable de los polinomios w y δ que a su vez, es el operador de retraso de la forma $B^r x_t = x_{t-r}$, donde r es un entero positivo.
- ε_t es el ruido que sigue proceso autoregresivo de medias móviles (ARMA).

Notemos que para el modelo (2.38) los polinomios

$$w(B) = w_0 + w_1 B + w_2 B^2 + \dots + w_{s-1} B^{s-1} + w_s B^s, \quad (2.39)$$

y

$$\delta(B) = \delta_0 + \delta_1 B + \delta_2 B^2 + \dots + \delta_{r-1} B^{r-1} + \delta_r B^r, \quad (2.40)$$

los coeficientes $w_0, w_1, \dots, w_{s-1}, w_s, \delta_0, \delta_1, \dots, \delta_{r-1}$ y δ_r , son parámetros del modelo.

Multiplicando el polinomio $w(B)$ por X_{t-r} , tenemos que:

$$\begin{aligned} w(B)X_{t-r} &= (w_0 + w_1 B + w_2 B^2 + \dots + w_{s-1} B^{s-1} + w_s B^s)X_{t-r}, \\ &= w_0 X_{t-r} + w_1 B(X_{t-r}) + w_2 B^2(X_{t-r}) + \dots + w_{s-1} B^{s-1}(X_{t-r}) + w_s B^s(X_{t-r}), \\ &= w_0 X_{t-r} + w_1 X_{t-r-1} + w_2 X_{t-r-2} + \dots + w_{s-1} X_{t-r-s+1} + w_s(X_{t-r-s}). \end{aligned} \quad (2.41)$$

Si en la ecuación (2.38), multiplicamos ambos lados por $\delta(B)$, tenemos:

$$\delta(B)Z_t = w(B)X_{t-r} + n_t, \quad (2.42)$$

donde $n_t = \delta(B)\varepsilon_t$. La ecuación (2.42) quedaría:

$$(1 - \delta_1 B - \delta_2 B^2 - \dots - \delta_r B^r)Z_t = (w_0 - w_1 B - w_2 B^2 + \dots + w_s B^s)X_{t-r} + \varepsilon_t \quad (2.43)$$

ó

$$Z_t - \delta_1 Z_{t-1} - \delta_2 Z_{t-2} - \dots - \delta_s Z_{t-s} = w_0 X_{t-r} + w_1 X_{t-r-1} + \dots + w_s X_{t-r-s} + \varepsilon_t. \quad (2.44)$$

De la ecuación (2.44), podemos ver que Z_t depende de los valores de s actuales y pasados de la variable de entrada (en el tiempo $t, t-1, \dots, t-s$) y también de los

últimos valores de r de la variable de salida en sí. Ahora, recordemos que la ecuación (2.36) puede escribirse también:

$$Z_t = v(B)X_t + \varepsilon_t, \quad (2.45)$$

donde, $v(B) = v_0 + v_1B + v_2B^2 + \dots$, es llamada la función impulso respuesta. De lo anterior, vemos que (2.38) toma la forma:

$$Z_t = \frac{w(B)}{\delta(B)}B^r X_t + \varepsilon_t, \quad (2.46)$$

donde, $B^r X_t = X_{t-r}$. La ecuación (2.46) recibe el nombre de función de transferencia. De las ecuaciones (2.45) y (2.46), tenemos lo siguiente

$$v(B) = \frac{w(B)}{\delta(B)}B^r X_t, \quad (2.47)$$

de lo anterior, resulta:

$$v(B)\delta(B) = w(B)B^r. \quad (2.48)$$

De la ecuación (2.48), se puede observar que la función impulso respuesta $v(B)$, que es un polinomio de grado muy elevado (k), se puede expresar como una proporción de dos polinomios de orden s y r , respectivamente. Los autores Box & Jenkins, 1976. [9], sugieren que s y r son raramente mayor que 2; por lo tanto, un sistema lineal como (2.36) que originalmente estaba descrito por un gran número de parámetros, ahora puede ser representado por un modelo de función de transferencia con la participación de sólo unos pocos parámetros. El cociente $\frac{w(B)}{\delta(B)}$, es llamado función de impulso-respuesta, o también se le llama función de transferencia del sistema. El concepto de un modelo de función de transferencia es como pasar la serie de entrada a través de un filtro dinámico estocástico para generar la serie de salida. Esta idea se muestra en la Figura 2-2.

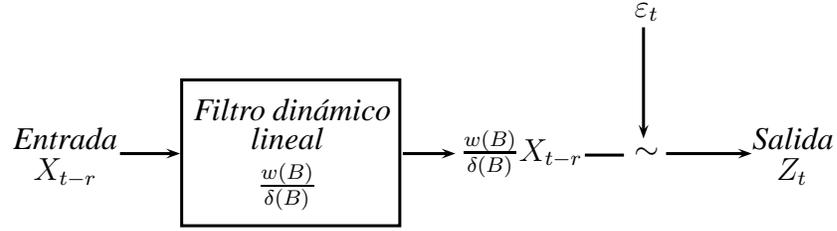


Figura 2-2: Relación entrada-salida representado por un filtro lineal dinámico, utilizando el enfoque de la función de transferencia.

En el caso de que existan múltiples entradas y salidas, entonces el modelo de función de transferencia toma la forma de la ecuación (2.37):

$$Z_t = v_1(B)B^{d_1}X_{1,t} + v_2(B)B^{d_2}X_{2,t} + \dots + v_J(B)B^{d_J}X_{J,t} + \varepsilon_t, \quad (2.49)$$

donde,

$$v_j(B) = \frac{w_j(B)}{\delta_j(B)} = \frac{w_{j,0} + w_{j,1}B + w_{j,2}B^2 + \dots + w_{j,s}B^s}{1 - \delta_{j,1}B - \delta_{j,2}B^2 - \dots - \delta_{j,r}B^r}, \quad (2.50)$$

$$\varepsilon_t = \frac{\Theta(B)}{\Phi(B)}a_t, \quad (2.51)$$

$$\Theta(B) = 1 + \theta_1B + \theta_2B^2 + \dots + \theta_qB^q, \quad (2.52)$$

$$\Phi(B) = 1 + \phi_1B + \phi_2B^2 + \dots + \phi_qB^q. \quad (2.53)$$

En la ecuación (2.49) se distinguen los siguientes elementos:

- $X_{j,t}$ es la j -ésima variable de entrada en el tiempo t ; $j = 1, 2, \dots, J$.
- a_t es una sucesión de variables aleatorias independientes con media cero y varianza constante.
- $v_j(B)$ es la función impulso respuesta para la j -ésima entrada de la variable X_{t-r} ; es decir, la función respuesta impulso es el cociente de los polinomios $w_j(B)$ y $\delta_j(B)$ proveniente de una serie temporal infinita.

Nótese que la serie infinita de la función de respuesta impulso converge cuando el sistema es asintóticamente estable (Pandit & Wu, 1983. [13]). La representación

racional de la función de transferencia es un modelo parsimonioso y la evaluación de la función de transferencia está dada por una división larga de dos polinomios, en el que sólo los valores más significativos se consideran (es decir, generalmente una serie mayor igual a 24 observaciones). La estabilidad asintótica de la función de transferencia se mide mediante el cálculo de las raíces del polinomio. El sistema es asintóticamente estable si las raíces quedan fuera del círculo unitario (Pandit & Wu, 1983. [13]).

Si en el modelo FET, consideramos en particular $r = s = 1$, entonces, tenemos

$$Z_t = \left(\frac{w_{1,0} + w_{1,1}B}{1 - \delta_{1,1}B} Z_{t-1} \right) + \left(\frac{w_{2,0} + w_{2,1}B}{1 - \delta_{2,1}B} Z_{t-2} \right) + \varepsilon_t, \quad (2.54)$$

donde,

- Z_t es la reflectividad en el pixel para el tiempo t .
- ε_t es el ruido que sigue proceso autoregresivo de medias móviles (ARMA) (Box & Jenkins, 1976. [9]).
- $\delta_{i,j}$ y $w_{i,j}$ son los parámetros de la función de respuesta impulso que se ha estimado por correlación cruzada.

La longitud de la memoria de la variable respuesta es controlada por los valores de δ . Por lo tanto, si δ se acerca a uno, el modelo tiene memoria larga y por lo tanto requiere muchas más observaciones anteriores para explicar la respuesta Z_t en el tiempo t . Por otro lado, si el δ se aproxima a cero, el modelo presentará memoria corta y sólo algunas observaciones recientes serán necesarias para expresar la respuesta del modelo en el tiempo t . Cada denominador de la ecuación (2.54) genera una serie geométrica, que convergen cuando el sistema es estable. Así, la expansión

de la serie geométrica de la ecuación (2.54), puede escribirse como sigue:

$$Z_t = (1 + \delta_{1,1}B + \delta_{1,1}^2B^2 + \dots)(w_{1,0} + w_{1,1}B)Z_{t-1} + \\ (1 + \delta_{2,1}B + \delta_{2,1}^2B^2 + \dots)(w_{2,0} + w_{2,1}B)Z_{t-2} + \varepsilon_t. \quad (2.55)$$

La anterior ecuación (2.55), nos brinda un primer enfoque a un modelo que se pueda implementar con el tipo de mediciones (datos) con las cuales contamos.

2.5.7. Modelo de Retraso en el Tiempo

Para desarrollar una rutina que sea mucho mas rápida en los cálculos, se implementa el modelo de **Retraso del Tiempo** (RT) como una alternativa a la ecuación (2.55). La aproximación se puede realizar después de ajustar los δ' s a cero; esta operación minimizaría la memoria del sistema. Cabe anotar, que el modelo RT requiere mucho menos esfuerzo computacional que el modelo FET. Además, el este modelo, RT, puede verse como una analogía al modelo de regresión lineal múltiple; y puede ser escrito de la siguiente forma:

$$Z_t = (w_{1,0} + w_{1,1}B) Z_{t-1} + (w_{2,0} + w_{2,1}B) Z_{t-2} + \varepsilon_t. \quad (2.56)$$

La representación del modelo general RT, se escribe así:

$$Z_t = \beta_0 + \beta_1 X_{1,t-d_1} + \beta_2 X_{2,t-d_2} + \dots + \beta_k X_{k,t-d_k} + \varepsilon_t, \quad d_i = 0, 1, 2, \dots, \quad (2.57)$$

donde Z_t , es reflectividad en el tiempo t ; $X_{i,t-d_i}$ son diferentes variables que pueden evaluarse en el modelo; k es el número total de variables que expresan mejor la reflectividad, β_i son los coeficientes del modelo RT y d_i es el retraso de la i -ésima variable de entrada. En nuestro caso particular, las variables usadas serán siempre reflectividad, pero en tiempos anteriores; es decir, $X_{i,t_i} = Z_{t-i}$, para $i = 1, 2$. Dado que las variables tienen información hasta tres pasos de tiempo anteriores, no hay información suficiente para identificar en el modelo el componente estocástico; así, se

asume que el componente estocástico está formado por una sucesión independiente con media cero una varianza constante. En un futuro cercano, existirá la posibilidad de contar con datos del radar cada 2 minutos, entonces, será asequible la identificación de la estructura del componente estocástico.

Según el Servicio Meteorológico Nacional, aproximadamente 62 personas mueren cada año en los tornados. En 2008, 124 personas perdieron la vida. y en ese mismo año, el THHC (Trauma, Health & Hazards Center) de la Universidad del Colorado diseñó un proyecto para la instalación de radares en el año de 2013, llamado CASA (Center for Collaborative Adaptive Sensing of the Atmosphere), que funciona en la Universidad de Puerto Rico en su recinto de Mayagüez y trabaja en colaboración con las Universidades del Estado de Colorado, y que une a un equipo multidisciplinario que incluye a investigadores de las Universidades de Massachusetts-Amherst, Oklahoma, Virginia, Indiana de Pennsylvania y la Universidad de Magill. En el RUM, los Dre(a)s. Rafael Rodriguez, José Colom, Sandra Cruz Polo y Eric Harmesen, del departamento de Ingeniería Eléctrica y los profesores Nazario D. Ramirez del Departamento de Ingeniería Industrial forman parte del equipo CASA. Este radar, arrojará datos a intervalos de tiempo mucho mas cortos (2 minutos).

2.5.8. Variograma

La variación en el espacio de una función aleatoria $Z(x)$, puede ser descrita por la diferencia entre los valores de un par de puntos x y $x + h$ en una región \mathcal{D} ; es decir, por:

$$Z(x + h) - Z(x), \tag{2.58}$$

el cual es llamado incremento. El variograma $\Upsilon(h)$, está definido por la hipótesis de estacionaridad débil de orden dos. Esta hipótesis (Definición 2.5), la cual es simplemente una característica de la función de variable aleatoria y está formada por las

dos suposiciones acerca del incremento h , para la función aleatoria y significan lo siguiente:

- a. La media $m(h)$ de los incrementos es invariante para toda translación de un valor h dentro del dominio \mathcal{D} .
- b. La varianza de los incrementos, está acotada por el valor $2\Upsilon(h)$, y depende de la longitud y la orientación de h , más no de su ubicación en el dominio \mathcal{D} .

Lo anterior significa que para un par de puntos $x, x+h \in \mathcal{D}$, y para toda $h \in \mathbb{R}$, se tiene lo siguiente:

$$E[Z(x+h) - Z(x)] = m(h) = 0, \quad (2.59)$$

$$\text{var}[Z(x+h) - Z(x)] = 2\Upsilon(h). \quad (2.60)$$

Estas dos propiedades de estacionaridad de funciones de variables aleatorias, nos llevan a la definición de variograma:

$$\Upsilon(h) = \frac{1}{2}E[(Z(x+h) - Z(x))^2]. \quad (2.61)$$

Para todo $h \in \mathbb{R}$, $h \geq 0$, se tienen las siguientes propiedades del variograma:

- a. $\Upsilon(0) = 0$,
- b. $\Upsilon(h) \geq 0$,
- c. $\Upsilon(-h) = \Upsilon(h)$.

2.5.9. Método de Interpolación de Kriging

Suponga que se tiene una data para un tipo de fenómeno aleatorio en el espacio, cuya medición es Z ; la cual se encuentra localizada en un punto x_α de una región \mathcal{D} , entonces, $Z(x_\alpha)$ será una variable aleatoria, la cual depende de la ubicación y su medida. En particular, sea x_0 un punto de la región \mathcal{D} , entonces, $Z(x_0)$ será la

variable aleatoria correspondiente al valor de Z en x_0 . De otro lado podemos asumir que existen infinitas variables aleatorias $Z(x)$ en la región \mathcal{D} las cuales serían una función aleatoria de la ubicación x . Se asume que la función aleatoria $Z(x_0)$ es *estacionaria de segundo orden*. Lo anterior puede ser escrito de la siguiente manera:

$$E[Z(x+h)] = E[Z(x)], \quad (2.62)$$

$$Cov[Z(x+h), Z(x)] = C(h). \quad (2.63)$$

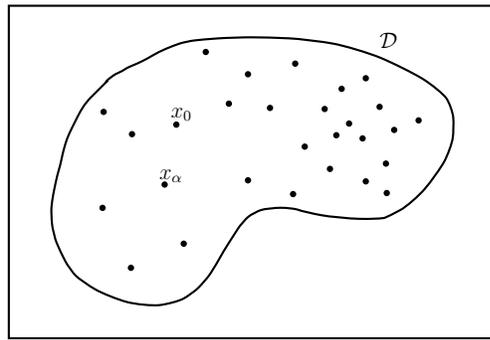


Figura 2-3: Punto x_α y su estimación para el punto x_0 en el dominio \mathcal{D} .

En primer lugar, se supone que el valor esperado $E[Z(x)] = m$, es el mismo para todo punto $x \in \mathcal{D}$; la covarianza entre un par de puntos x_i y x_j depende únicamente de la distancia h a la cual se encuentran separados. El problema consiste en construir los pesos promedios que permitan estimar el valor de $Z(x_\alpha)$ en cada punto x_α , para $\alpha = 1, 2, \dots, n$ (Figura 2-3). Este tipo de estimación se basa en el conocimiento de la covarianza entre las variables aleatorias en su respectivas ubicaciones. La respuesta a este problema está ligado a la regresión múltiple en la cual $Z(x_\alpha)$ hace las veces de regresores y $Z(x_0)$ es la estimación. Si se usa la hipótesis de que la media m es constante en la región \mathcal{D} y es calculado como el promedio de la data en la región. El método de interpolación Kriging simple es usado para calcular los residuales usando el valor a priori de la media m . En analogía con la regresión lineal múltiple, los pesos

promedios para estas estimaciones se define como:

$$Z^*(x_0) = m + \sum_{\alpha=1}^n w_{\alpha}(Z(x_{\alpha}) - m), \quad (2.64)$$

donde, w_{α} son los pesos correspondientes a los residuales $Z(x_{\alpha}) - m$. Nótese que a diferencia de la regresión múltiple, en nuestro caso la media m es la misma para todas $x \in \mathcal{D}$ esto debido a las suposición de estacionariedad. El error de estimación en un punto x_0 viene dado por:

$$Z^*(x_0) - Z(x_0). \quad (2.65)$$

Podemos ver que el estimador es insesgado ya que:

$$\begin{aligned} E[Z^*(x_0) - Z(x_0)] &= m + \sum_{\alpha=1}^n w_{\alpha}E[Z(x_{\alpha}) - m] - E[Z(x_0)], \\ &= m + \sum_{\alpha=1}^n w_{\alpha}(m - m) - m, \\ &= 0. \end{aligned} \quad (2.66)$$

De otro lado, la varianza del error de estimación viene dada por:

$$\sigma_e^2 = \text{var}[Z^*(x_0) - Z(x_0)] = E[(Z^*(x_0) - Z(x_0))^2]. \quad (2.67)$$

Expandiendo la ecuación (2.67), se tiene

$$\begin{aligned} \sigma_e^2 &= E[((Z^*(x_0))^2 + (Z(x_0))^2 - 2Z^*(x_0)Z(x_0))], \\ &= \sum_{\alpha=1}^n \sum_{\beta=1}^n w_{\alpha}w_{\beta}C(x_{\alpha} - x_{\beta}) + C(x_0 - x_0) - 2 \sum_{\alpha=1}^n w_{\alpha}C(x_{\alpha} - x_0); \end{aligned} \quad (2.68)$$

realizando algunos cálculos en la ultima ecuación, podemos escribir

$$\text{Cov}[Z(x_{\alpha}), Z(x_{\beta})] = C(x_{\alpha} - x_{\beta}), \quad (2.69)$$

puesto que la covarianza espacial depende solo de la diferencia de los vectores x_{α} y x_{β} , se sigue que el proceso es estacionario en el sentido de (2.62) y (2.63). Calculemos

cuando la varianza es mínima:

$$\frac{\partial \sigma_e^2}{\partial w_\alpha} = 0, \text{ para } \alpha = 1, 2, \dots, n. \quad (2.70)$$

De manera mas explicita, se tiene que para cada $\alpha = 1, 2, \dots, n$,

$$2 \sum_{\beta=1}^n w_\beta C(x_\alpha - x_\beta) - 2C(x_\alpha - x_0) = 0, \quad (2.71)$$

y así, el sistema de ecuaciones para el método de interpolación de Kriging simple, se escribe

$$\sum_{\beta=1}^n w_\beta C(x_\alpha - x_\beta) = C(x_\alpha - x_0), \text{ para } \alpha = 1, 2, \dots, n. \quad (2.72)$$

El lado izquierdo de la ecuación (2.72) describe la covarianza entre las variables $Z(x_\alpha)$ y $Z(x_\beta)$. El lado derecho describe la covarianza entre cada localización del dato y la localización donde un estimado es buscado. La solución del sistema (2.72) consiste en los pesos w_α del método de Kriging simple. Para un estudio más detallado del método de interpolación de Kriging, puede ver Wackernagel, 1994. [14].

Kriging Ordinario

El método de Kriging ordinario, es normalmente el más usado de los métodos de interpolación de Kriging. Este sirve para estimar el valor de un punto en una región para el cual su variograma es conocido, usando los datos en la vecindad mas cercana del punto. El Kriging ordinario evalúa implícitamente el significado de una vecindad en movimiento.

Deseamos estimar un valor de x_0 como esta representado en la Figura 2-4 usando los valores de los datos de los n puntos x_α de la muestra más cercanos y combinándolos linealmente con pesos w_α .

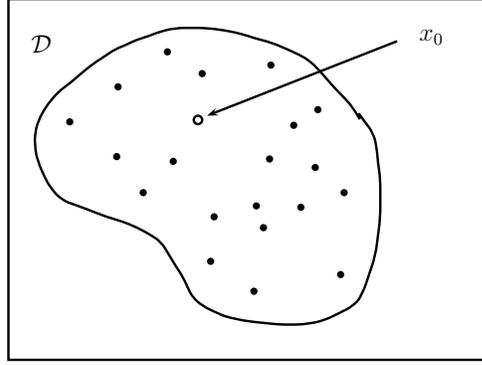


Figura 2–4: Un dominio con una muestra de puntos espaciados irregularmente (puntos negros) y una ubicación de interés x_0 .

$$Z_{OK}^*(x_0) = \sum_{\alpha=1}^n w_{\alpha} Z(x_{\alpha}). \quad (2.73)$$

Obviamente se debe restringir la anterior suma de los pesos a uno; porque en el caso particular cuando todos los valores de los datos son iguales a una constante, el valor estimado también debe ser igual a esta constante.

Asumiremos que los datos son parte de una realización de una función aleatoria intrínseca $Z(x)$ con un variograma $\gamma(h)$.

El insesgamiento se justifica con una suma de pesos unitarios.

$$E[Z^*(x_0) - Z(x_0)] = E\left[\sum_{\alpha=1}^n w_{\alpha} Z(x_{\alpha}) - Z(x_0) \cdot \underbrace{\sum_{\alpha=1}^n w_{\alpha}}_1\right], \quad (2.74)$$

$$= \sum_{\alpha=1}^n w_{\alpha} E[Z(x_{\alpha}) - Z(x_0)], \quad (2.75)$$

$$= 0, \quad (2.76)$$

Lo anterior se debe a que las esperanzas de los incrementos son cero.

La estimación de la varianza $\sigma_e^2 = \text{var}(Z^*(x_0) - Z(x_0))$ es la varianza de la combinación lineal:

$$Z^*(x_0) - Z(x_0) = \sum_{\alpha=1}^n w_{\alpha} Z(x_{\alpha}) - 1 \cdot Z(x_0) = \sum_{\alpha=0}^n w_{\alpha} Z(x_{\alpha}), \quad (2.77)$$

con un peso $w_0 = -1$ y

$$\sum_{\alpha=0}^n w_\alpha = 0; \quad (2.78)$$

por lo tanto la condición de que los pesos numerados del 1 al n , de la suma anterior (2.78) sean iguales a uno, implica también, que el uso del variograma está autorizado en el cálculo de la varianza del error de estimación.

La varianza es correcta para Kriging ordinario, pero no para Kriging simple, porque este último no incluye cualquier restricción en los pesos. la condición 2.78 corresponde a una función definida negativa condicionalmente (Wackernagel, 1994. [14]).

La estimación de la varianza es:

$$\sigma_e^2 = E[(Z^*(x_0) - Z(x_0))^2], \quad (2.79)$$

$$= -\gamma(x_0 - x_0) - \sum_{\alpha=1}^n \sum_{\beta=1}^n w_\alpha w_\beta \gamma(x_\alpha - x_\beta) + 2 \sum_{\alpha=0}^n w_\alpha \gamma(x_\alpha - x_\beta). \quad (2.80)$$

Minimizando la estimación de la varianza con la restricción de los pesos, obtenemos el *Sistema Kriging Ordinario*(OK).

$$\begin{pmatrix} \gamma(x_1 - x_1) & \cdots & \gamma(x_1 - x_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n - x_1) & \cdots & \gamma(x_n - x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1^{OK} \\ \vdots \\ w_n^{OK} \\ \mu_{OK} \end{pmatrix} = \begin{pmatrix} w_1^{OK} \\ \vdots \\ w_n^{OK} \\ \mu_{OK} \end{pmatrix}, \quad (2.81)$$

donde los w_α^{OK} , son pesos asignados a los valores de los datos y μ_{OK} es el parámetro de Lagrange. El lado izquierdo del sistema describe las diferencias entre los puntos de la muestra, mientras que el lado derecho muestra la diferencia entre cada punto de la muestra y el punto de estimación x_0 .

Realizando la multiplicación de matrices, el sistema Kriging ordinario puede ser reescrito de la forma:

$$\left\{ \begin{array}{l} \sum_{\beta=1}^n w_{\beta}^{OK} \gamma(x_{\alpha} - x_{\beta}) \gamma(x_{\alpha} - x_{\beta}) + \mu_{OK} = \gamma(x_{\alpha} - x_0), \quad \text{para } \alpha = 1, \dots, n \\ \sum_{\beta=1}^n w_{\beta}^{OK} = 1. \end{array} \right. \quad (2.82)$$

La estimación de la varianza del Kriging ordinario es:

$$\sigma_{OK}^2 = \mu_{OK} - \gamma(x_0 - x_0) + \sum_{\alpha=1}^n w_{\alpha}^{OK} \gamma(x_{\alpha} - x_0). \quad (2.83)$$

El kriging ordinario es un *interpolador exacto* en el sentido de que si x_0 es exactamente un punto de la muestra, entonces el valor estimado será el mismo valor del dato en ese punto.

$$Z^*(x_0) = Z(x_{\alpha}), \text{ si } x_0 = x_{\alpha}. \quad (2.84)$$

Esto puede ser visto fácilmente. Cuando x_0 es uno de los puntos de la muestra, el lado derecho del sistema Kriging es igual a una columna del lado izquierdo de la matriz. Un vector de pesos \mathbf{w} , con un peso para que la columna igual a 1 y todos los otros pesos (incluyendo μ_{OK}) igual a cero es una solución del sistema. Como la parte izquierda de la matriz es no singular, esta es la única solución.

Validación Cruzada

La Validación cruzada, es una forma sencilla de comparar varios supuestos; ya sea sobre el modelo (por ejemplo, el tipo de variograma y sus parámetros, el tamaño de la vecindad kriging) o los datos (por ejemplo, valores que no encajan en su vecindario como: valores atípicos y anomalías puntuales)

En el procedimiento de validación cruzada, cada valor de la muestra $Z(x_{\alpha})$, se retira del conjunto de datos y un valor $Z^*(x_{[\alpha]})$ en esa ubicación se estima mediante las otras $n - 1$ muestras.

los corchetes del índice alfa, simbolizan el hecho de que la estimación se realiza en la ubicación x_{α} excluyendo el valor de muestra.

La diferencia entre los valor de la data observado y el valor estimado,

$$Z(x_\alpha) - Z^*(x_{[\alpha]}), \quad (2.85)$$

este valor nos da un indicativo de que tan bien trabaja la estimación del valor en la ubicación de la región dada. Si el promedio del error de los valores de la validación cruzada no es lejano a cero; es decir,

$$\frac{1}{n} \sum_{\alpha=1}^n (Z(x_\alpha) - Z^*(x_{[\alpha]})) \cong 0. \quad (2.86)$$

Se puede decir que no existe un aparente sesgo en la estimación de los valores. Sin embargo, si los valores de la ecuación (2.86), son significativamente negativos (o positivos), entonces, este promedio significa que existe una sobre-estimación (respectivamente una sub-estimación) de los valores reales en la vecindad.

EL promedio de los cuadrados de los errores de la validación cruzada, ecuación (2.87), resulta muy apropiada para comparar diferentes estimaciones,

$$\frac{1}{n} \sum_{\alpha=1}^n (Z(x_\alpha) - Z^*(x_{[\alpha]}))^2. \quad (2.87)$$

La desviación estandar $\sigma_{[\alpha]}$, del Kriging representa el error predicho por el modelo cuando la interpolacion Kriging se realiza en un punto x_α (omitiendo el valor de la muestra en x_α). Dividiendo la ecuación (2.85) por $\sigma_{[\alpha]}$, podemos comparar las magnitudes del error predicho y el observado:

$$\frac{Z(x_\alpha) - Z^*(x_{[\alpha]})}{\sigma_{[\alpha]}}. \quad (2.88)$$

Si el promedio de los errores estandarizados de la validación cruzada son cercanos a uno,

$$\sum_{\alpha=1}^n \frac{(Z(x_\alpha) - Z^*(x_{[\alpha]}))^2}{\sigma_{[\alpha]}^2} \cong 1, \quad (2.89)$$

entonces, la estimación actual del error es igual al promedio del error descrito por el modelo. Esta ultima ecuación (2.89), nos da una idea acerca de lo adecuado que

es el modelo y de sus parámetros.

2.5.10. “Forward Selection” (Selección Hacia Adelante)

Para ejecutar este algoritmo de optimización de un modelo lineal múltiple, se empieza con la regresión lineal simple que considera como variable predictora a aquella que esta más altamente correlacionada (sin tomar en cuenta el signo) con la variable de respuesta. Si esta primera variable no es significativa entonces se considera el modelo $\hat{Y} = \bar{Y}$ y se detiene el proceso; de lo contrario se sigue y en el siguiente paso se añade al modelo la variable que reúne cualquiera de estos requisitos equivalentes entre sí:

- a. Aquella variable que tiene el estadístico de t , en valor absoluto, más grande entre las variables no incluidas aún en el modelo. Es decir, la variable con el F-parcial más grande.
- b. Aquella variable que produce el mayor incremento en el R^2 al ser añadida al modelo. Es decir, aquella variable que produce la mayor reducción en la suma de cuadrados del error.
- c. Aquella variable que tiene la correlación parcial más alta (en valor absoluto) con la variable de respuesta, tomando en cuenta las variables ya incluidas en el modelo. Aquí también está presente el efecto de anidamiento ya que toda variable que es añadida al modelo ya no puede salir del mismo. El proceso termina cuando se cumple una de las siguientes condiciones:
 - Se llega a un modelo con un número prefijado p^* de variables predictoras.
 - El valor de la prueba de F parcial para cada una de las variables no incluidas aun en el modelo es menor que un número prefijado F-in (por lo general este valor es 4). O en forma equivalente se para cuando el valor absoluto del estadístico de t es menor que la raíz cuadrada de F-in (por lo general, $|t| < 2$). Algunas veces se prefija de antemano un nivel de significancia dado α^* (digamos del 15%) para la

prueba de t o de F parcial en cada paso y en este caso se termina el proceso cuando todos los p - *values* de la prueba t de las variables no incluidas aún son mayores que α^* .

Puede consultar (Mongomery et al. 2006, [23]), para un estudio mas detallado sobre análisis de regresión lineal.

2.6. Técnicas de Procesamiento de Imágenes (Morfología)

La palabra morfología hace referencia, dentro del contexto de morfología matemática, a una herramienta para extraer componentes de una imagen, que sean útiles en la representación de la forma de una región; tales como contornos, esqueletos y cerco convexo. También son de interés las técnicas morfológicas para el pre o postprocesado, tales como el filtrado morfológico, la reducción y el recortado. La morfología se basa en la teoría de conjuntos para realizar ciertas operaciones que afecten a estos conjuntos. Como tal, la morfología ofrece un método poderoso y único de abordar muchos problemas del procesado de imágenes. Los conjuntos en la morfología matemática son representados como objetos dentro de una imagen. Por ejemplo, el conjunto de todos los píxeles negros de una imagen binaria es una descripción completa de ella. Puede consultar González et al. 2006, [15], para ver mas detalles sobre procesamiento de imágenes.

En la morfología binaria, una imagen se ve como un subconjunto de un espacio euclidiano \mathbb{R}^d o la cuadrícula entera \mathbb{Z}^d , para alguna dimensión d . Denotemos este espacio o cuadrícula por E . Un elemento estructurante C es una forma simple, predefinida, representada como una imagen binaria, que se utiliza para probar (comparar) con otra imagen binaria en las operaciones morfológicas, tales como erosión, dilatación, apertura y cierre. Todas estas operaciones morfológicas utilizan un elemento estructurante, que dependiendo de su tamaño y forma, el resultado de la operación es diferente. Realmente se trabaja con un conjunto de píxeles, no con la propia

imagen. El resultado al aplicar el elemento estructurante es otro conjunto, que es considerado como una imagen. La operación morfológica que será utilizada en este trabajo, es la conectividad; la que describiremos luego de aclarar lo siguiente:

Vecinos de un pixel

Un pixel p de coordenadas (x, y) tiene cuatro vecinos horizontales y verticales cuyas coordenadas son: $(x + 1, y)$, $(x - 1, y)$, $(x, y + 1)$ y $(x, y - 1)$. Este conjunto de pixeles recibe el nombre de $4 - vecinos$ de p , y se nota por $N_4(p)$. Cada pixel está a una distancia unitaria de (x, y) , y algunos de los vecinos de p caen fuera de la imagen digital si (x, y) está en el borde de la imagen.

Los pixeles diagonales vecinos de p tienen coordenadas: $(x + 1, y + 1)$, $(x - 1, y - 1)$, $(x - 1, y + 1)$ y $(x + 1, y - 1)$ y se notan por $N_D(p)$. Estos puntos, junto con los cuatro vecinos, se llaman $8 - vecinos$ de P y se denotan por $N_8(p)$. Debe verse, que igualmente como se dijo antes, algunos puntos de $N_D(p)$ y $N_8(p)$ caen fuera de la imagen si (x, y) está en el borde de la misma.

Conectividad

La conectividad entre píxeles es un concepto importante empleado para establecer los límites de los objetos y los componentes de áreas en una imagen. Para determinar si dos píxeles están conectados, debe determinarse si son adyacentes en algún sentido (como ser $4 - vecinos$) y si sus niveles de gris cumplen un criterio específico de similitud (como ser iguales). Por ejemplo, en una imagen binaria con valores 0 y 1, dos píxeles pueden ser $4 - vecinos$ pero no estarán conectados a menos que tengan el mismo valor. Sea V el conjunto de valores de nivel de gris empleados para definir la conectividad; por ejemplo, en una imagen binaria, se tendrá $V = \{1\}$ para la conectividad entre píxeles con valor 1. En una imagen con escala de grises, para la conectividad entre píxel con un rango de valores de intensidad de, por ejemplo,

32 a 64, se tiene $V = \{32, 33, \dots, 63, 64\}$. Se consideran tres tipos de conectividad:

1. *4 – conectividad*: Dos píxeles p y q con valores dentro de V están *4 – conectados* si q pertenece a $N_4(p)$.
2. *8 – conectividad*: Dos píxeles p y q con valores dentro de V están *8 – conectados* si q pertenece a $N_8(p)$.
3. *m – conectividad* (conectividad mixta): Dos píxeles p y q con valores dentro de V están *m – conectados* si:
 - a) q pertenece a $N_4(p)$, o bien
 - b) q pertenece a $N_D(p)$ y además el conjunto $N_4(p) \cap N_4(q)$ es vacío. (Este es el conjunto de píxeles que son *4 – vecinos* de p y q cuyos valores están en V .)

Un píxel p es adyacente de un píxel q si están conectados. Se puede definir 4, 8 o m adyacencia, dependiendo del tipo de conectividad especificada. Dos subconjuntos de la imagen, S_1 y S_2 , son adyacentes si algún píxel de S_1 es adyacente a algún píxel de S_2 .

2.7. Medición y Cobertura del Radar Doppler

La lluvia se mide en un punto sobre la superficie mediante el instrumento conocido como pluviómetro. El radar mide en forma indirecta la cantidad de lluvia a través de la diferencia de potencia que emite y regresa al radar. Se ha mostrado que la diferencia de potencia que recibe el radar es proporcional a la cantidad de precipitación (Doviak & Zrnić, 1993. [16]). La medida que realiza el pluviómetro es velocidad de lluvia y se expresa en mm/hr ; mientras que el radar mide la diferencia de potencia, la cual se le conoce como reflectividad y es proporcional a la velocidad de lluvia.

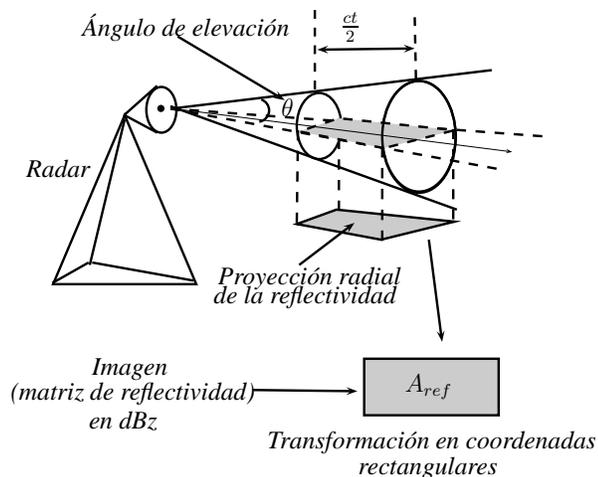
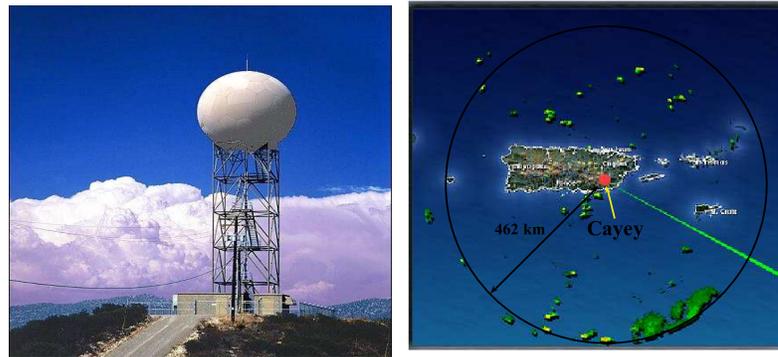


Figura 2–5: Área de reflectividad resultante (A_{res}), de tamaño radial $c\frac{t}{2}$, que se corresponde con la energía medida en un instante t y asociada a un distancia dada.

La reflectividad Z , se puede expresar en decibeles o en mm^6/m^3 . Se ha mostrado que la reflectividad es proporcional a la velocidad de lluvia y se puede representar mediante una relación de potencia, que viene dada por: $Z = aR^b$, en la que los parámetros a y b varían dependiendo de la región y condiciones climatológicas; R representa la medida de un punto en la superficie terrestre, que son aproximadamente 50 cm^2 . Debe notarse que esta relación es empírica y además no es una relación perfecta e inequívoca debido a que los valores de Z representan una información promedio de lo que ocurre en un volumen de la troposfera, la cual se mide a determinado ángulo de elevación, proporcionando así la información de reflectividad a una resolución radial (Figura 2–5). La resolución radial resultante se aproxima a coordenadas rectangulares de 1 km aproximadamente.

El área de estudio está limitada por la cobertura de radar (121×131 píxeles) que corresponde aproximadamente a un área circular con 462 km de radio a partir de la localización del radar el cual está ubicado en Cayey, PR ($18.12^\circ N, 66.08^\circ W$), a una altura de 886.63 m) este es un radar Doppler (WSR-88D) y trabaja a una frecuencia de 2.7 GHz en la banda S con una longitud de onda de 10 cm (Figura 2–6) y hace parte de la red de radares NEXRAD de EEUU a través de diversos

canales proporciona de forma gratuita la información de todos los radares.. En el Handbook No.11 del Federal Meteorological, 2005. [17], se encuentra una explicación mas detallada sobre el funcionamiento, mediciones y características de este tipo de radares .



(a) Radar Doppler

(b) Ubicación del Radar en PR.

Figura 2-6: Radar Doppler (WSR-88D).

Capítulo 3

METODOLOGIA

La lluvia es un evento que se origina a partir de un proceso termodinámico, el cual resulta del enfriamiento de una nube. En un inicio estas partículas son moléculas de vapor; posteriormente al enfriarse y al estar en continuo movimiento, estas entran en fricción con el aire y quedan cargadas eléctricamente; sumado a estas condiciones, si se tiene además que el ambiente está supersaturado de vapor, entonces, esto contribuye a que las partículas choquen, se fusionen y luego se condensen formando las gotas. Al ir creciendo con el tiempo, llega el momento en el que estas partículas, tienen un peso superior a la tensión superficial del aire, en un área definida, entonces, estas se precipitan y forman la lluvia. Podemos señalar, que este proceso es netamente estocástico discreto; además, está fuertemente influenciado por el tamaño de las gotas (partículas) de agua. Para inicios de la década de los 90's, Biggerstaff & Houze, 1993. [18], desarrollaron un modelo apoyados en la física y los procesos estocásticos, para representar la evolución de dos gotas; y encontraron que una partícula de $0.1 \mu m$ de radio, crecerá hasta llegar a unos $10 \mu m$ en promedio en 30 minutos. No obstante, una partícula de agua requerirá de un radio de $15 \mu m$ como mínimo para iniciar su descenso.

El período de tiempo que va desde la unificación de las gotas hasta su precipitación, puede durar de 30 minutos a varias horas dependiendo de las condiciones propias de la tropósfera. Es por esto, que para llegar a realizar predicciones de lluvia, se deben tener en cuenta los periodos posteriores al inicio de la misma; ya que el proceso no

depende únicamente del momento y condiciones actuales de la nube, sino de la evolución e instantes anteriores de ésta. Otro factor preponderante en la predicción, es el vector desplazamiento de la lluvia; pues este determinará la posible ubicación de las coordenadas (puntos o en nuestro caso pixeles) en las cuales habrá precipitación. En la búsqueda de un modelo adecuado, que represente los factores inherentes al fenómeno de la precipitación podemos hacer el siguiente análisis. Supongamos una variable continua Y , en un tiempo t , denotemos esta relación por $Y(t)$. Suponga ahora, que el tiempo t es discreto, cuando la variable Y es observada; entonces, escribiríamos simplemente Y_t . Ahora, si la variable Y , se ve influenciada por el de otra variable X , para cada instante t , entonces, podemos escribir $Y_t(X)$, para enfatizar que en cada instante de tiempo t la variable respuesta dependerá no solo del tiempo(discreto) sino también de la variable de entrada X . De lo anterior, podemos ver, que un modelo que se ajusta razonablemente a esta discusión es la función estocástica de transferencia FET (Box & Jenkins, 1976. [9] y Brockwell & Davis, 2002. [19]). Para la implementación óptima del modelo FET, se requiere una secuencia mínima de 24 observaciones (imágenes); sin embargo, las lecturas del radar Doppler (WSR-88D) ocurren en promedio cada 5-10 minutos; por tanto, se requieren casi 3 horas de observaciones por archivo para trabajar de manera adecuada con el modelo FET. No obstante, la mayoría de las tormentas (lluvia) no tienen una duración de este tiempo. Lo anterior indica, que el modelo no se adecua cuando las tormentas tienen prevaecía en periodos mas cortos. Por tanto, se requiere implementar una aproximación al modelo FET; este puede ser obtenido mediante el modelo de Decremento o Decrecimiento Exponencial o también utilizando el modelo de Retraso del Tiempo RT. Las diferencias en los modelos anteriormente nombrados, radica principalmente en la longitud de las series temporales utilizadas. El modelo FET utiliza series infinitas. Para fines prácticos, cuando hablamos de “serie infinita”, estaríamos hablando de mínimo 24 observaciones. De otro lado, los modelos RT y DE, utilizan

series finitas.

En este trabajo, se estudiarán las condiciones y estrategias necesarias para implementar el modelo RT, de acuerdo a las restricciones que nos presenta el modelo FET. Para esto, se mostrará paso a paso como se realiza la construcción del modelo bajo el análisis previo de las imágenes sucesivas del radar. En general, se realizarán tres etapas del trabajo. La primera parte del problema, se divide en dos fases: Aumento del tamaño de la imagen y eliminación de sectores o regiones de lluvia aisladas y de poca intensidad. En esta primera parte, se utilizará el algoritmo de punto medio aleatorio (PMA); con el cual, a partir de una imagen ZI , de dimensión $n \times m$, se pueda generar una nueva imagen con dimensión $(2n - 1) \times (2m - 1)$. En la segunda fase de la primera parte, se utilizarán herramientas de procesamiento de imágenes para reducir las regiones donde se centrará el estudio del fenómeno.

La segunda parte del trabajo, consiste en desarrollar un algoritmo basado en la correlación cruzada de imágenes; el cual, a partir de las imágenes ZI logre obtener una serie de vectores que describan el desplazamiento de la nube a través del tiempo. La tercera y última etapa del trabajo, se propone utilizar el modelo de retraso RT para estimar la intensidad de la lluvia, y luego realizar predicciones (una o dos horas como máximo) hacia adelante del final del periodo de muestra del evento de lluvia que se esté estudiando.

El siguiente, diagrama de flujo (Figura 3-1) ilustra el proceso que se utiliza para encontrar las predicciones de la intensidad de la lluvia y los vectores de desplazamiento.

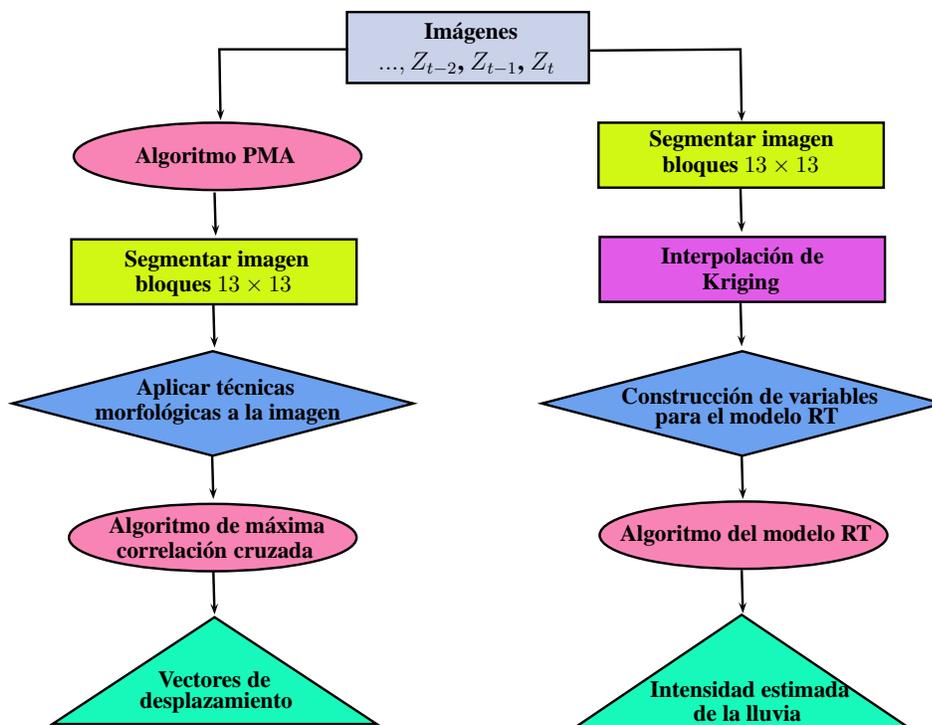


Figura 3–1: Diagrama de flujo para los procesos de predicción de intensidad y vectores de desplazamiento de la lluvia.

3.1. Procesamiento de la Imagen

A continuación, presentamos algunos aspectos que fueron usados para el procesamiento de cada imagen de reflectividad.

3.1.1. Imágenes de Mayor Pixelado

La pregunta que surge en primer lugar es: ¿porque no utilizar las imágenes originales y en cambio transformarlas en imágenes con mayor pixelado? La respuesta es la siguiente: las imágenes que se obtienen del radar, tienen una resolución de (121×131) pixeles), cada uno de estos pixeles, representa un área aproximada de 16 km^2 . Por lo tanto, al realizar un desplazamiento de dos pixeles en cualquier dirección, por ejemplo, se estaría abarcando un área aproximada de 400 km^2 ; esto representa un problema al recorrer los pixeles y medir el desplazamiento de la nube; ya que en un área tan extensa, la variabilidad de movimiento e intensidad de

la nube podría presentar cambios abruptos. La forma de minimizar este problema, es generando una nueva imagen cuya cantidad de pixeles sea mayor y permita un menor desplazamiento y que no afecte la medida de la reflectividad en los pixeles de lluvia. La descripción del proceso por el cual se obtienen las imágenes con una mayor pixelado está descrito en León, 2005. [20].

3.1.2. Eliminación de Regiones

Generalmente, las tormentas presentan sectores en los cuales la intensidad de la lluvia es relativamente bajo ($< 15 \text{ dBz}$); de otro lado, también se presentan sectores aislados, los cuales al ser relativamente pequeños en comparación con el tamaño de otros sectores de la tormenta son de poca significancia. Por tal razón, es conveniente eliminar estas regiones que solo contribuyen al “ruido” de la medición de los vectores de desplazamiento. Por lo tanto, son útiles las técnicas de morfología (conectividad) de imágenes descritas en la sección 2.6 del Capítulo 2.

3.2. Vector de Movimiento de la Tormenta

Para la construcción y posterior visualización de los vectores de desplazamiento que describan de manera más adecuada el movimiento de una tormenta, se utilizará en primer lugar un pre-procesamiento de la imagen del evento (tormenta) en cuestión. Puesto que la lluvia es un fenómeno que tiene implícito diferentes elementos (temperatura, humedad, viento, altitud, etc.), se debe pensar que la nube se comporta como un “fluido gaseoso”; que está sujeta a mucha variabilidad; además, tener siempre presente que, en situaciones de desarrollo convectivo, el problema del pronóstico involucra la iniciación, evolución y movimiento de las tormentas, que la hacen deformar constantemente a través del tiempo (Figura 3-2). Estas deformaciones son intermitentes y se dan a través del tiempo y en diferentes regiones

de la tormenta. Esto genera una dificultad clara al momento de intentar encontrar similitudes del evento en diferentes lugares y tiempos.

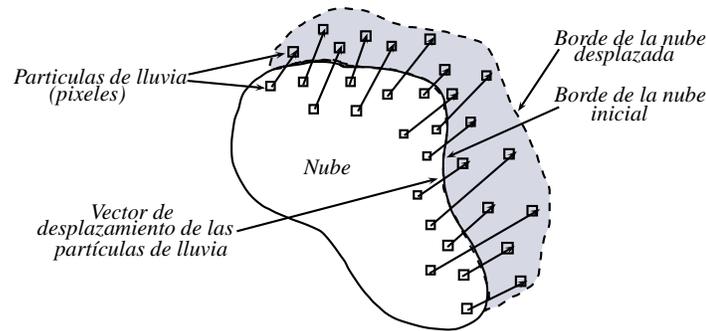


Figura 3–2: Esquema representativas del movimiento de las partículas de lluvia.

A continuación, un listado de los aspectos del pre-procesamiento necesario en cada imagen, para la mejor adaptación de los algoritmos implementados.

1. Eliminar los datos de reflectividad que estén por fuera de un rango establecido.
2. Encontrar las regiones máximas donde la tormenta se encuentre conectada.
3. Eliminar las regiones en las cuales la tormenta sea menor a un área específica. Esto con el fin de reducir al máximo el ruido en las similitudes entre imagen e imagen.

En el momento que encontremos tormentas en el área de interés es de mucha importancia, tanto como desde el punto de vista técnico como científico, que el procedimiento por el cual se detecte y se analice la evolución de la tormenta sea lo más sencillo y robusto posible, ya que la simplicidad redundará en la rapidez del algoritmo; lo que será decisivo al momento de la implementación en tiempo real. Ahora, en cuanto a la robustez del algoritmo, se puede decir que brindará un mejor análisis y pronóstico final. Debe tenerse en cuenta que las tormentas individuales (celdas simples) poseen un tiempo de vida cercano a los 20 minutos (Wilson et al, 1998. [6]), y estas se mueven según el viento medio y son capaces de unirse para formar grupos de tormentas (con un tiempo de vida mucho mayor y movimiento muy diferente al de las celdas simples), que pueden a su vez subdividirse con posterioridad. Se puede notar, que hay una clara diferencia entre el pronóstico del movimiento de las

tormentas como un todo y aquellas que permiten distinguir su movimiento interno.

3.2.1. Correlación Cruzada e Imágenes

La correlación cruzada (a veces denominada “covarianza cruzada”) puede ser vista como una medida de la similitud entre dos señales (imágenes para nuestro caso particular); por tal razón, es usada frecuentemente para encontrar características relevantes en una imagen de la cual se conocen pocos elementos de sus componentes y de su estructura. Generalmente, las imágenes que se someten a comparación, están en función del tiempo; y este tiempo es relativo entre las imágenes. En este trabajo, se usará la correlación cruzada como una herramienta para encontrar la relación (similitudes) entre imágenes (o sectores de la imagen) de radar que son consecutivas en el tiempo, y así de esta forma encontrar el desplazamiento de la nube. En el artículo de Leese et al, 1970. [21], se puede ver una discusión detallada sobre éste método. A continuación, se ilustrará de manera más detallada el método de correlación cruzada aplicado a las imágenes de radar; teniendo en cuenta que cada imagen de radar, es propiamente una matriz de datos de reflectividad.

Sean Z_{t-1} y Z_t , matrices de igual dimensión. Consideremos los bloques (submatrices) A_{t-1} y A_t de Z_{t-1} y Z_t , de dimensión $r \times r$ en los tiempos $t - 1$ y t respectivamente. Cada submatriz A_t , para todo $t \in \mathbb{Z}$ y con centro en (i_0, j_0) , estará dispuesta de la siguiente forma: de la fila nr hasta la fila $nr + 1$; y de la columna nc hasta la columna $nc + 1$, donde $r \in \{0, 1, 2, \dots\}$ y $c \in \{0, 1, 2, \dots\}$ son fijos y $n \leq 121$ impar. De manera análoga se hace para A_{t-1} . Se construye una matriz S_t de tamaño $m \times m$, $m \leq 121$ impar y cuyo centro estará ubicado en cada componente de la submatriz A_t . Posteriormente, se encuentra la correlación cruzada, ecuación (3.1), entre todas las matrices $S_t^{i,j}$ con $(i, j) \in A_t$ y $S_{t-1}^{i,j}$, $(i, j) \in A_{t-1}$. A la matriz $S_t^{i,j}$ la llamaremos rastreador. La dimensión de la celda rastreadora, será $m \times m$, con $m \leq 121$. De lo

anterior, se tiene que la correlación de las celdas viene dada por:

$$R(i, j) = \frac{Cov[S_t^{i',j'}, S_{t-1}^{i,j}]}{\sigma(S_t^{i',j'})\sigma(S_{t-1}^{i,j})}, \quad (3.1)$$

para todo punto $(i, j) \in A_{t-1}$ y $(i', j') \in A_t$ fijo; $Cov[S_t^{i',j'}, S_{t-1}^{i,j}]$ es la covarianza entre las matrices vectorizadas $S_t^{i',j'}$ y $S_{t-1}^{i,j}$; además, $\sigma(S_t^{i',j'})$ y $\sigma(S_{t-1}^{i,j})$ serán las desviaciones estandar de las matrices vectorizadas $S_t^{i',j'}$ y $S_{t-1}^{i,j}$, respectivamente. El valor $R(i, j)$ en la ecuación (3.1), es el coeficiente de correlación cruzada entre las matrices $S_{t-1}^{i,j}$ de A_{t-1} y $S_t^{i',j'}$ de A_t , en el tiempo t y $t-1$ respectivamente. La máxima correlación cruzada será entonces, $R(p, q) = \max\{R(i, j) : (i, j) \in A_{t-1}\}$. Por lo tanto, será en la componente (p, q) de A_{t-1} , donde estará ubicada $(S_t^{i',j'}$ superpuesta sobre $S_{t-1}^{p,q}$) la matriz (región) $S_t^{i',j'}$ de mayor similitud para la submatriz A_t de Z_t . Ahora bien, el vector de desplazamiento (p, q) del patrón de reconocimiento (matriz $S_{t-1}^{p,q}$ con mayor correlación cruzada sobre todos los (i, j) de A_t), en el intervalo de tiempo $[t-1, t]$, es proporcional al punto de localización de correlación máxima; es decir, $(p, q) = (i'\Delta x, j'\Delta y)$. Por lo tanto, la velocidad y dirección del movimiento de la nube, vienen dadas por las ecuaciones:

$$|C| = \frac{[(i'\Delta x)^2 + (j'\Delta y)^2]^{\frac{1}{2}}}{\Delta t}, \quad (3.2)$$

$$\theta = \arctan \left[\frac{i'\Delta x}{j'\Delta y} \right], \quad (3.3)$$

donde, $|C|$ es la norma del vector velocidad (la dimensión del vector), θ es la dirección, Δx y Δy son los incrementos de intervalo en la matriz Z_t en las direcciones I y J respectivamente, y Δt es el intervalo de tiempo entre dos imágenes sucesivas. En la Figura 3-3, podemos ilustrar gráficamente lo discutido hasta el momento.

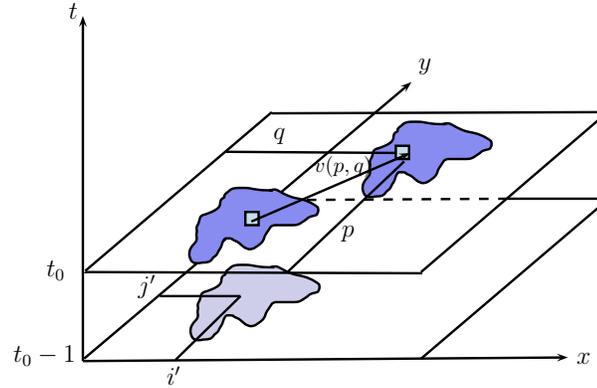


Figura 3-3: Desplazamiento de la nube sobre las matrices A_{t_0-1} y A_{t_0} ; en los tiempo $t_0 - 1$ y t_0 respectivamente. El vector $v(x, y)$ es el vector de desplazamiento de la nube.

Los coeficientes de la correlación cruzada, son calculados para una plantilla $S_t^{i,j}$, desplazándola en cada componente de A_{t-1} de Z_{t-1} ; luego se toma el punto (p, q) , donde la correlación de $S_{t-1}^{p,q}$ y $S_t^{i',j'}$ fue máxima (Figura 3-4).

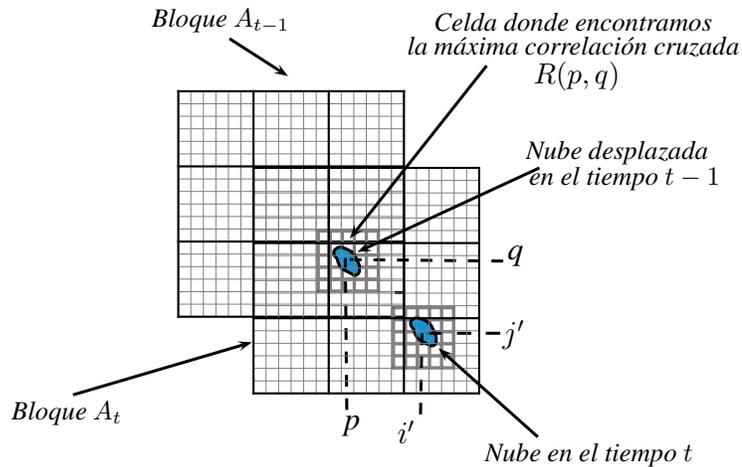


Figura 3-4: La celda $S_{t-1}^{p,q}$, tiene su máxima correlación con la celda $S_t^{i',j'}$ en el tiempo t .

Para la implementación del algoritmo de desplazamiento de la nube, se hace necesario encontrar las dimensiones óptimas de las matrices A_{t-1} y A_t ; así como también las dimensiones de la celda rastreadora $S_t^{i,j}$. Otro de los posibles inconvenientes, puede ser la medida del intervalo de tiempo que se escoja; ya que éste, incide directamente en la posibilidad de encontrar celdas cuya correlación cruzada sea máxima, pero ficticia. Es decir, es posible que exista una superposición o traslapeo de imágenes de otros sectores de la nube dentro de la misma submatriz A_{t-1} , y que

esto incida en el cálculo del máximo. Esto puede ocurrir debido a varios factores, como por ejemplo, que las nubes no siempre actúan como cuerpos rígidos. Algunas nubes pueden disiparse y puede formarse otras. Además, por la acción del viento, o por el mismo desarrollo intrínseco de las nubes, estas pueden cambiar de forma, y de esta forma, la máxima correlación podría aparecer en algún lugar falso.

Luego de varias pruebas, se encontró que las dimensiones ($n \times n$) óptimas para los bloques A_{t-1} y A_t son 13×13 píxeles; y las dimensiones ($m \times m$) del rastreador $S_t^{i,j}$ serán iguales a 51×51 píxeles. El centro del rastreador (i, j) , se ubicará en todos los (i, j) de la submatriz A_{t-1} y luego se calcula la correlación cruzada para cada par de matrices (rastreadores). Al finalizar el proceso, se tomará el valor el máximo de todas las correlaciones. No obstante, podemos disminuir la posibilidad de encontrar correlaciones máximas que generen vectores que no se encuentren en concordancia con la dirección real de la nube. Las siguientes condiciones; algunas de las cuales fueron adaptadas de Kishtawal et al., 2007. [22], son impuestas para los puntos de máxima correlación en imágenes sucesivas, para los tiempos $t - 1$ y t ; que son encontradas por el rastreador $S_t^{i,j}$ sobre la submatriz A_{t-1} , sirven como base para la consecución de este objetivo.

- I. Si el valor de máxima correlación cruzada, $R(p, q) < 0.8$, entonces, se rechazan los índices de la ubicación correspondiente a esa correlación.
- II. Cualquier vector que resulte máximo, no debe desviarse más de 60° de la dirección media de los vectores de un vecindario de submatrices A_t de dimensión 3×3 en un mismo tiempo t .
- III. Cualquier magnitud del vector máximo, no debe superar en más de 10 m/s de la magnitud media los vectores de un vecindario de submatrices A_t de dimensión 3×3 en un mismo tiempo t .

- IV. Cualquier dirección del vector máximo en una submatriz A_{t-1} , no debe desviarse más de 60° de la dirección del vector máximo correspondiente en la submatriz A_t .
- V. Cualquier magnitud del vector máximo en una submatriz A_{t-1} , no debe superar en más de 10 m/s de la magnitud del vector máximo correspondiente en la submatriz A_t .
- VI. Si sobre una imagen para un tiempo $t-1$, se coloca un rastreador $S_t^{i,j}$ de A_t ; y éste, se ubica sobre A_{t-1} , y produce una correlación máxima $R(p, q)$; y en una segunda imagen para un tiempo t , un rastreador $S_{t+1}^{i,j}$ de A_{t+1} , se ubica sobre A_t , y produce una correlación máxima $R(p', q')$ igual a $R(p, q)$; y además, se satisfacen las condiciones (IV.) y (V.), en los dos vectores generados por este máximo, entonces, se promedian estos dos vectores para producir el vector final en ese lugar. Si únicamente en una, de dos imágenes sucesivas en los tiempos $t-1$ y t , existe un único punto $(p, q) = (p', q')$ en A_t ó A_{t+1} ; cuyo rastreador $S_t^{i,j}$ tiene correlación cruzada máxima sobre A_{t-1} ó $S_{t+1}^{i,j}$ tiene correlación cruzada máxima sobre A_t , entonces, este vector máximo se mantiene, pero debe pasar las pruebas de los incisos I., II. y III..

3.3. Estimación de la Intensidad de Lluvia

Como ya se discutió al inicio del Capítulo, no es posible la implementación inmediata del modelo FET; debido a la falta de información del radar en periodos mucho más cortos. Por lo tanto, procedemos a implementar el modelo RT. A continuación, se construirán las variables que alimentaran el modelo RT, para lograr posteriormente su implementación. Las secuencias de datos de reflectividad fueron filtradas, de tal manera que solo se tienen en cuenta aquellos pixeles que tienen una reflectividad de 15 a 55 dBz . Puesto que la reflectividad del radar varía en el tiempo y el espacio, es necesario dividir el área en regiones más pequeñas (13×13 pixeles); además, se restringen aquellas áreas en las cuales solo existan menos de 5 pixeles

de lluvia en el tiempo t , debido a que se considera que en éstas no hay suficiente información para las estimaciones. Los predictores potenciales son las observaciones previas de reflectividad localizadas en una vecindad con centro en un pixel proyectado. Aunque no se logró separar el aspecto de la variabilidad del crecimiento de la reflectividad y la variabilidad del vector de desplazamiento, el modelo RT, considera simultáneamente la variación total de la intensidad de la reflectividad en el espacio y el tiempo. Para reducir el esfuerzo computacional en el modelo RT, se utilizarán tres imágenes sucesivas para la predicción de la reflectividad. La variabilidad temporal es expresada por el subíndice t y la variabilidad espacial será denotada por superíndices (i, j) ; por lo tanto, un pixel de reflectividad observada, se denotará por $z_t^{i,j}$. El modelo incluye un gran número de variables; sin embargo, se utilizará el algoritmo de *Forward Selection* para accionar los parámetros de las variables más representativas (máximo 5) y descartar aquellas que no sean importantes para el modelo lineal. A continuación describiremos de forma detallada el proceso aplicado para construir el modelo RT y mostrar como se obtienen las variables usadas para el mismo.

En una imagen Z_{t-1} , ubicamos el centro (i_0, j_0) de una submatriz (bloque) A_{t-1} , de dimensión 13×13 . Posteriormente, recorreremos todos los pixeles $(i, j) \in A_{t-1}$ con una celda de dimensión 5×5 , que llamaremos $C_{t-1}^{i,j}$ y cuyo centro estará en cada (i, j) de A_{t-1} . De forma análoga, pero sobre la imagen Z_{t-2} , se recorren los pixeles de A_{t-2} ; pero esta vez, con una celda de dimensión 9×9 , que llamaremos $C_{t-2}^{i,j}$ y cuyo centro estará en cada (i, j) de A_{t-2} (Figura 3-5).

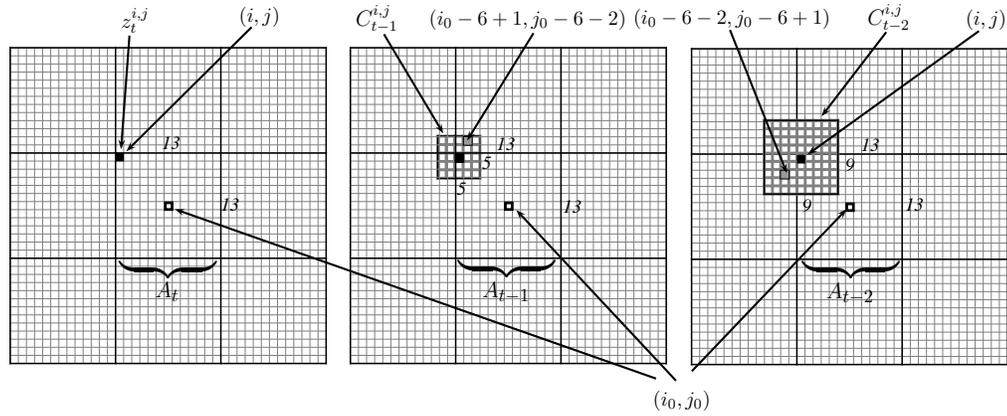


Figura 3-5: Construcción de las variables del modelo RT. Puede ver el recorrido de las celdas $C_{t-1}^{i,j}$ y $C_{t-2}^{i,j}$ sobre los bloques A_{t-1} y A_{t-2} respectivamente.

Cada celda $C_{t-1}^{i,j}$, tendrá un máximo total de 25 valores de reflectividad. Si el número r de pixeles, con reflectividad diferente de cero en $C_{t-1}^{i,j}$ es $r_{t-1} < 5$, entonces, se descarta esa celda y pasamos a la siguiente. Si $5 \leq r_{t-1} < 25$, entonces los valores de reflectividad faltantes en la celda, se estiman por medio de interpolación de Kriging ordinario; ver sección 2.5.9. De manera similar, el bloque A_{t-2} , tendrá un máximo total de 81 valores de reflectividad. Si el número r_{t-2} de pixeles con reflectividad diferente de cero en $C_{t-2}^{i,j}$ es menor que 9, entonces, se descarta esa celda y pasamos a la siguiente celda. Si por el contrario, $9 \leq r_{t-2} < 81$, entonces, interpolamos los valores faltantes en la celda. Cabe anotar, que si alguna de las celdas $C_{t-1}^{i,j}$ ó $C_{t-2}^{i,j}$, en la misma ubicación (i, j) , no cumplen las condiciones $5 \leq r_{t-1} < 25$ y $9 \leq r_{t-2} < 81$, entonces, se descartan ambas. Al final del proceso se tendrán m vectores fila, donde $1 \leq m \leq 169$. Los vectores resultantes, para una celda en los tiempos $t, t-1$ y $t-2$, se muestran en la Tabla 3-1; los vectores encontrados forman una matriz; y las columnas de esta matriz serán las variables del modelo. En general, este proceso es iterativo en todos los bloques de las matrices, a través del tiempo.

Los 106 vectores formados por todos los valores en cada $C_{t-1}^{i,j}$ y $C_{t-2}^{i,j}$ y cada submatriz A_{t-1} y A_{t-2} respectivamente, serán las variables con las cuales se buscará evaluar el modelo para cada submatriz A_t . No obstante, para reducir el esfuerzo

Tabla 3–1: Construcción de las variables predictoras X_i . Estas serán discriminadas con Forward Selection y solo se tomarán como máximo 5 variables para el modelo RT.

		Z_{t-1}				Z_{t-2}			
		X_1	X_2	\dots	X_{25}	X_{26}	X_{27}	\dots	X_{106}
Z_t	$z_1^{i,j}$	$z_{t-1}^{i_0-6-2, j_0-6-2}$	$z_{t-1}^{i_0-6-1, j_0-6-2}$	\dots	$z_{t-1}^{i_0-6+2, j_0-6+2}$	$z_{t-2}^{i_0-6-4, j_0-6-4}$	$z_{t-2}^{i_0-6-3, j_0-6-4}$	\dots	$z_{t-2}^{i_0-6+4, j_0-6+4}$
	$z_2^{i,j}$	$z_{t-1}^{i_0-5-2, j_0-6-2}$	$z_{t-1}^{i_0-5-1, j_0-6-2}$	\dots	$z_{t-1}^{i_0-5+2, j_0-5+2}$	$z_{t-2}^{i_0-5-4, j_0-6-4}$	$z_{t-2}^{i_0-5-3, j_0-6-4}$	\dots	$z_{t-2}^{i_0-5+4, j_0-6+4}$
	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\ddots	\vdots
	$z_m^{i,j}$	$z_{t-1}^{i_0-i-2, j_0-j-2}$	$z_{t-1}^{i_0-i-1, j_0-j-2}$	\dots	$z_{t-1}^{i_0-i+2, j_0-j+2}$	$z_{t-2}^{i_0-i-4, j_0-j-4}$	$z_{t-2}^{i_0-i-3, j_0-j-4}$	\dots	$z_{t-2}^{i_0-i+4, j_0-j+4}$

computacional, se realiza una rutina de depuración de las variables; que funciona de la siguiente forma: primero, se toman grupos de 5 variables y se escoje un maximo de 5 variables por cada grupo. posteriormente se vuelve a tomar grupos de 5 variables y nuevamente, se escoje un maximo de 5 variables por cada grupo. Esto se realizara hasta llegar a un número de variables que será ≤ 5 . Luego de todo el proceso, nos quedaremos sólo con las variables más significativas (a lo más 5 de las 106 variables). Esto se hará con el algoritmo de Forward Selection. Para más detalles puede ver Montgomery et al. 2006, [23] y Wilks, 1995. [24].

Al final del proceso para cada submatriz A_t con centro en el pixel $z_t^{i,j}$, se obtendrá la siguiente ecuación:

$$Z_t^{i,j} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{25} X_{25} + \beta_{26} X_{26} + \dots + \beta_{105} X_{105} + \beta_{106} X_{106} + \varepsilon_t. \quad (3.4)$$

En ésta, se distinguen los siguientes elementos:

$$Z_t^{i,j} = \{z_r^{i,j}\}, \text{ para } r = 1, 2, 3, \dots, m, \quad (3.5)$$

$$X_n = \{z_{t-1}^{i+K, j+L}\}, \text{ para } K, L \in \Omega \text{ y } n = 1, 2, \dots, 25, \quad (3.6)$$

$$X_n = \{z_{t-2}^{i+M, j+N}\}, \text{ para } M, N \in \Xi \text{ y } n = 26, 27, \dots, 106, \quad (3.7)$$

En síntesis, reescribiendo el modelo RT (ecuación (3.7)), éste tendrá la forma:

$$Z_t^{i,j} = \beta_0 + \sum_{K \in \Omega} \sum_{L \in \Omega} \beta_{KL} Z_{t-1}^{i+K, j+L} + \sum_{M \in \Xi} \sum_{N \in \Xi} \beta_{MN} Z_{t-2}^{i+M, j+N} + \varepsilon_t, \quad (3.8)$$

donde,

$$\begin{aligned}
i &\in \{i_0 + 0, i_0 \pm 1, i_0 \pm 2, i_0 \pm 3, i_0 \pm 4, i_0 \pm 5, i_0 \pm 6\}, \\
j &\in \{j_0 + 0, j_0 \pm 1, j_0 \pm 2, j_0 \pm 3, j_0 \pm 4, j_0 \pm 5, j_0 \pm 6\}, \\
\Omega &= \{0, \pm 1, \pm 2\}, \\
\Xi &= \{0, \pm 1, \pm 2, \pm 3, \pm 4\};
\end{aligned}$$

(i_0, j_0) es el centro de cada bloque A_t ; además, para $s = 0, 1, 2, \dots, 106$, β_s , son los coeficientes del modelo lineal con los retrasos, en los tiempos $t - 1$ y $t - 2$ de las submatrices A_{t-1} y A_{t-2} respectivamente.

El error

$$\varepsilon_t = \phi_2 \varepsilon_{t-2} + \phi_1 \varepsilon_{t-1} + a_t, \quad (3.9)$$

se ha estimado, que tiene un comportamiento autorregresivo de orden AR(2) (sección 2.5.2). Sin embargo, no es posible construir una función de autocorrelación, debido a que se dispone de solo los tiempos t , $t - 1$ y $t - 2$; así que, el orden del modelo se determinó por inspección. Esto es; se construyó un modelo y se observó el desempeño del mismo. En la ecuación (3.9), a_t , es una sucesión de errores independientes con media cero y varianza constante.

De forma análoga, se obtienen las demás ecuaciones para cada $Z_t^{i,j}$ con $(i, j) \in A_t$.

En resumen, el proceso para realizar las estimaciones de la reflectividad para cada imagen es el siguiente:

- I. Se dividen las imágenes (matrices) Z_{t-1} y Z_{t-2} en bloques de 13×13 .
- II. Para cada bloque A_{t-1} y A_{t-2} , ubicados en la misma posición (i_0, j_0) , se encuentran las variables X_n (Tabla 3-1).
- III. Se aplica el algoritmo de Forward Selection para capturar las variables significativas (de las 106 variables se tomará un máximo de 5).

iv. Para cada bloque A_t , se realiza la estimación utilizando la ecuación (3.8).

Capítulo 4

ANÁLISIS DE RESULTADOS DEL MODELO RT Y VECTOR DE DESPLAZAMIENTO

A continuación, mostraremos el desarrollo de la aplicación del modelo teórico, propuesto en el Capítulo 2. Para validar nuestro modelo se realizaron pruebas en 2 tormentas ocurridas en Puerto Rico durante los años 2003 y 2007 respectivamente. Los detalles de cada una de ellas se muestran en la Tabla 4-1.

Tabla 4-1: Evento usado para validación del modelo RT.

Evento para validación del modelo			
No.	Fecha del evento	Tipo de evento	Impacto humano o económico
1	Abril 17,18-2003	Onda tropical	Inundaciones repentinas grave se concentra a través de varios municipios Las excesivas lluvias obligaron a varios cientos de familias a salir de sus hogares; además, causó un estimado de 15 millones de dolares en daños a la infraestructura y la agricultura de EE.UU..
2	Octubre 27,29-2007	Fuerte onda tropical (huracán Noel)	Las lluvias torrenciales de Noel produjeron daños generalizados y pérdida de vidas en la República Dominicana, Haití, Jamaica, el este de Cuba y las Bahamas. Se estima que Noel causo un total de 163 muertes, mientras que 59 siguen desaparecidas.

Los dos eventos (tormentas ocurridas el 17 y 18 de abril del 2003 y 27 y 29 de octubre del 2007), se seleccionaron aleatoriamente de una muestra total de 10; éstos se encuentran en la página web del National Weather Service de Puerto Rico (<http://www.srh.noaa.gov>). La Tabla A-1 del Apéndice, muestra el resumen del total de eventos.

Los algoritmos utilizados para la implementación de los modelos y composición de gráficos, fueron creados con el software matemático MATLAB R2010a.

Para los eventos de abril 17, 18-2003 y octubre 27, 29-2007, podemos notar los siguientes aspectos: En forma general se puede observar que en ambos eventos la estimación de la reflectividad es menos precisa a medida que pasa el tiempo con respecto al punto observado donde se hizo la medición inicial. Esto implica que nuestro modelo de estimación de reflectividad a los 15 minutos es notablemente mejor que el estimado a los 30, 60 y 120 minutos. La Tabla 4-2, muestra el valor del coeficiente de determinación R^2 para cada evento en los diferentes tiempos de estudio; este valor mide la proporción de la variación de la reflectividad dada por el radar y que es explicada por los predictores seleccionados en las diferentes celdas para los modelos RT.

Ajuste del modelo RT por evento de lluvia		
Evento	tiempo de estimación	R^2
Abril 17,18-2003	15 min	0.9446
	30 min	0.8074
	60 min	0.4701
	120 min	0.4134
Octubre 27,29-2007	15 min	0.9124
	30 min	0.6281
	60 min	0.5701
	120 min	0.3400

Tabla 4-2: Coeficiente de determinación R^2 , para cada modelo RT (validación para cada tormenta).

Resultados de la estimación de la reflectividad para el evento 1, ocurrido en Puerto Rico el 17 y 18 de abril del 2003

A continuación se muestran las gráficas de la reflectividad observada y estimada para cada uno de los tiempos seleccionados en nuestro estudio.

La Figura 4-1, muestra la reflectividad observada a las 00:45 horas , el cual representará para nosotros el momento inicial del estudio.

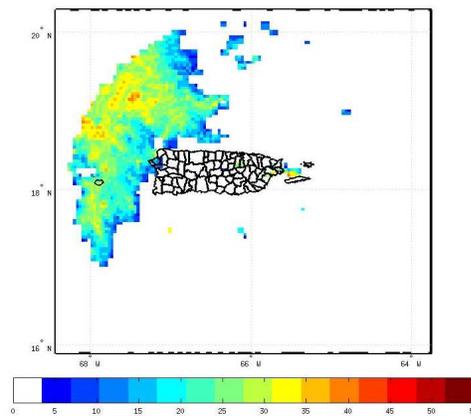


Figura 4-1: Reflectividad observada, para el evento del 2003/04/17 a las 00:45 horas en PR.

La Figura 4-2 (a), nos muestra la reflectividad observada luego de transcurridos 15 minutos del momento inicial. En ella se puede observar que pese a que el área de presencia de la lluvia no ha cambiado significativamente respecto al inicial, la intensidad de la misma si ha disminuido. Por otro lado, la Figura 4-2 (b), de reflectividad estimada por el modelo RT, se asemeja mucho a la observada, pues los valores de reflectividad varían entre 22 y 38 dBz aprox. en ambos casos; es necesario aclarar que el algoritmo construido para la estimación, no toma en consideración valores de reflectividad menores a 20 dBz de aquí que este segundo gráfico no muestra la reflectividad donde la intensidad de lluvia es baja; esto es, en los extremos de la región observada; lo cual podría generar conclusiones erróneas de la aparente reducción del área de lluvia en dicha región.

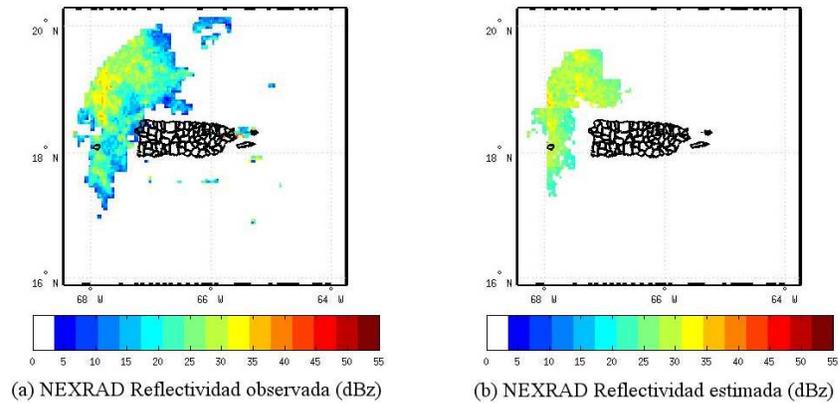


Figura 4-2: Reflectividad observada y estimada 15 min hacia adelante, para el evento del 2003/04/17 a las 01:00 horas en PR.

Las cuatro últimas Figuras (4-3 y 4-4), muestran los valores de reflectividad observadas y estimadas transcurridos 30 y 60 minutos del momento inicial de medición; las dos primeras son similares a las encontradas a los 15 minutos; sin embargo cabe señalar que la reflectividad es sobre-estimada en una pequeña región de coordenadas (19N, 67W) (Figura 4-4 (b)). Podemos notar que en la Figura 4-4 (a) de la reflectividad observada en ese momento, muestra pequeñas regiones de precipitación moderada alrededor de Puerto Rico; esto, no se registra en la estimación de reflectividad (Figura 4-4 (b)), debido a que el algoritmo implementado, elimina las regiones en las cuales el número de pixeles con valores de reflectividad es menor de 50.

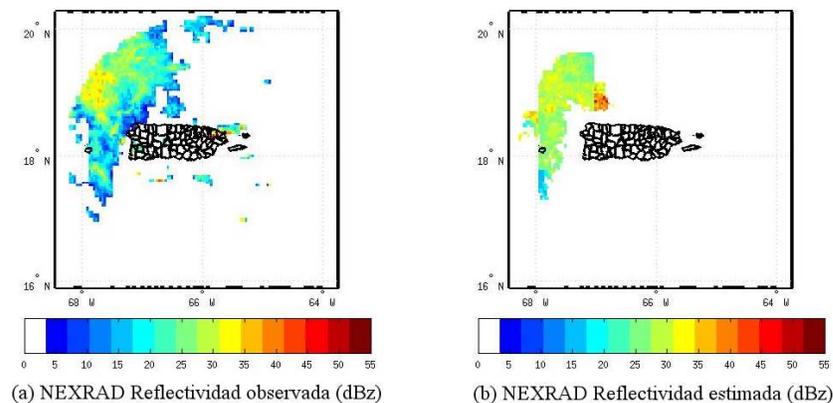


Figura 4-3: Reflectividad observada y estimada 30 min hacia adelante, para el evento del 2003/04/17 a las 01:15 horas en PR.

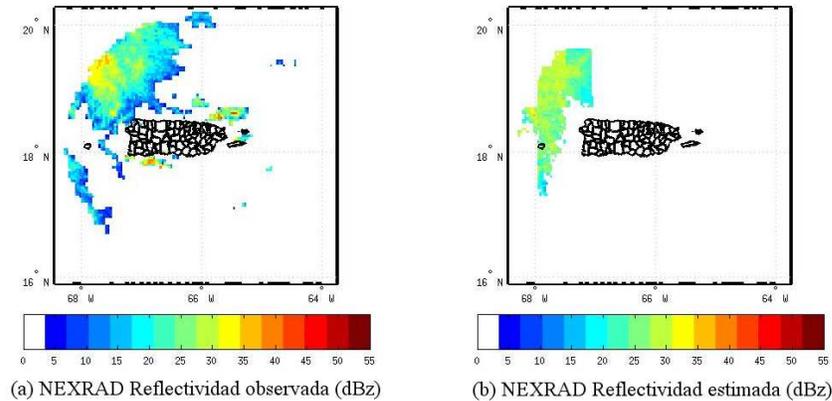


Figura 4-4: Reflectividad observada y estimada 60 min hacia adelante, para el evento del 2003/04/17 a las 01:45 horas en PR.

Finalmente en la Figura 4-5 mostrada a continuación se observa que la reflectividad observada (Figura 4-5 (a)), se encuentra muy dispersa en comparación con el momento inicial de medición (Figura 4-1). En ese momento se observa también presencia de precipitación en el nor-este de Puerto Rico, lo cual no es mostrado por la reflectividad estimada (Figura 4-5 (b)).

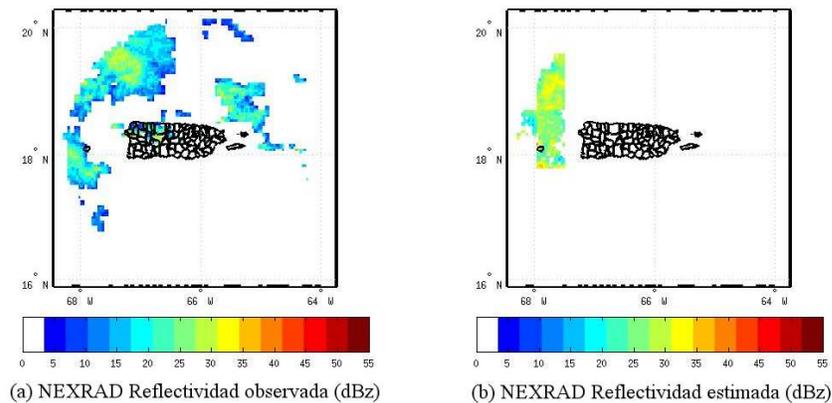


Figura 4-5: Reflectividad observada y estimada 120 min hacia adelante, para el evento del 2003/04/17 a las 02:45 horas en PR.

Las Figuras 4-6 a la 4-9, representan el contraste de reflectividad y validación cruzada de los eventos para cada uno de los tiempos de estimación (15, 30, 60 y 120 minutos) considerados en el estudio.

De la Figura 4-6 (a), se puede observar que la reflectividad estimada se superpone de manera ajustada a la observada; aunque, se puede ver que la amplitud de la curva de reflectividad de esta última es mayor. Con respecto a la gráfica de validación cruzada Figura 4-6 (b), asociada a este evento, se tiene que ambos valores de reflectividad se encuentran distribuidos de manera homogénea con respecto a la línea de 45° ; esto significa, que el promedio de los valores del error entre observados y predichos son muy cercanos a cero; eso nos indica, que la estimación es muy buena.

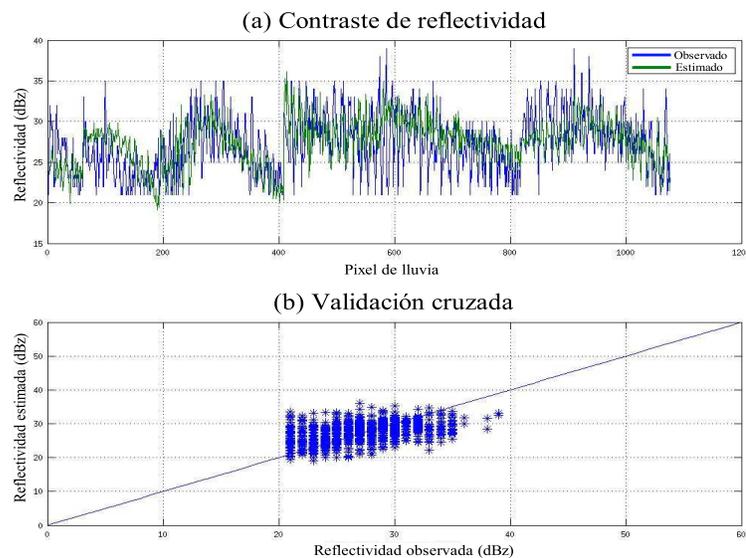


Figura 4-6: Contraste de reflectividad y validación cruzada para los valores observados y estimados 15 min hacia adelante, para el evento del 2003/04/17 a las 01:00 horas en PR.

La figura siguiente (Figura 4-7), muestra mayor variabilidad entre los valores de reflectividad estimados y observados, con respecto a la estimación a los 15 minutos. Vemos, por ejemplo, en la Figura 4-7 (a), que inicialmente (píxeles < 50 aprox), la reflectividad estimada es menor a la observada y tal como se había mostrado anteriormente en la Figura 4-3 (b) se observa una sobre-estimación de la reflectividad entre los píxeles de lluvia 700 y 780 aprox. De igual forma este hecho, se refleja en la Figura 4-7 (b) de validación cruzada, en donde se observan valores subestimados y sobre-estimados de reflectividad, con respecto a los demás píxeles que se mantienen dentro del rango aceptable de estimación.

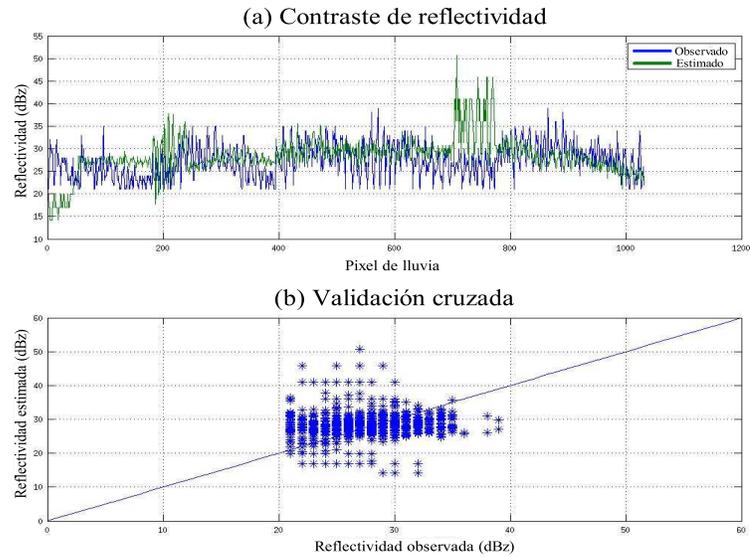


Figura 4–7: Contraste de reflectividad y validación cruzada para los valores observados y estimados 30 min hacia adelante, para el evento del 2003/04/17 a las 01:15 horas en PR.

En las Figura 4-8, se observa que la variación entre la reflectividad estimada y observada es mayor, destacando la presencia de una ligera subestimación de la reflectividad en los el rango de los pixeles 0-50 y del 600 en adelante; sin embargo la estimación sigue siendo relativamente buena, teniendo en cuenta que ésta se realiza a los 60 minutos.

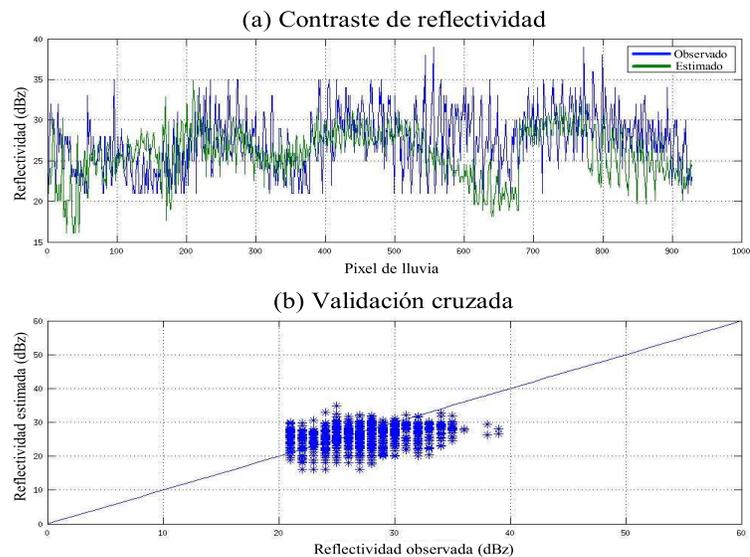


Figura 4–8: Contraste de reflectividad y validación cruzada para los valores observados y estimados 60 min hacia adelante, para el evento del 2003/04/17 a las 01:45 horas en PR.

Por último, tenemos la Figura 4-9 (a), en ella vemos que los valores de la reflectividad están sobre-estimados (píxeles < 120). De otro lado, tenemos también subestimación de la reflectividad (píxeles 150-200 y del 650-700). Lo anterior puede corroborarse al observar la Figura 4-9 (b); donde los valores se encuentran dispersos en torno a la recta de 45° ; nos indica, que los errores no son cercanos a cero y por tanto, la estimación no es buena en ese instante (120 minutos). Podemos anotar, que la estimación posterior a los 120 minutos, presenta fuertes deficiencias; por lo tanto el modelo RT no trabaja de manera óptima para predicciones mas allá de 120 minutos.

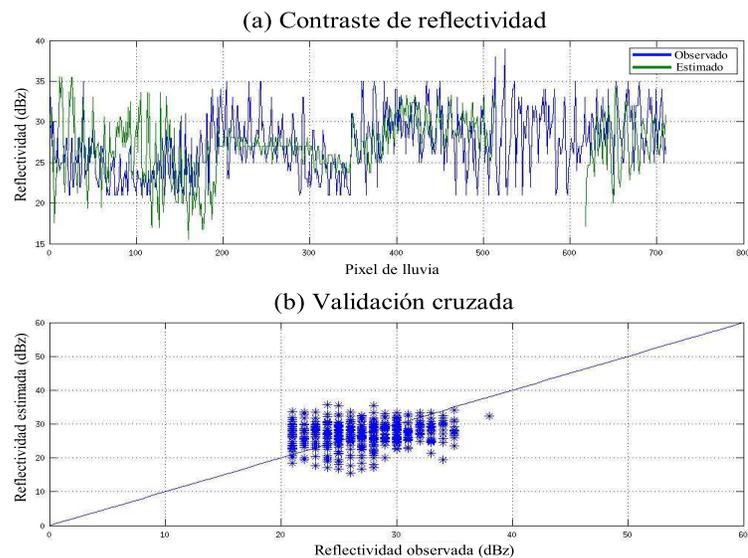


Figura 4-9: Contraste de reflectividad y validación cruzada para los valores observados y estimados 120 min hacia adelante, para el evento del 2003/04/17 a las 02:45 horas en PR.

Vectores de desplazamiento para el evento 1, ocurrido en Puerto Rico el 17 y 18 de Abril del 2003 a las 00:45 horas

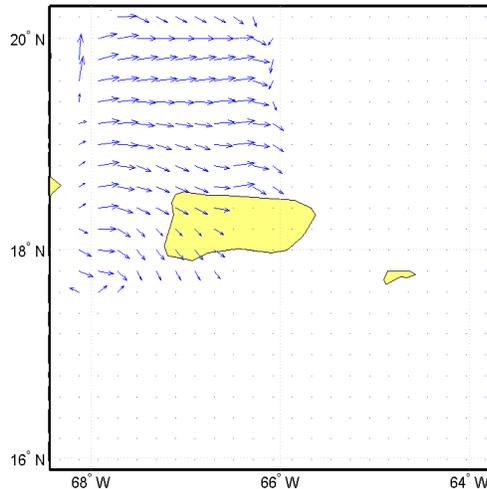


Figura 4–10: Vectores de desplazamiento de la nube, para el evento del 2003/04/17 a las 00:45 horas en PR.

Los vectores de desplazamiento para la Figura 4–10, cuya velocidad y dirección promedio son $v = 76,2$ Km/h y $\theta = -23,3^\circ$ aprox. respectivamente. Estos vectores, ilustran de forma más clara uno de los múltiples factores que causan las diferencias (grandes o pequeñas) entre los valores observados y los estimados de la reflectividad; en este sentido, es importante mostrar los dos procesos: crecimiento de la reflectividad causada por fenómenos físicos, que son propios de la evolución de la nube y el crecimiento o disipación de la nube debido al factor viento. Estos dos factores, se encuentran embebidos en el desarrollo de una tormenta, y es complicado separarlos utilizando únicamente las mediciones de la reflectividad que brinda el radar. No obstante, en el trabajo presente, se logran encontrar los vectores del desplazamiento, los cuales pueden diferir en poca medida con los vectores de viento. Podemos recalcar, que para la evaluación de futuros trabajos de predicción en espacio, tiempo y cantidad de lluvia, es importante incluir los vectores de desplazamiento de la tormenta.

**Resultados de la estimación de la reflectividad para el evento 2,
ocurrido en Puerto Rico
el 28 y 29 de octubre del 2007**

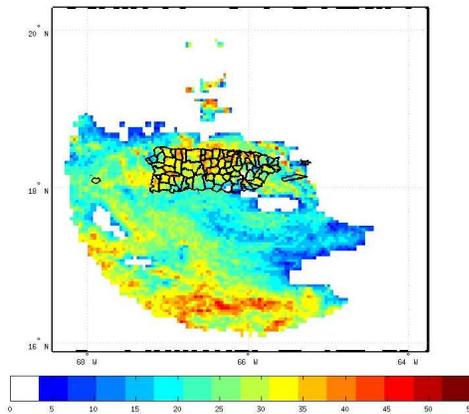


Figura 4–11: Reflectividad observada, para el evento del 2007/10/28 a las 00:00 horas en PR.

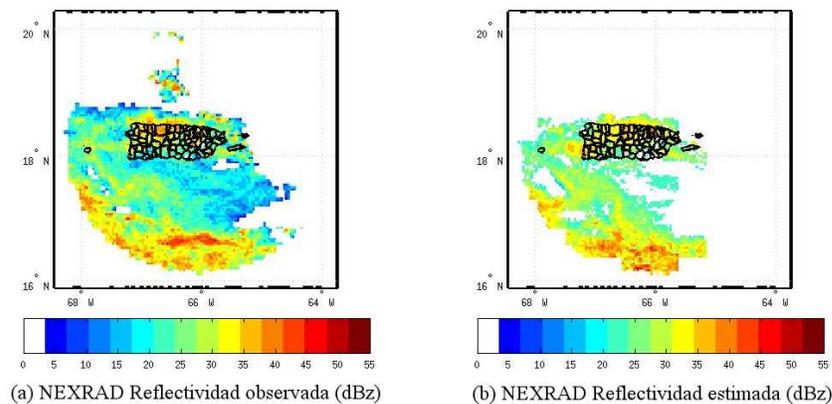


Figura 4–12: Reflectividad observada y estimada 15 min hacia adelante, para el evento del 2007/10/28 a las 12:15 horas en PR.

En la Figura 4–11, del límite en el cual se inició la estimación de la reflectividad (00.00 horas), observamos un gran área de fuertes lluvias sobre Puerto Rico y sus alrededores en particular en el sur de la isla ($16.5N, 66W$). De otro lado, en la Figura 4–12 (a), a los 15 minutos la reflectividad observada no muestra una variación significativa. Del mismo modo, la reflectividad estimada se asemeja mucho a la real (Figura 4–12 (b)); podemos notar que el modelo RT, es muy eficiente en la predicción de reflectividad en un corto intervalo de tiempo (15 minutos).

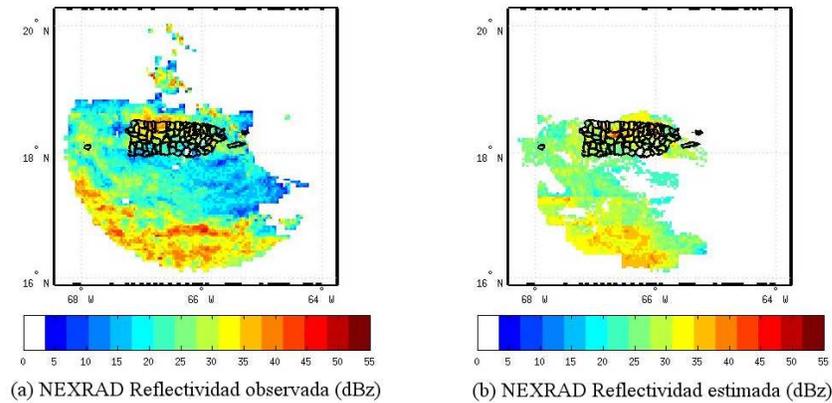


Figura 4-13: Reflectividad observada y estimada 30 min hacia adelante, para el evento del 2007/10/28 a las 12:30 horas en PR.

Notemos que no hay cambios significativos en la reflectividad observada (Figura 4-13 (a)) luego de un periodo de 30 minutos. Sin embargo, para este mismo lapso de tiempo, el modelo subestima la reflectividad; sobre todo en el área sur de la región de estudio (Figura 4-13 (b)). Por otro lado, en la Figura 4-14 (b), la reflectividad estimada, muestra una mayor variabilidad; esto se refleja en la sobre-estimación de la reflectividad en un área pequeña de la región de estudio, cuyas coordenadas son (16.5N, 67.5W).

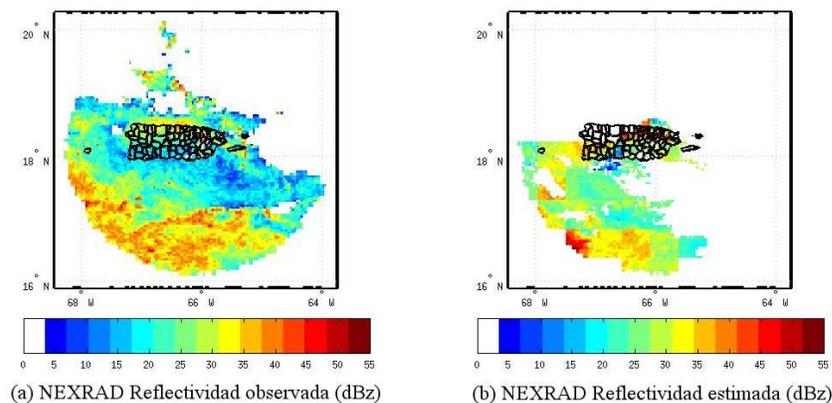


Figura 4-14: Reflectividad observada y estimada 30 min hacia adelante, para el evento del 2007/10/28 a las 00:00 horas en PR.

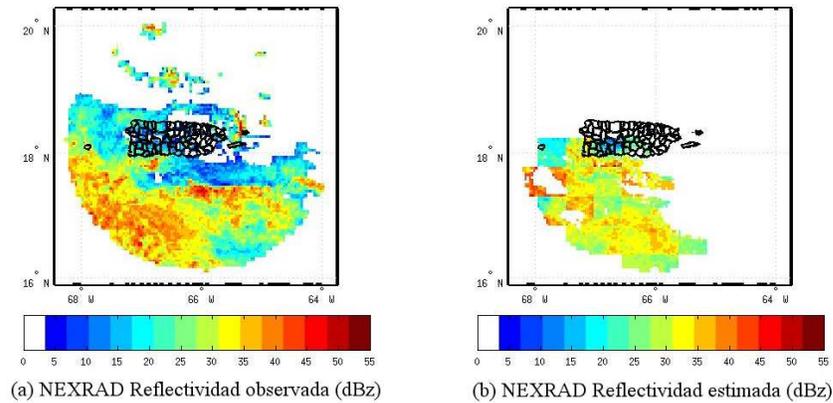


Figura 4–15: Reflectividad observada y estimada 120 min hacia adelante, para el evento del 2007/10/28 a las 02:00 horas en PR.

En la Figura 4–15(a), vemos un ligero cambio en la reflectividad observada con respecto a la anterior (Figura 4–14 (a)); mientras que la Figura 4–15 (b), muestra una mala estimación de la reflectividad; esto es; subestimación y sobre-estimación; distribuida en pequeñas áreas de la región de estudio.

A continuación, vemos la Figura 4–16 (a) y (b); en ella, se observa una muy buena estimación de la reflectividad. Nótese, que este evento, presenta precipitación mas copiosa con respecto al evento 1.

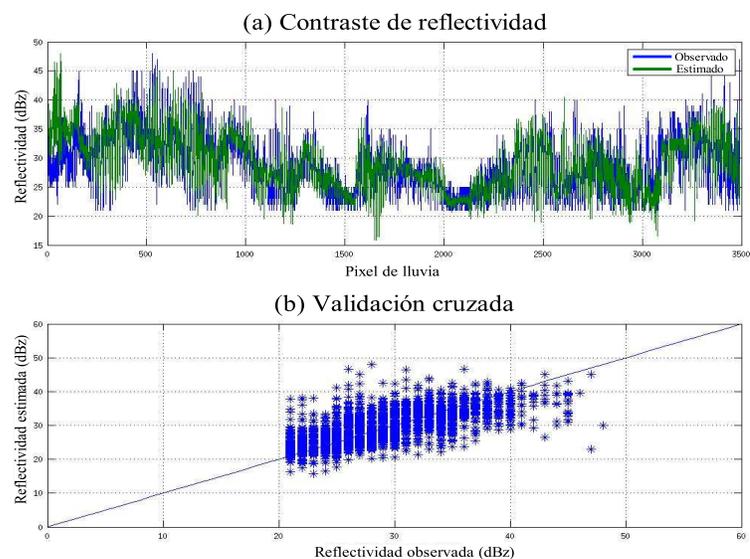


Figura 4–16: Contraste de reflectividad y validación cruzada para los valores observados y estimados 15 min hacia adelante, para el evento del 2003/12/7 a las 12:15 horas en PR.

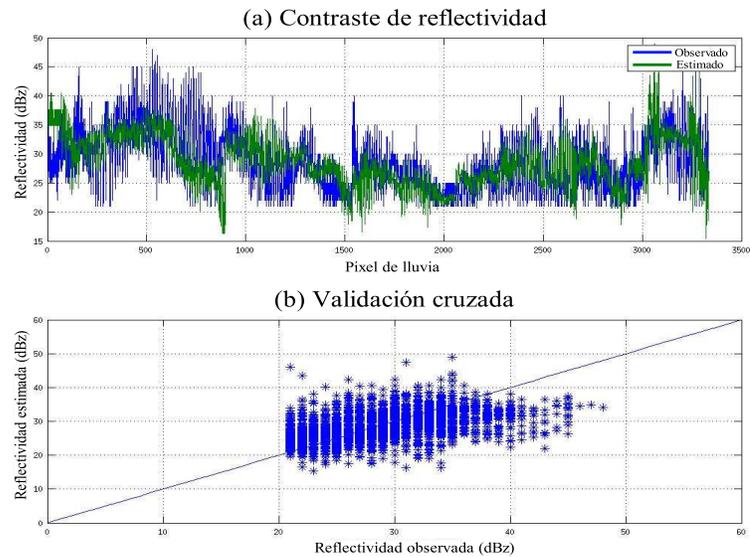


Figura 4–17: Contraste de reflectividad y validación cruzada para los valores observados y estimados 30 min hacia adelante, para el evento del 2003/12/7 a las 12:30 horas en PR.

El contraste de los valores de reflectividad observada vs estimada (Figura 4–17 (a)), presentan un buen ajuste a los 30 minutos; salvo en algunos tramos de la curva. Esto también se refleja en la grafica de validación cruzada (Figura 4–17 (b)). Por otro lado, el ajuste de las curvas de estimación vs observados en la Figura 4–18 (a), es aceptable.

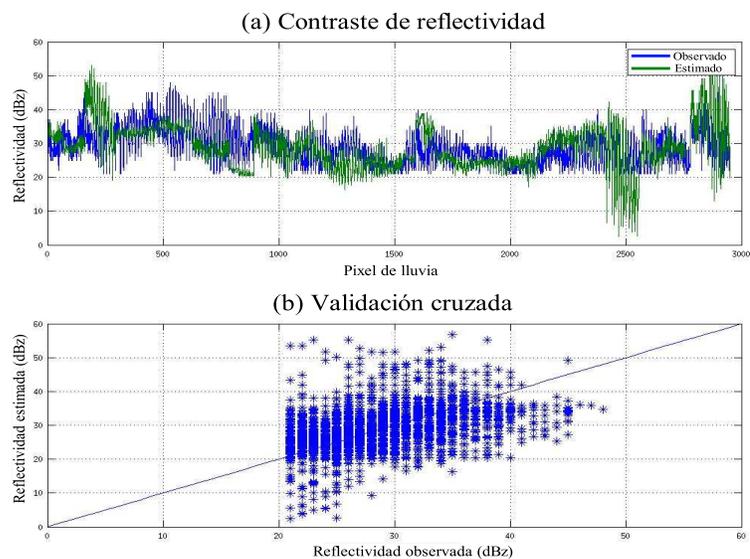


Figura 4–18: Contraste de reflectividad y validación cruzada para los valores observados y estimados 60 min hacia adelante, para el evento del 2003/12/7 a las 01:00 horas en PR.

No obstante, existen sobre-estimaciones en el rango de pixeles del 200-300 y 2800-2950; y la subestimación en el rango 2450-2550. Lo anterior, se corrobora en la Figura 4-18 (b), donde hay una gran dispersión de los valores estimados con respecto a los observados; esto es: el promedio de los errores no es cercano a cero.

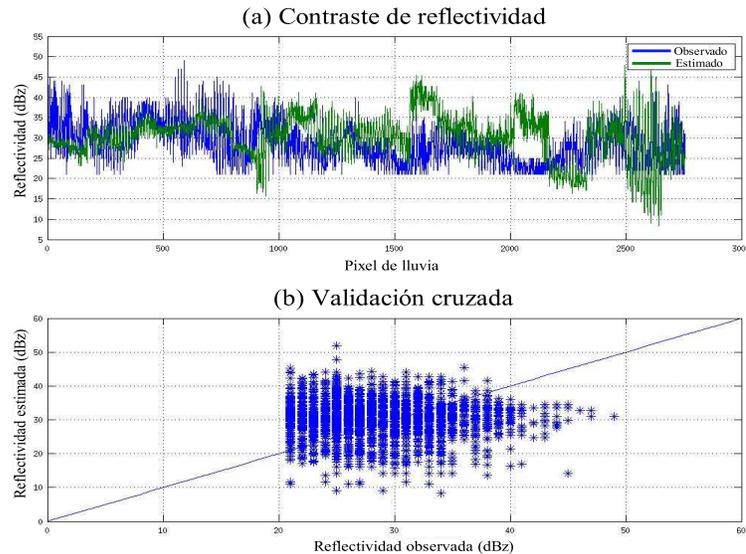


Figura 4-19: Contraste de reflectividad y validación cruzada para los valores observados y estimados 120 min hacia adelante, para el evento del 2003/12/7 a las 02:00 horas en PR.

Por último, tenemos la Figura 4-18 (a), en la cual se observa el contraste de la reflectividad a los 120 minutos, es alto; ya que ambas curvas se traslapan únicamente para los pixeles < 850 . Además, el modelo evidencia una sobre-estimación notable que llega al punto máximo de 45 dBz , en una pequeña vecindad del pixel 1750. En la Figura 4-18 (b), vemos que la sobre-estimación se hace notable; además, se observa una considerable cantidad de puntos con desviaciones positivas. Así, concluimos que la estimación a los 120 minutos no es buena; y por tanto no se recomienda el modelo RT para la estimación de la lluvia (reflectividad) más allá de los 120 minutos.

Vectores de desplazamiento para el evento 2, ocurrido en Puerto Rico
el 28 y 29 de octubre del 2007 a las 00:00 horas,

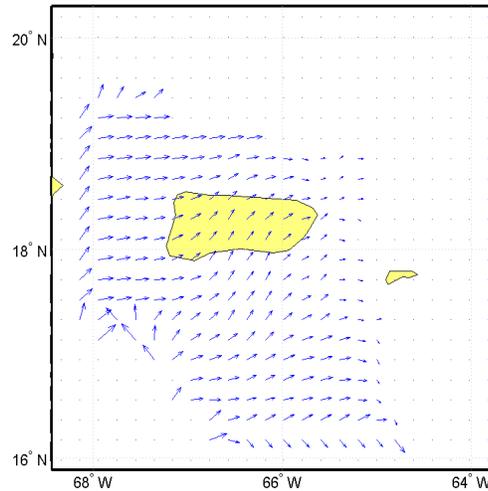


Figura 4–20: Vectores de desplazamiento de la nube, para el evento del 2007/10/28 a las 00:00 horas en PR.

La Figura 4–20, muestra los vectores de desplazamiento de la nube; cuya velocidad y dirección promedio son $v = 93,4$ Km/h y $\theta = 37,5^\circ$ aprox. respectivamente. Nuevamente, (como se dijo anteriormente para la Figura 4–10) podemos resaltar, que estos vectores **no** representan una estimación del desplazamiento en el tiempo. Estos fueron desarrollados para tener una "idea" de cual podría ser el desplazamiento (velocidad y dirección) de la nube en un instante del tiempo. La intuición nos dice que, basados en periodos de tiempo corto (< 2 horas), estos vectores no presentan regularmente cambios significativos en su magnitud y dirección. Sin embargo, estos vectores constituyen una primera fase para desarrollar un modelo en el cual se puedan combinar los dos factores importantes de la lluvia (intensidad y dirección) y otros variables como temperatura ambiente y reflectancia; los cuales permitan realizar predicciones conjuntas de intensidad de la lluvia y velocidad de la nube.

Capítulo 5

CONCLUSIONES Y TRABAJOS FUTUROS

5.1. Conclusiones

- En este trabajo, se discutieron aspectos relacionados con la estimación de la cantidad de lluvia precipitada y vectores de desplazamiento en los eventos (tormentas) del 17-18 de abril del 2003 y 27-29 de octubre del 2007 presentados en la región de Puerto Rico. Se logró crear un modelo (RT) óptimo para la estimación de la lluvia a corto plazo; así como también, diseñar un algoritmo que permite modelar el campo de vectores de desplazamiento de la nube.
- El modelo RT tiene un desempeño óptimo en la estimación de la lluvia a corto plazo; sin embargo, este se vuelve deficiente para periodos que excedan los 120 minutos.
- El algoritmo creado para la estimación de la lluvia, presenta un porcentaje de decremento del grado de variabilidad explicada por cada modelo en aproximadamente 20 % en promedio para cada incremento en el periodo de estimación de 15, 30 y 60 minutos,
- El uso de herramientas como las imágenes con mayor número de píxeles y la eliminación de las regiones en las cuales la tormenta sea menor a un área específica, resulto útil para el mejor desempeño del algoritmo de desplazamiento de la nube.
- La posibilidad de crear un modelo en el cual se pueda utilizar la función de transferencia, mostraría una mejora significativa en la estimación de la lluvia a corto plazo; ya que este tendría en cuenta una sucesión de datos en periodos mas cortos

de tiempo; que en el caso precedente, no fue posible implementarlo debido a la falta de estos.

- El trabajo de realizar predicciones de la intensidad y desplazamiento de las nubes en una tormenta, será mucho mas fructífero con el apoyo y cooperación directa de personas que se desempeñen en diferentes áreas del conocimiento (física, matemáticas, estadística, meteorología y computación) y así, lograr la construcción de un modelo lo suficientemente confiable y un algoritmo óptimo, robusto y rápido que se pueda implementar de forma sencilla.

5.2. Trabajos Futuros

Los siguientes son algunos aspectos que se debe implementar para mejorar el rendimiento del modelo RT o aplicar lo que se ha trabajado en la consecución de nuevos modelos mas precisos y rápidos.

- Utilizar otras variables para el modelo (humedad relativa, reflectancia, velocidad del viento, etc.) y buscar una clasificación de los pixeles donde se encuentre la reflectividad, para lograr una mejor adaptación del modelo RT.
- Aplicar el modelo FET; en la medida de las posibilidades brindadas por el radar que se instalará en el RUM, para el año 2013, que tendrá mediciones en intervalos de 2 minutos.
- Encontrar un modelo Bayesiano, en el que se tengan en cuenta la distribución de probabilidad apriori de intensidad de lluvia en una región y la magnitud y dirección de los vectores de desplazamiento en una imagen Z_{t-1} , para luego hacer inferencia en estos dos aspectos.

APÉNDICES

Apéndice A

ALGORITMOS

A.1. Algoritmos para encontrar vectores de desplazamiento

A.1.1. Función que interpola una matriz usando el método PMA.

```

1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  %% Función que interpola una matriz usando el método del desplazamiento %%
3  %% del punto medio aleatorio. %%
4  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
5
6  function B=pma(B1)
7
8  n=size(B1);
9  B=zeros(2*n(1)-1,2*n(2)-1,n(3));
10 for k=1:n(3)
11     m = B1(:, :, k);
12     M = -2*ones(2*n(1)-1,2*n(2)-1);
13     for i=1:n(1)
14         for j=1:n(2)
15             M(2*i-1,2*j-1)=m(i, j);
16         end
17     end
18     for i=1:n(1)-1
19         for j=1:n(2)-1
20             M(2*i,2*j)=(m(i, j)+m(i+1, j)+m(i, j+1)+m(i+1, j+1))/4+0.6089*randn;
21         end
22     end
23     for i=1:n(2)-1
24         M(1,2*i)=(m(1, i)+m(1, i+1))/2+0.4471*randn;
25     end
26     for i=1:n(1)-1
27         M(2*i,1)=(m(i,1)+m(i+1,1))/2+0.4471*randn;
28     end
29     for i=1:n(2)-1
30         M(2*n(1)-1,2*i)=(m(n(1), i)+m(n(1), i+1))/2+0.4471*randn;
31     end
32     for i=1:n(1)-1
33         M(2*i,2*n(2)-1)=(m(i, n(2))+m(i+1, n(2)))/2+0.4471*randn;
34     end
35     for i=1:2*n(1)-1
36         for j=1:2*n(2)-1
37             if(M(i, j)==-2)
38                 t=0.6089*randn*(1/sqrt(2))^(5/3);
39                 M(i, j)=(M(i, j+1)+M(i, j-1)+M(i+1, j)+M(i-1, j))/4+t;
40             end
41         end
42     end
43 B(:, :, k)=M;
44 end

```


A.1.4. Función que genera el vector asociados a cada pixel.

```

1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  %% Función que genera el vector de valores asociados a cada pixel. %%
3  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4
5  function [aso1] = aso(A1,i,j)
6
7  [m,n]=size(A1); % dimensión de la matriz A
8  A=A1(:, :);
9  % aso1= pixeles asociados al pixel ij
10
11  if j==1
12      if i==1
13          aso1=[A(i,j) A(i+1,j) A(i+1,j) A(i+1,j+1)];
14      elseif i==m
15          aso1=[A(i,j) A(i-1,j) A(i,j+1) A(i-1,j+1)];
16      else
17          aso1=[A(i,j) A(i-1,j) A(i-1,j+1) A(i,j+1) A(i+1,j+1) A(i+1,j)];
18      end
19  elseif j==n
20      if i==1
21          aso1=[A(i,j) A(i,j-1) A(i+1,j) A(i+1,j-1)];
22      elseif i==m
23          aso1=[A(i,j) A(i,j-1) A(i-1,j) A(i-1,j-1)];
24      else
25          aso1=[A(i,j) A(i-1,j) A(i,j-1) A(i-1,j-1) A(i+1,j-1) A(i+1,j)];
26      end
27  else
28      if i==1
29          aso1=[A(i,j) A(i,j-1) A(i+1,j-1) A(i+1,j) A(i+1,j+1) A(i,j+1)];
30      elseif i==m
31          aso1=[A(i,j) A(i,j-1) A(i-1,j-1) A(i-1,j) A(i-1,j+1) A(i,j+1)];
32      else
33          aso1=[A(i,j) A(i-1,j-1) A(i-1,j) A(i-1,j+1) A(i,j-1) A(i,j+1)
34                A(i+1,j-1) A(i+1,j) A(i+1,j+1)];
35      end
36  end

```

A.1.5. Función que genera los vectores de desplazamiento.

```

1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  %% Función que genera las matrices ZU y ZV de vectores de desplazamiento. %%
3  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
4
5  % b=9; % dimension del bloque
6  % dimras=19; % dimension del rastreador
7  % B: Arreglo de matrices aumentadas de ZI
8
9  function [ZU,ZV]=vectores(B, dimras, b)
10
11  n=size(B);
12  a1=floor(dimras/2)+1; % centro del rastreador
13  a=floor(b/2)+1; % centro del bloque
14  radbloque=floor(b/2); % radio del bloque
15  radiatoras=floor(dimras/2); % radio del rastreador
16  c1=floor(n(1)/b); % c1 y c2 dimensiones de la matriz ZU y ZV
17  c2=floor(n(2)/b);
18
19  %% Indices de los centros de cada bloque Bt: indi, indj para la matriz ZI
20  indi=zeros(c1,1);
21  for i=1:c1
22      indi(i)=a+(i-1)*b;
23  end
24
25  indj=zeros(c2,1);
26  for j=1:c2
27      indj(j)=a+(j-1)*b;

```

```

28 end
29
30 k1=size(indi(1)-(radbloque+radoras):indi(c1)+radbloque+radoras);
31 k2=size(indj(1)-(radbloque+radoras):indj(c2)+radbloque+radoras);
32 A=zeros(k1(:,2),k2(:,2),n(3));
33 A(a1:a1+n(1)-1,a1:a1+n(2)-1,1:n(3))=B(1:n(1),1:n(2),1:n(3));
34 n1=size(A);
35
36 %% Índices(indaumi,indaumj) de centros para bloque de la matriz A aumentada
37 indaumi=zeros(c1,1);
38 for i=1:c1
39     indaumi(i)=a+(i-1)*b+radoras;
40 end
41
42 indaumj=zeros(c2,1);
43 for j=1:c2
44     indaumj(j)=a+(j-1)*b+radoras;
45 end
46
47 %% En esta parte se encuentran los índices de los bloques
48 ind_bloquei=zeros(b,1,c1,c2);
49 ind_bloquej=zeros(b,1,c1,c2);
50 [Y,X]=meshgrid(1:n1(2),1:n1(1));
51
52 for i=1:c1
53     for j=1:c2
54         ind_bloquei(:,1,i,j)=reshape(X(indaumi(i)-radbloque:indaumi(i)+radbloque,
55                                     indaumj(j)-radbloque),b,1);
56         ind_bloquej(:,1,i,j)=reshape(Y(indaumi(i)-radbloque,
57                                     indaumj(j)-radbloque:indaumj(j)+radbloque),b,1);
58     end
59 end
60
61 %% Índices del rastreador: rastreadorij
62
63 [Y1,X1]=meshgrid(1:b,1:b);
64 ind_rastreadori=reshape(X1(1:b,1),b,1);
65 ind_rastreadorj=reshape(Y1(1,1:b),b,1);
66 rastreadorij=zeros(b*b,2);
67     k=0;
68     while k<b
69         rastreadorij(k*b+1:(k+1)*b,1)=ind_rastreadori(k+1)*ones(b,1);
70         rastreadorij(k*b+1:(k+1)*b,2)=ind_rastreadorj;
71         k=k+1;
72     end
73 centroras=[a*ones(b*b,1) a*ones(b*b,1)];
74
75 %% En esta parte eliminamos los NaN y los valores de reflec < constante
76 for i=1:n1(3)
77     for i1=1:n1(1)
78         for j1=1:n1(2)
79             if isnan(A(i1,j1,i))==1 || A(i1,j1,i)<25
80                 A(i1,j1,i)=0;
81             end
82         end
83     end
84 end
85
86 %% Correlaciones para los índices de cada bloque
87 % bloques: bloques de la matriz A1 en el tiempo t
88
89 corrBt=zeros(b*b,1,c1,c2,n1(3)); % correlaciones de cada bloque
90     t=2;
91 while t<n1(3)
92     for i=1:c1 % ciclos for : recorrerán los bloques de A.
93         for j=1:c2
94             bloques=[ind_bloquei(:,1,i,j),ind_bloquej(:,1,i,j)];

```

```

95     corrBt (:,1,i,j,t-1)=Rpq(A,t,b,bloques,dimras,radioras);
96     end
97     end
98     t=t+1;
99 end
100
101 %% Dirección velocidad y ángulo del vector (ZU,ZV) de desplazamiento
102 % Las siguientes son las matrices correspondientes a cada elemento
103 % del vector de desplazamiento
104
105 ZU1=zeros(b*b,c1,c2,nl(3)-1); % matriz de indices i
106 ZV1=zeros(b*b,c1,c2,nl(3)-1); % matriz de indices j
107 vel1=zeros(b*b,c1,c2,nl(3)-1); % matriz de velocidad
108 angulo1=zeros(b*b,c1,c2,nl(3)-1); % matriz de angulos
109 maxBtcorr1=zeros(b*b,c1,c2,nl(3)-1); % matriz de correlaciones
110 ZU=zeros(c1,c2,nl(3)-1); % matriz de indices i
111 ZV=zeros(c1,c2,nl(3)-1); % matriz de indices j
112 vel=zeros(c1,c2,nl(3)-1); % matriz de velocidad
113 angulo=zeros(c1,c2,nl(3)-1); % matriz de angulos
114 maxBtcorr=zeros(c1,c2,nl(3)-1); % matriz de correlaciones
115
116 % ciclos que recorran los bloques de A para llenar las
117 % matrices de correlaciones, velocidad, ángulo y vectores
118 % desplazamiento
119 t=1;
120 while t<nl(3)
121     for i=1:c1
122         for j=1:c2
123             [maxBt,IX]=sort(corrBt(:,1,i,j,t),'descend');
124             indicesij=rastreadorij(IX(:,:),:);
125             [v x y theta]=indmaxangulo(indicesij,centroras,maxBt,b);
126             maxBtcorr1(:,i,j,t)=maxBt;
127             vel1(:,i,j,t)=v;
128             angulo1(:,i,j,t)=theta;
129             ZU1(:,i,j,t)=x;
130             ZV1(:,i,j,t)=y;
131         end
132     end
133     t=t+1;
134 end
135
136     maxBtcorr(:,:,:)=maxBtcorr1(1,:,:,:);
137     vel(:,:,:)=vel1(1,:,:,:);
138     angulo(:,:,:)=angulo1(1,:,:,:);
139     ZU(:,:,:)=ZU1(1,:,:,:);
140     ZV(:,:,:)=ZV1(1,:,:,:);

```

A.1.6. Función que genera los gráficos de vectores desplazamiento de la nube

```

1  %%%%%%%%%%%
2  %% Rutina que genera los gráficos desplazamiento %%
3  %%%%%%%%%%%
4
5  clear
6  clc
7  close
8  load ZI1
9  n2=size(ZI1);
10
11  ZI2=dpma(ZI1(:, :, 100:120));
12  % [ZU,ZV]=vectores1(B1,dimras,b)
13  % b=7; % dimension del bloque
14  % dimras=15; % dimension del rastreador

```

```

15 n1=size(ZI2);
16 B1=ZI2(:, :, 1:n1(3));
17 %B=pma(B1);
18 [ZU,ZV]=vectores(B1,23,7);
19 n=size(ZU);
20 res=2; % Resolucion de la imagen
21 lim=[-68.43 -63.73;15.89 20.3]; % Limites de lat y lon de la imagen
22 dLa=(lim(2,2)-lim(2,1))/n(1); % Incremento de lat y lon de la imagen
23 dLo=(lim(1,2)-lim(1,1))/n(2);
24 La=lim(2,1)+0.5*dLa:dLa:lim(2,2)-0.5*dLa; % vector final de lat para ubicar los
25 % centros de los vectores de desplazamiento
26 Lo=lim(1,1)+0.5*dLo:dLo:lim(1,2)-0.5*dLo;
27 [XR,YR]=meshgrid(Lo,La); % Rejilla de lat y lon final
28
29 % Rejilla de lat y lon de la imagen ampliada
30 [XI,YI] = meshgrid(lim(1,1):res/deg2km(1):lim(1,2),lim(2,1):res/deg2km(1):lim(2,2));
31
32 for i=1:n(3)
33     figure(i)
34     subplot(1,2,1)
35     worldmap(limites(2,:),limites(1,:)) % Las esquinas del mapa
36     setm(gca,'mapprojection','mercator') % Cambiar proyección del globo a rectangular
37     geoshow('landareas.shp') % Superpone el mapa del satélite.
38     geoshow(YI,XI,ZI2(:, :, i),'DisplayType','texturemap')
39
40     subplot(1,2,2)
41     worldmap(limites(2,:),limites(1,:)) % Las esquinas del mapa
42     setm(gca,'mapprojection','mercator') % Cambiar proyección del globo a rectangular
43     geoshow('landareas.shp') % Superpone el mapa del satélite
44     quiverm(YR,XR,ZU(:, :, i),ZV(:, :, i));
45     saveas(figure(i),['subplot_',int2str(i),'.png'])
46     close
47 end

```



```

65 %This is command (LINEM(Longitude, Latitude))
66     for j=1:78 %Puerto Rico tiene 78 poblaciones.
67         %Usar la función de límites por población
68         linem(S(j).Y,S(j).X, 'k', 'linewidth',1) %
69     end
70     caxis([0 55])
71     subplot(1,3,2)
72     worldmap(latlim, lonlim);
73     surfacem(La,Lo,Zo);
74     setm(gca, 'mapprojection', 'mercator')
75 %%%%%%%%%%%
76 %This is command (LINEM(Longitude, Latitude))
77     for j=1:78
78         linem(S(j).Y,S(j).X, 'k', 'linewidth',1)
79     end
80     caxis([0 55])
81 %%%%%%%%%%%
82     subplot(1,3,3)
83     worldmap(latlim, lonlim);
84     surfacem(La,Lo,Ze);
85     setm(gca, 'mapprojection', 'mercator')
86 %%%%%%%%%%%
87 %This is command (LINEM(Longitude, Latitude))
88     for j=1:78
89         linem(S(j).Y,S(j).X, 'k', 'linewidth',1)
90     end
91     caxis([0 55])
92 saveas(figure(1), ['Compare_Reflectivity_year_'_mes_'_day_'_hour_'_minute'.jpg'])
93
94     figure(2)
95     subplot(2,1,1)
96     plot([outdata estdata]);
97     grid
98     subplot(2,1,2)
99     plot(outdata, estdata, '*');
100    grid
101 saveas(figure(2), ['Function_Reflectivity_year_'_mes_'_day_'_hour_'_minute'.jpg'])
102 close all
103 end

```

A.2.2. Función que genera la interpolación en las regiones de entrada.

```

1 %%%%%%%%%%%
2 %% Función que genera las regiones de entrada y se interpolan %%
3 %% los valores faltantes de cada bloque para el algoritmo principal %%
4 %% Algoritmo diseñado por Joan Castro %%
5 %%%%%%%%%%%
6
7 function winzone=rainfall_study_area(dataserie, area, La, Lo, res, winsel, winvar)
8
9 [row, col]=size(La);
10
11 limlat=minmax(La(:,1)');
12 limlon=minmax(Lo(1,:));
13
14 [Lo01, La01]=meshgrid(limlon(1)-2*res:res:limlon(2)+2*res, limlat(1)
15                     -2*res:res:limlat(2)+2*res);
16 [Lo02, La02]=meshgrid(limlon(1)-4*res:res:limlon(2)+4*res, limlat(1)
17                     -4*res:res:limlat(2)+4*res);
18
19 colref=1:area:col;
20 rowref=1:area:row;
21
22 indref=1:row*col;
23 indref=reshape(indref, row, col);

```

```

24
25 indref01=1:(row+4)*(col+4);
26 indref01=reshape(indref01',row+4,col+4);
27 indref01=indref01(3:row+2,3:col+2);
28
29 indref02=1:(row+8)*(col+8);
30 indref02=reshape(indref02',row+8,col+8);
31 indref02=indref02(5:row+4,5:col+4);
32
33
34 data=dataserie{3,1};
35 lag01=zeros(row+4,col+4);
36 lag01(3:row+2,3:col+2)=dataserie{2,1};
37 lag02=zeros(row+8,col+8);
38 lag02(5:row+4,5:col+4)=dataserie{1,1};
39
40 winzone={};
41 count=1;
42
43 for i=1:length(rowref)-2
44     for j=1:length(colref)-2
45         study=data(rowref(i):rowref(i+1)-1,colref(j):colref(j+1)-1);
46         zone=indref(rowref(i):rowref(i+1)-1,colref(j):colref(j+1)-1);
47         zone01=indref01(rowref(i):rowref(i+1)-1,colref(j):colref(j+1)-1);
48         zone02=indref02(rowref(i):rowref(i+1)-1,colref(j):colref(j+1)-1);
49         winner=find(study>20);
50         outdata=[];
51         indata=[];
52         if length(winner)>=50
53             for k=1:length(winner)
54                 [x1,y1]=ind2sub([row+4 col+4],zone01(winner(k)));
55                 [x2,y2]=ind2sub([row+8 col+8],zone02(winner(k)));
56                 x1=x1-2:x1+2;
57                 y1=y1-2:y1+2;
58                 x1=repmat(x1',5,1);
59                 y1=reshape(repmat(y1,5,1),25,1);
60                 vector01=sub2ind([row+4 col+4],x1,y1);
61                 x2=x2-4:x2+4;
62                 y2=y2-4:y2+4;
63                 x2=repmat(x2',9,1);
64                 y2=reshape(repmat(y2,9,1),81,1);
65                 vector02=sub2ind([row+8 col+8],x2,y2);
66         if length(find(lag01(vector01)>0))>=5 && length(find(lag02(vector02)>0))>=9
67             istudy01=krig_version2([La01(vector01) Lo01(vector01) lag01(vector01)],0);
68             istudy02=krig_version2([La02(vector02) Lo02(vector02) lag02(vector02)],0);
69             indata=[indata; fix(istudy01(:,3)') fix(istudy02(:,3)')];
70             outdata=[outdata; study(winner(k))];
71
72         end
73     end
74     [var, indatad]=division_sel(winsel, winvar, indata, outdata, 0.95);
75
76     if var(1)==0
77
78         winzone{count,1}=[i j];
79         winzone{count,2}=mean(outdata);
80         winzone{count,3}=var;
81         winzone{count,4}=0;
82         winzone{count,5}=0;
83         count=count+1;
84     else
85         [indataw, vars, beta, R_2, VIF, p_val]=variance_factor(outdata, indatad);
86
87         winzone{count,1}=[i j];
88         winzone{count,2}=beta;
89         winzone{count,3}=var(vars);
90         winzone{count,4}=R_2;
91         winzone{count,5}=p_val;

```

```

92         count=count+1;
93
94     end
95
96     end
97 end
98 end

```

A.2.3. Función para interpolación en las regiones.

```

1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  %% Función que genera las regiones de entrada y se interpolan %%
3  %% Algoritmo diseñado por Joan Castro   %%
4  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
5
6  function [data_f] = krig_version2(data , miss_v)
7
8  %'DATA' es una matriz que tiene tres columnas. Las primeras dos
9  % columnas corresponden a la localizacion de los datos (posicion geografica).
10 % La ultima columna corresponde a la
11 % informacion reservada en la posicion.
12
13 %'miss_v' es un codigo numerico que indica un dato no tabulado
14 %'missing value '.
15
16 data_f=data;
17 % Definicion de variables y matrices
18
19 reg1=1;           % Punto inicial para el contador de data no tabulador
20 reg2=reg1;       % Punto inicial para el contador de data tabulada
21 V=[];            % Registro de la data tabulada (9x3)
22 un_loc=[];       % Contador de la data no tabulada (3)
23 un_data=[];      % Registro de la data no tabulada (3x2)
24 matriz_V=[];    % Variograma : registro entre las distancias conocidas (9x9)
25 matriz_U=[];    % Registro entre las distancias desconocidas (9x9)
26 Z_value=[];     % Solucion de la interpolacion espacial (3)
27 eq1=[];         % Añade una columna al variograma (landa)
28 eq2=[];         % Añade una fila al variograma [U] (w1+w2+w3=1)
29 inv_matriz_V=[]; % Matriz inversa del variograma [V^-1] (10x10)
30 Pesos_W=[];     % Solucion de los Pesos en el variograma [W= V^-1 * U] (10x3)
31
32 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
33 %Primer Paso: Hallar la localizacion de los datos no registrados
34 %%
35 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
36 %% Rutina que separa la posicion con los datos conocidos y la
37 %% Posicion con los datos desconocidos.
38
39 estructura=size(data);
40
41 for count=1:estructura(1) %Tamaño del banco de datos
42     %Evalua si el dato fue o no fue tabulado
43     if(data(count,estructura(2))==miss_v) %Si el dato no fue tabulado
44         un_data(reg1,:)=data(count,1:estructura(2)-1);
45         %Guarda la posicion no tabulada
46         un_loc(reg1)=count; %Cuenta la cantidad de datos no tabulados
47         reg1=reg1+1;
48     else %Si el dato fue tabulado
49         V(reg2,:)=data(count,:); %Guarda la posicion tabulada
50         reg2=reg2+1;
51     end
52 end
53
54 m_any=size(un_data); %El tamaño de la matriz de los datos incompletos.
55

```

```

56 if m_any(1)==0
57     data_f=data;
58
59 else
60
61     %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
62     %Segundo Paso: Organización de las matrices V y U
63     %Para hallar los valores estimados
64     %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
65
66     m_any2=size(V);
67     matriz_V=dist(V(:,1:estructura(2)-1)');
68     matriz_U=dist(un_data,V(:,1:estructura(2)-1)');
69
70     m_any3=size(matriz_V);
71
72     %Genera el variograma
73     for count=1:length(matriz_V)
74         eq1(:,count)=[matriz_V(:,count);1];
75     end
76
77     %Añade la columna de la variable error (landa)
78     %eq1(:,length(matriz_V)+1)= [ones(length(matriz_V),1);0];
79     eq1(:,m_any3(1)+1)=[ones(m_any3(1),1);0];
80     inv_matriz_V=inv(eq1);
81
82     matriz_U=matriz_U';
83     %Añade la fila de la variable error landa (w1+w2+w3=1)
84     eq2=[matriz_U;ones(1,m_any(1))];
85
86     %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
87     %Tercer Paso: Guardar los valores estimados en el nuevo banco de datos 'data'
88     %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
89
90     %sumatoria de los datos interpolados espacialmente
91     for check=1:m_any(1) %Cantidad de data no tabulados
92         Pesos_W(:,check)= inv_matriz_V * eq2(:,check);
93         %valores de los pesos por cada dato no tabulado
94
95         Z_value(check) = Pesos_W(1:m_any3(1),check)'*V(:,estructura(2));
96         %Interpolaciones de datos espaciales
97         data_f(un_loc(check),estructura(2))=Z_value(check);
98         %Sustitucion de los datos no tabulados al os datos tabulados
99     end
100
101 end

```

A.2.4. Función interpolación de Kriging.

```

1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  %% Función interpolación de Kriging. %%
3  %% Algoritmo diseñado por Joan Castro   %%
4  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
5
6  function fdata=krig_knn(data,vecino,miss_v)
7
8  n=size(data);
9  nodata=find(data(:,n(2))==miss_v);
10 yesdata=find(data(:,n(2))~=miss_v);
11
12 if length(nodata)==0
13     fdata=data;
14 else
15     fdata=data;
16     mdata=data(nodata,:);

```

```

17 vdata=data(yesdata ,:);
18 for counter=1:length(nodata)
19     distancia=dist(vdata(:,1:n(2)-1),mdata(counter,1:n(2)-1)');
20     [valor,kvecinos]=sort(distancia);
21     indata=vdata(kvecinos(1:vecino),:);
22     n2=size(indata);
23     outdata=krig_version2([indata;mdata(counter,:)],miss_v);
24     fdata(nodata(counter),n(2))=outdata(n2(1)+1,n(2));
25 end
26 end

```

A.2.5. Función multicolumna.

```

1  %%%%%%%%%%%
2  %% Función multicolumna %%
3  %% Algoritmo diseñado por Joan Castro   %%
4  %%%%%%%%%%%
5  function [Xsel_f,beta,gamma,elima,vars,R_2]=multimulticol(Y,Xa, restr)
6
7  n1=size(Xa);
8  if n1(2)==1
9      if Xa(:,1)==ones(n1(1),1);
10         Xsel_f(:,1)=ones(n1(1),1);
11         R_2=0;
12         vars=0;
13         gamma=0;
14         elima=0;
15         beta=mean(Y);
16     else
17         Xsel_f=Xa;
18         vars=1;
19         gamma=0;
20         elima=0;
21         Xa=[ones(length(Y),1) Xa];
22         beta=(Xa'*Xa)^-1*Xa'*Y;
23         n2=size(Y);
24         SST=sum(Y.^2)-n2(1)*mean(Y).^2;
25         Ye=Xa*beta;
26         E=Y-Ye;
27         SSE=sum(E.^2);
28         R_2=1-SSE/SST;
29         beta=beta';
30     end
31 else
32     [Xsel,beta,gamma,elima]=multicol(Y,Xa, restr);
33     %%%%%%%%%%%
34     %%%%%%%%%%%
35     while gamma>restr
36         n=size(Xsel);
37         Xn=Xsel(:,2:n(2));
38         [Xsel,beta,gamma,elima]=multicol(Y,Xn, restr);
39     end
40     %%%%%%%%%%%
41     %%%%%%%%%%%
42     n2=size(Y);
43     SST=sum(Y.^2)-n2(1)*mean(Y).^2;
44     Ye=Xsel*beta;
45     E=Y-Ye;
46     SSE=sum(E.^2);
47     R_2=1-SSE/SST;
48     %%%%%%%%%%%
49     %%%%%%%%%%%
50     n=size(Xsel);
51     Xsel=Xsel(:,2:n(2));
52     for j=1:n(2)-1

```

```

53         Xsel_f(:,j)=Xsel(:,j);
54         vars(j)=eliminate_2(Xa,Xsel(:,j));
55     end
56     beta=beta';
57 end

```

A.2.6. Función factor de la varianza.

```

1  %%%%%%%%%%%
2  %% Función factor de la varianza %%
3  %% Algoritmo diseñado por Joan Castro   %%
4  %%%%%%%%%%%
5
6  function [X,vars,beta,R_2,VIF,p_val,t_0]=variance_factor(Y,X)
7
8  n=size(X);
9  vars=1:n(2);
10 xm=mean(X);
11 d=X-repmat(xm,n(1),1);
12 sjj=(d'*d);
13 sjj=sqrt(diag(sjj));
14 %X
15 %%%%%%%%%%%
16 W=(X-repmat(xm,n(1),1))./repmat(sjj',n(1),1);
17 C=W*W;
18 VIF=abs(diag(inv(C)));
19 %%%%%%%%%%%
20 while max(VIF)>5
21     [vifval,vifpos]=max(VIF);
22     detection=hist(vifpos,1:size(VIF));
23     winners=find(detection==0);
24     vars=vars(winners);
25     length(winners)
26     X=X(:,winners);
27     xm=mean(X);
28     size(X);
29     size(xm);
30     d=X-repmat(xm,n(1),1);
31     sjj=(d'*d);
32     sjj=sqrt(diag(sjj));
33 %%%%%%%%%%%
34 W=(X-repmat(xm,n(1),1))./repmat(sjj',n(1),1);
35 C=W*W;
36 VIF=abs(diag(inv(C)));
37 end
38 %%%%%%%%%%%
39 Xs=[ones(n(1),1) X];
40 beta=(Xs'*Xs)^-1*Xs'*Y;
41 n2=size(Y);
42 SST=sum(Y.^2)-n2(1)*mean(Y).^2;
43 Ye=Xs*beta;
44 E=Y-Ye;
45 SSE=sum(E.^2);
46 R_2=1-SSE/SST;
47 beta=beta';
48 C=(Xs'*Xs)^-1;
49 for r=1:size(Xs,2)
50     desv_b(r)=sqrt(SSE/(n(1)-length(beta))*C(r,r));
51     t_0(r)=abs(beta(r))./desv_b(r);
52     p_val(r)=2*(1-tcdf(t_0(r),n(1)-length(beta)));
53 end

```

A.2.7. Función que selecciona las mejores variables para el modelo.

```

1  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
2  %% Función que selecciona las mejores variables para el modelo %%
3  %% Algoritmo diseñado por Joan Castro   %%
4  %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
5
6  function [vars,Xs,beta,R_2,p_val,t0]=select_best(method,Xsa,Y,alfa)
7
8  nsa=size(Xsa);
9
10 if nsa==[0 0]
11     Xs=ones(nsa(1),1);
12     vars=0;
13     R_2=0;
14     beta=(Xs'*Xs)^-1*Xs'*Y;
15     p_val=Inf;
16     t0=0;
17 else
18     if method==1
19         [R_2,Xs,beta,Ye,p_val,Fu,vars,t0]=forward_sel(Y,Xsa,alfa);
20     elseif method==2
21         [R_2,Xs,beta,Ye,p_val,Fu,vars,t0]=backward_sel(Xsa,Y,alfa);
22     elseif method==3
23         [vars,R_2,Xs,beta,p_val,t0]=stepwise_method(Xsa,Y,alfa);
24     end
25
26     if vars==0
27         Xs=ones(nsa(1),1);
28         R_2=0;
29         beta=(Xs'*Xs)^-1*Xs'*Y;
30         p_val=Inf;
31         t0=0;
32     end
33 end

```

A.3. Eventos de lluvias y tormentas

Tabla A-1: Muestra de los eventos de lluvias y tormentas usadas en el trabajo.

Eventos totales				
No.	Fecha del evento	Duración del evento (horas)	Tipo de evento	Impacto humano o económico
1	Abril 17-18, 2003	48 horas	Onda tropical	Inundaciones repentinas grave se concentra a través de varios municipios Las excesivas lluvias obligaron a varios cientos de familias a salir de sus hogares; además, causó un estimado de 15 millones de dolares en daños a la infraestructura y la agricultura de EE.UU..
2	Noviembre 11-18, 2003	192 horas	Onda tropical activa de nivel superior	Fuertes lluvias sobre Puerto Rico y las Islas Vírgenes de los EE.UU., durante la semana del 10 al 15 noviembre 2003; estas dieron lugar a inundaciones repentinas generalizada y numerosos deslizamientos de lodo , tierra y roca. En Puerto Rico, 19 municipios fueron declarados una zona de desastre federal, con daños totales estimados en alrededor 45 millones de dolares.
3	Diciembre 5-8, 2003	96 horas	Efectos colaterales causados por latormenta tropical Odette	Odette fue una tormenta tropical atípica (diciembre), que tocó tierra en la República Dominicana y fue responsable de ocho muertes. Además, fue la primera tormenta tropical de diciembre en el registro que se forman en el mar Caribe.
4	Abril 19-24, 2005	144 horas	Eventos aislados y lluvias locales durante la tarde	
5	Mayo 17, 2005	24 horas	Precipitaciones locales	
6	Octubre 9-13, 2005	120 horas	Presión baja de nivel Superior	Grandes pérdidas en los sectores agrícolas de la isla por los daños causados a las cosechas. Se estimaron pérdidas en 20 millones de dolares.
7	Octubre 27-29, 2007	72 horas	Fuerte onda tropical (huracán Noel)	Las lluvias torrenciales de Noel produjeron daños generalizados y pérdida de vidas en la República Dominicana, Haití, Jamaica, el este de Cuba y las Bahamas. Se estima que Noel causo un total de 163 muertes, mientras que 59 siguen desaparecidas.
8	Septiembre 20-23, 2008	96 horas	Tormenta tropical Kyle	El 24 de septiembre de 2008, el gobernador Anibal Acevedo Vilá solicitó una declaración de desastre debido a las fuertes tormentas e inundaciones que comenzaron el 21 de septiembre de 2008. El Gobernador solicitó una declaración de Asistencia Individual para ocho municipios y mitigación de riesgos en todos los demas municipios de Puerto Rico; en total fueron 45 millones de dolares en asistencia.
9	Julio 18-24, 2010	168 horas	Fuerte onda tropical	La onda tropical que se muevió a través de Puerto Rico y las Islas Vírgenes de EE.UU. en la semana del 19 de julio al 23 de julio generó fuertes aguaceros y tormentas eléctricas aisladas. Como resultado, hubo inundaciones en muchos sectores de Puerto Rico y las Islas Vírgenes de EE.UU..
10	Agosto 21-24, 2011	24 horas	Huracan Irene	A través de Puerto Rico, las lluvias torrenciales causaron grandes daños a las carreteras, mientras que los vientos con fuerza de huracán, derribó árboles y postes de electricidad, dejando a más de 1 millón de viviendas sin energía eléctrica. Las mayores cantidades de precipitación cayeron sobre parte del este del territorio, con un máximo total de 22 pulg. Estimaciones preliminares indican daños estructurales, que podrían llegar a 500 millones de dolares, con pérdidas adicionales debido a las suspensiones de tres días laborales fijadas en más de 60.4 de dolares.

Bibliografía

- [1] Browning K.A. *Nowcasting*. Academic Press, first edition, 1982.
- [2] Browning K.A and Conway B.J. Wather forecasting by interactive analysis of radar and satellite imagery. *Philos. Trans. Roy. Soc. London*, 1579(324):299–315, 1987.
- [3] Li P.W. and Lai S.T. Short range queantitative precipitation forecasting in Hong Kong. *Journal of Hydrology*, 288(3):189–209, 2004.
- [4] Germann U. and Zawadzki I. Scale-dependence of the predictability of precipitation from continental radar images. part ii: Probability forecasts. *Atmospheric Sciences Division, Meteorology and Oceanography Group, Space Applications Centre*, 2004.
- [5] Germann U. and Zawadzki I. Scale-dependence of the predictability of precipitation from continental radar images. part i: Description of the methodology. *Atmospheric Sciences Division, Meteorology and Oceanography Group, Space Applications Centre*, 2002.
- [6] Wilson J.W.; Crook N.A.; Mueller C.K.; Sun J.; Dixon M. Nowcasting thunders-torms: A status report. b. *American Society of Meteorology*, 79(99):2079–2086, 1998.
- [7] Engle R.F. Autorregresive conditional heteroskedasticity with estimates of united kingdom inflation. *Econometrica*, 50:987–1008, 1982.
- [8] Engle R.F and Granger C.W. Cointegration and error correction : Representa-tion, estimation and testing. *Econometrica*, 55:251–276, 1987.
- [9] Box G. and Jenkins G. *Time Series Analysis Forecasting and Control*. Holden-Day, third edition, 1976.

- [10] Box G. and Jenkins G. *Time Series Analysis Forecasting and Control*. Prentice Hall, third edition, 1994.
- [11] Ramírez-Beltran N.D. y Sastri T. Transient detection with an application to a chemical process. *Computers Industrial Engineering*, 32(4):891–908, 1997.
- [12] Ramírez-Beltran N.D.; Castro J.M.; Harmsen E.; Vasquez R. Stochastic transfer function models and neural networks to estimate soil moisture. *J. of the Amer. Water Resour. Assoc.*, 44(4):847–865, 2008.
- [13] Pandit S.M. and Wu S.M. *Time Series and System Analysis Applications*. John Wiley and Sons, first edition, 1983.
- [14] Wackernagel H. *Multivariate Geostatistics, An Introduction with Applications*. Springer, third edition, 1994.
- [15] González A.; Martínez F.J.; Pernía A.V.; Alba F.; Castejón M. ; Ordieres J. and Vergara E.G. *Técnicas y Algoritmos Básicos de Visión Artificial*. Univesidad de la Rioja; Servicio de publicaciones, first edition, 2006.
- [16] Doviak R.J. and Zrnić D.S. *Doppler Radar and Weather Observations*. Academic Press, second edition, 1993.
- [17] Office of the Federal Coordinator for Meteorological Services and Supporting Research. *Doppler Radar Meteorological Observations*. Federal Meteorological, handbook no. 11 edition, 2005.
- [18] Biggerstaff M. and Houze R. Kinematics and microphysics of the transition zone of the 10-11 june 1985 squall line. *J. Atmos. Sci.*, 50:3091–3110, 1993.
- [19] Brockwell P. and Davis R. *Introduction to Time Series and Forecasting*. Springer-Verlag, second edition, 2002.
- [20] León J.C. *Simulación de la distorsión atmosférica en un frente de onda electromagnética*. Biblioteca UPRM, tesis de maestría, 2005.
- [21] John Leese; Charles Novak and Bruce Clark. An automated technique for obtaining cloud motion from geosynchronous satellite data using cross correlation.

- IBM Federal Systems Division. Gaithersburg, Md.*, 10:118–132, 1970.
- [22] Kishtawal C.M.; Deb S.K.; Pal P.K. and Joshi P.C. Estimation of cloud motion winds from kalpana vhr. *Atmospheric Sciences Division, Meteorology and Oceanography Group, Space Applications Centre*, 2007.
- [23] Montgomery D.C.; Peck El. A. and Vining G.G. *Introduction to Linear Regression Analysis*. John Wiley and Sons, fourth edition, 2006.
- [24] Wilks D.S. *Statistical Methods in the Atmospheric Sciences*. Academic Press, first edition, 1995.