Algoritmos de Estimación e Interpolación de Parámetros Geofísicos

Por

Jones Harold Cruzado Japan

Tesis Sometida Para la Obtención del Grado de Maestro
En
Ciencias de Ingeniería de Computadoras

Universidad de Puerto Rico
Recinto de Mayagüez

20004

Aprobada por:

***Ramón Vásquez E*. PhD** Presidente del comité	**Fecha**
***Jorge Gonzalez,* PhD** Miembro del comité	**Fecha**
***Wilson Rivera G.* PhD** Miembro del comité	**Fecha**
***Nazario Ramírez,* PhD** Miembro del comité	**Fecha**
***Jorge L. Ortiz,* PhD** Director del departamento	**Fecha**
Representante de Estudios Graduados ***Eric Harmsen,* PhD**	**Fecha**

Resumen

Este trabajo de tesis está encaminado a encontrar modelos para estimar e interpolar espacialmente algunos parámetros de la superficie de suelo, específicamente para la región de Puerto Rico, tales como temperatura, índice de vegetación, y humedad de suelo, siendo este último el de mayor interés en este trabajo. Se proponen dos modelos empíricos basados en Funciones de Transferencia para la estimación de humedad de suelo y se construyen los correspondientes algoritmos de interpolación espacial de estos parámetros usando como herramienta de programación los software MATLAB, C/C++ y FORTRAN principalmente. Los resultados muestran una aproximación al 86% usando datos observados de Raingauge , por otro lado una aproximación al 40% usando datos observados de Radar lo cual es bueno si se toma en consideración la resolución espacial del Radar (aprox. 5*Km*).

Abstract

This thesis work is directed to find models and algorithms to estimate and interpolate some parameters of the land surface, especially for the Puerto Rico region. The studied parameters are: land surface temperature (LST), Vegetation Index (NDVI), and Soil Moisture, the last one being of the greater interest in this work for which two models based on Transfer Functions are proposed. The algorithms of spatial interpolation of these parameters are implemented using the software's C/C++, FORTRAN and MATLAB. The results show a good approach of approximately 86% using Raingauge data, on the other hand an approach to 40% was obtained using Radar data which are good if the spatial resolution of the Radar is considered (aprox. 5*Km*).

Contenido

1 Introducción 14

2 Revisión Bibliográfica 18

2.1 Modelo de Constante Dieléctrica 18

2.2 Modelos Basados en Redes Neurales Artificiales 22

2.3 Modelo de Balance de Agua 25

3 Teoría y Obtención del Modelo 28

3.1 Sistemas Dinámicos 28

3.1.1 Sistema Dinámico Continuo 28

3.1.2 Sistema Dinámico Discreto 32

3.2 Funciones de Transferencia 34

3.2.1 Función de Transferencia en Paralelo 35

3.2.2 Función de Trasferencia en Serie 35

3.2.3 Función de Transferencia con Realimentación 36

3.3 Modelos de Función de Transferencia 37

3.3.1 Modelos de Intervención 42

3.3.2 Modelo de Función de Transferencia de Múltiples Entradas 44

3.3.3 Modelo de Función de Transferencia Autoregresivo 45

3.4 Algoritmos de Optimización 48

3.4.1 Simplex Serach 48

3.4.2 Algoritmos Genéticos 50

3.5 Métodos de Interpolación 52

3.5.1 Método de Kriging 53

3.5.2 Método de Vecinos Más Cercanos 55

3.6 Redes Neurales Artificiales 57

3.6.1 Elementos de Una Neurona 58
3.6.2 Tipos de Funciones de Activación 59
3.6.3 Algoritmo "Back Propagation" 60

3.7 Regla de Kohonen 63

3.8 Algoritmos de Estimación de Temperatura de Superficie e Índice de Vegetación 65

3.8.1 Selección de Píxeles, Máscara 66
3.8.2 Temperatura de Superficie (LST) 67
3.8.3 Índice de Vegetación (NDVI) 68

3.9 NEXRAD (Next generation Radar) 70

4 Materiales y Procedimiento 74

4.1 El Laboratorio de Climatología 75

4.2 El Laboratorio PaSCoR 75

4.3 Fuentes de Datos 75

4.3.1 DAAC 76
4.3.2 SIL 76
4.3.3 WCC 77
4.3.4 USGS 78
4.3.5 NCDC 79

4.4 Otras Fuentes de Datos 80

4.5 Regiones Climatológicas 82

5 Interpolación Espacial de Parámetros Geofísicos 84

5.1 Interpolación de NDVI 84

5.2 Interpolación de Temperatura Usando Kriging 87

5.3 Interpolación de Temperatura Usando RNA 88

5.4 Interpolación de Observaciones de Radar 90

5.5 Interpolación de Observaciones de Raingauge 92

5.6 Interpolación Espacial Humedad de Suelo..... 93

6 Resutados 102

7 Conclusiones 105

8 Trabajo Futuro..... 107

9 Bibliografía 108

10 Apéndice..... 113

A Métodos e Instrumentación de Medición de Humedad de Suelo 113

A.1 Métodos..... 113

A.1.1 Método Gravimétrico..... 113

A.1.2 Método de Dispersión de Neutrones..... 113

A.1.3 Método de Atenuación de Rayos Gamma 114

A.1.4 Método de Conductividad Eléctrica 114

A.1.5 Probetas de Conductividad Eléctrica 114

A.2 Bloques de Resistencia Eléctrica 115

A.3 Inducción Electromagnética 115

A.4 Método de Tensiómetro..... 115

A.5 Método Hygrometric..... 116

A.5.1 Método Dieléctrico 116

A.5.2 Probetas de Capacitancia..... 117

A.6 Probetas TDR..... 117

B Trabajo de Campo 118

B.1 Instrumentación *118*

B.1.1 Medidor de Temperatura *118*

B.1.2 Medidor de Humedad de Suelo (ECHO) *119*

B.1.3 Medidor de Lluvia (ECHORain) *119*

B.1.4 Data Logger (EM5) *120*

C Percepción Remota *122*

C.1 TERRA/MODIS *123*

C.2 NOAA/AVHRR *124*

C.3 Radares *126*

C.4 RADARSAT/SAR *126*

Lista de Figuras

Fig. 2.1-1 Imágenes de ESTAR (iz.), Modelo de Constante Dieléctrica (der.) 19
Fig. 2.2-1 RNA, *Atluri* 23
Fig. 2.3-1 Balance de agua 25
Fig. 3.1-1 Sistema dinámico 28
Fig. 3.1-2 Sistema de dos tanques 29
Fig. 3.2-1 sistema general (izq.), sistema más concreto (der.) 34
Fig. 3.2-2 Función de transferencia en paralelo 35
Fig. 3.2-3 Función de transferencia en serie 36
Fig. 3.2-4 funciones de transferencia con realimentación 36
Fig. 3.3-1 Humedad de suelo 37
Fig. 3.3-2 Lluvia acumulada 38
Fig. 3.3-3 Temperatura de aire 38
Fig. 3.3-4 Humedad de suelo (Maricao) 39
Fig. 3.3-5 Lluvia acumulada (Maricao) 39
Fig. 3.3-6 Temperatura de aire (Maricao) 39
Fig. 3.3-7 Humedad de suelo (Isabela) 40
Fig. 3.3-8 Lluvia acumulada (Isabela) 40
Fig. 3.3-9 Temperatura de aire (Isabela) 40
Fig. 3.3-10 Algoritmo Optimización de la FT 44
Fig. 3.4-1 Un simplex, caso n=2 50
Fig. 3.4-2 Estructura general de algoritmos genéticos 51
Fig. 3.4-3 Pseudo código de algoritmos genéticos 52
Fig. 3.5-1 Interpolación Kriging 53
Fig. 3.5-2 Vecinos más cercanos 56
Fig. 3.6-1 RNA de L+1 capas 57
Fig. 3.6-2 Una Neurona y sus elementos 58
Fig. 3.7-1 Regla de Kohonen 64
Fig. 3.8-1 Reflectancia de la vegetación 69

Fig. 3.8-2 Ejemplo de NDVI 70
Fig. 3.9-1 NEXRAD 71
Fig. 4.3-1 SIL 76
Fig. 4.3-2 Cinco estaciones WCC 78
Fig. 4.3-3 Clases de suelos 79
Fig. 4.3-4 Localización de las estaciones NCDC 80
Fig. 4.4-1 Clases de vegetación 81
Fig. 4.4-2 Colores asignados a las clases 81
Fig. 4.5-1 Cinco regiones climatológicas 82
Fig. 5.1-1 Algoritmo de interpolación de NDVI 85
Fig. 5.1-2 Composición de 16 días de NDVI, Nov. 2003 86
Fig. 5.1-3 Composición de 16 días de NDVI, Interpolación, Nov. 2003 86
Fig. 5.2-1 Temperatura °C de superficie, 1 Nov. 2003 87
Fig. 5.2-2 Temperatura °C de superficie, 1 Nov. 2003 87
Fig. 5.3-1 Algoritmo RNA de obtención de temperatura de superficie 89
Fig. 5.3-2 Temperatura (°C) de superficie, Nov 2003 89
Fig. 5.4-1 Procesamiento de datos de radar 91
Fig. 5.4-2 Observaciones de Radar, Nov 1, 2003 91
Fig. 5.5-1 Observaciones de *Raingauge*, Nov 5 2003 92
Fig. 5.6-1 Aproximación mediante FT, Maricao, Dic 2003 94
Fig. 5.6-2 Aproximación mediante FT, Bosque Seco, Nov-Dic. 2003 94
Fig. 5.6-3 Tres niveles de humedad de suelo (moist1, moist2, moist3), y Lluvia 96
Fig. 5.6-4 Aproximación con datos de Raingauge, Dic. 2003 97
Fig. 5.6-5 Aproximación con datos de Radar, Dic. 2003 98
Fig. 5.6-6 Humead en la región asociado a *Bosque Seco* 98
Fig. 5.6-7 Aproximación con datos de Raingauge, Dic 2003 99
Fig. 5.6-8 Aproximación con datos de Radar, Dic. 2003 99
Fig. 5.6-9 Humead en la región asociado a *Maricao* 100
Fig. 5.6-10 Algoritmo de estimación e interpolación humedad de suelo 101
Fig. 5.6-1 Promedio aproximado cada 24h, Raingauge 102

Fig. 5.6-2 Promedio aproximado cada 24h, Radar 103
Fig. 5.6-3 Promedio aproximado cada 24h, Raingauge..... 103
Fig. 5.6-4 Promedio aproximado cada 24h, Radar 104
Fig. 5.6-1 Sensor d e Temperatura..... 118
Fig. 5.6-2 Sensor de humedad de suelo 119
Fig. 5.6-3 Raingauge, mecanismo interno 120
Fig. 5.6-4 Datalogger 120
Fig. 5.6-5 MODIS 124
Fig. 5.6-6 NOAA 125
Fig. 5.6-7 RADARSAT 126

Lista de tablas

Tabla 2.1-1 Valores de C 20
Tabla 3.9-1 Variaciones con la distancia del radar 73
Tabla 4.3-1 Localización y características de las estaciones experimentales 77
Tabla 5.6-1 Coeficientes óptimos de la FT 94
Tabla 5.6-1 Promedios mensuales, Dic. 2003 104
Tabla 5.6-1 Designación de banda de microonda 123
Tabla 5.6-2 Bandas del sensor MODIS 123
Tabla 5.6-3 Bandas del sensor AVHRR 125
Tabla 5.6-4 Características del satélite NOAA 125

Dedicatoria

Dedico este trabajo a mis padres, mis hermanos como una muestra de empeño y superación.

Agradecimientos

Agradezco al Dr. Ramon Vasquez E. Co-PI del proyecto NASA-EPSCOR y PI del Proyecto NOAA-CREST por haberme dado la oportunidad de avanzar y dar un gran paso más en mi formación profesional.

Agradezco al Dr. Jorge Gonzales, PI del proyecto NASA-EPSCOR por el apoyo economico.

Agradezco al Dr. Nazario Ramirez B. por darme la iniciativa y las bases fundamentales para dar una solución al problema del cual trata esta tesis.

Agradezco a Pieter Van Der Meer coordinador del grupo de investigación, y todo el miembro comité que formaron parte de esta tesis.

Gracias al departamento de Ingeniería Eléctrica y Computadoras por haberme dado la oportunidad de superarme.

1 Introducción

Existen muchos parámetros que influyen en el estado de la superficie de la tierra, algunos de los parámetros más conocidos son: vegetación, humedad y temperatura de superficie de suelo, lluvia acumulada, tipos de suelos, vientos, temperatura del aire, velocidad y dirección de viento, presión atmosférica entre otros, siendo los tres primeros de los mencionados que trata este trabajo, principalmente la estimación de humedad de suelo; gran parte del trabajo es estimarlos por medio de sensores en contacto con el objeto en observación para el estudio de los fenómenos y la validación de modelos, y los datos de percepción remota son usados en este trabajo para la aplicación del modelo. Estos parámetros son de mucha importancia en diferentes campos así como en climatología (en modelos de predicción), la agricultura (crecimiento de las plantas), el comercio (transporte aéreo), entre otros.

La búsqueda de modelos para la estimación de humedad de suelo condujo a proponer un modelo, y la necesidad de su prueba y aplicación condujo a trabajos de preparación de los datos de entrada. Se proponen dos modelos los cuales son llamados: "Modelo de Función de Transferencia de Múltiples Entradas", y el "Modelo de Función de Transferencia Autoregresivo". Estos modelos están basados en la construcción de una función de transferencia con variables de intervención externa (temperatura de aire, lluvia acumulada), cada uno de los modelos mencionados tienen formas similares y sus coeficientes se determinan por un mismo método, sin embargo tienen concepciones diferentes y que presentan sus ventajas una sobre la otra.

La construcción de la función de transferencia es local puesto que las series de tiempo correspondientes a la humedad están dadas en forma local (en un punto de la superficie). Una misma función es aplicada dentro de una región determinada por las clases de suelos o vegetación o las determinadas por la clasificación en regiones de similaridad climatológicas, ésta última es previamente determinada mediante clasificación teniendo como datos la lluvia acumulada, temperatura, índices de vegetación, y elevación., captados durante periodos largos de tiempo (datos por cada día o mes del año, durante 1 año o más). El método de clasificación usado en casi todo este trabajo es el no supervisado debido a que se disponen de pocas muestras experimentales para los algoritmos de clasificación supervisada. La humedad está influenciada por otros parámetros tales como: temperatura, velocidad de viento, vegetación, presión atmosférica etc. Sin embargo la intervención de todos estos parámetros hace difícil la construcción de la función de transferencia.

También se proponen algoritmos de estimación e interpolación de temperatura usando datos de percepción remota (específicamente MODIS), la interpolación es importante especialmente en la región de Puerto Rico debido a la presencia de nubes en la atmósfera que crea dificultades en las observaciones de satélite.

Un método de interpolación del Índice de Vegetación (conocido como NDVI) es también considerado, ya que esta variable presenta el mismo problema anterior (presencia de nubes) por lo que se hace necesario tener una técnica eficiente y confiable de interpolación.

Debido al interés en el uso de datos de percepción remota, se toman en consideración principal el sensor MODIS del satélite TERRA (TERA/MODIS)

describiéndose algunos algoritmos para el procesamiento y la utilización de tales datos. También se consideran las observaciones de Radar (NEXRAD) el cual llega a ser fundamental en la interpolación espacial de la humedad de suelo.

Existen en la actualidad métodos diferentes de estimación de humedad de suelo, uno de ellos se basa en la Temperatura de Brillantez, ésta temperatura es captada mediante radiómetros de la banda *L* o *C*, siendo el largo de onda de observación óptima es el de la banda L (~21 cm. de longitud de onda), y requiere de características detalladas de vegetación así como su transmitancia o su contenido de agua para estimar la perdida de de la señal a través de su paso por ella, además se requieren de las características de la superficie observada como la rugosidad, vegetación, y suelos; aunque la señal es transparente a la atmósfera (la perdida de señal a través de la atmósfera es despreciable), debido a la falta de los parámetros requeridos no es posible su aplicación en el área de Puerto Rico. Un segundo método es basado en medida de dispersión de señal de radar con longitudes de ondas de observación en la bandas *C* o *L*, este último es más factible debido a que además de ser transparente a la atmósfera no requiere de muchos parámetros de entrada como el anterior. La desventaja con estos métodos está en que los instrumentos de observación en las longitudes de ondas mencionadas no son muy comunes, hasta ahora solo con fines prueba y validación de modelos son montados en aviones y por corto tiempo, en el caso de datos de radar en estas mismas longitudes de onda son también poco comunes y por lo general son comerciales y de muy alto costo. Debido a las condiciones de la superficie del área de Puerto Rico, así como la densa y compleja vegetación,

pendientes abruptas, entre otras dificultades mencionadas hacen difícil la aplicación de estos métodos en dicha región.

2 Revisión Bibliográfica

Los siguientes modelos que se mencionan en seguida son a cerca de diversos modelos relacionados a la estimación, predicción de humedad de suelo.

2.1 Modelo de Constante Dieléctrica

Este algoritmo es basado en la inversión de la ecuación de Fresnel [18], el cual estima la constante dieléctrica del suelo conociendo la reflectancia de la superficie de suelo [20]; este método sigue los siguientes pasos:

1. Estimar la reflectancia conociendo las temperaturas de la superficie de profundidad de suelo
2. Corregir los efectos ocurridos en la trayectoria de la señal para obtener la reflectividad efectiva de la superficie:
    i. Corregir las perdidas causadas por la vegetación
    ii. Corregir las perdidas causadas por la rugosidad de la superficie
3. estimación de la constante dieléctrica efectiva mediante la inversión de la ecuación de Fresnel.
4. computar el valor de humedad volumétrica de suelo, usando el método de mezcla. Para el cual se requieren datos de la composición de suelos, así como porcentaje de barro, porcentaje de arena, porosidad, entre otros, y la constante dieléctrica.

Estos pasos son resumidos en el diagrama de la figura siguiente (Fig. 2.1-1) las imágenes que se muestran allí son obtenidas en una campaña realizada al sur de

Oklahoma, con avión de la nasa ESTAR en el año 1997 [19], este instrumento trabaja en la banda *L* del espectro.

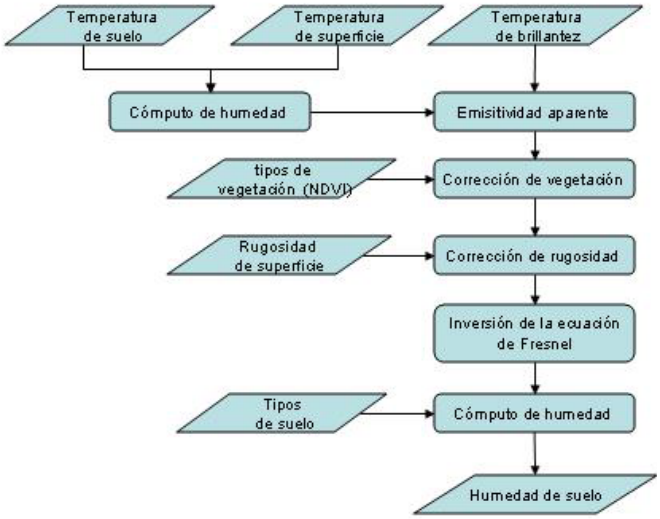

Fig. 2.1-1 Imágenes de ESTAR (iz.), Modelo de Constante Dieléctrica (der.)

A continuación se describe con más detalle las entradas del algoritmo:

a) *Temperatura de suelo* (*Soil Temperatura,* T_0), esta temperatura es obtenida mediante interpolación de datos tomadas en contacto directo con el suelo.

b) *Temperatura de superficie* (*Surface Temperatura,* T_∞), es posible obtener desde los sensores satelitales, MODIS [7], Landsat, NOAA.

c) *Temperatura efectiva* (*Efefective Temperatura,* T_{eff}), también llamada temperatura neta, esta temperatura es relativo al sensor y dado en función de la longitud de onda de observación y es dado como

$$T_{eff} = T_\infty + C(T_0 - T_\infty)$$

donde T_0 y T_∞ son temperaturas profundidad de suelo y de superficie de suelo respectivamente. La tabla siguiente muestra algunos valores de C para longitudes de ondas más usadas en percepción remota

Long. de onda (*cm*)	***C***
2.8	0.802±0.006
6.0	0.667±0.008
11.0	0.480±0.010
21.0	0.246±0.009
49.0	0.084±0.005

Tabla 2.1-1 Valores de C

Para un sensor en la banda L, C es el correspondiente a la longitud de onda de 21cm (1.4GHz), también puede ser usada la tabla para longitudes de ondas en la banda C ~6 *cm.* (6.6GHz). Por otro lado para aplicaciones prácticas las temperaturas de profundidad de suelo y de superficie puede considerarse la diferencia aproximada entre estas, ~2°C.

d) *Emisitividad aparente* ($emisitividad = 1 - reflectividad$), L emisitividad aparente es estimado en función de la temperatura de brillantez: $e = \frac{T_b}{T_s}$, donde T_b es la temperatura de brillantez, y T_s es la temperatura de superficie suave[20].

e) *Tipos de Vegetación* (*Vegetation Types*), los tipos de vegetación pueden ser caracterizados por el índice de vegetación NDVI (Normalized Difference Vegetation Index) [8].

f) *Transmitancia de la vegetación* (*Transmissivity of vegetation, γ*), debido a la presencia de vegetación compleja (gras, entre otras mezclas de cúpulas de vegetación) una tabla de valores de transmitancia no puede ser usado aquí, una formulación dependiente del índice de vegetación es construida,

$$\gamma = 0.7049 - 0.6141\,(NDVI)$$

La corrección es entonces echa por $R_v = R/\gamma^2$, donde R_v es la reflectancia efectiva

g) *Rugosidad de superficie, corrección de rugosidad*, para corregir las pérdidas causadas por la rugosidad de la superficie [24], el parámetro h es obtenida mediante el parámetro estadístico σ (desviación estándar de la rugosidad),

$$R_s = R_r \exp(h \cos\theta)$$
$$h = 4\sigma^2 \left(\frac{2\pi}{\lambda}\right)^2$$

donde θ es el ángulo de vista del sensor, λ la longitud de onda de observación, R_r la reflectancia de la superficie rugosa, y R_s sería la reflectancia de la superficie suave.

h) *Tipo de suelo, y cómputo de humedad*, el tipo de suelo es requerido por el algoritmo, específicamente sus componentes en porcentaje de arena, de barro, y la densidad bruta, el cómputo de la humedad es hecho usando el modelo de mezcla [21].

Las dificultades encontradas con respecto a este modelo fueron: El número de parámetros a conocer, la poca probabilidad de detección en regiones de densa vegetación, y otro es el costo de su aplicación debido a que los sensores de la banda *L* no son de libre acceso o tienen un alto costo en la actualidad.

Cuando el interés es monitorear grandes áreas, es necesario el uso de sensores remotos. En la banda *L* además de ser óptimo (mayor correlación con la humedad), las pérdidas de señal en la atmósfera es casi despreciable ~4Kelvin.

Otras alternativas de estimar la humedad existen, usando radares satelitales [22]. El problema con ello es mayor aún, el costo es elevado, ya que estos en su mayoría son comerciales.

2.2 Modelos Basados en Redes Neurales Artificiales

Otra alternativa de obtener estimaciones de humedad de suelo es usando redes neurales (RNA o simplemente red, ver seccion 3.6), existen diversos métodos de obtener estimaciones de humedad de suelo usando una RNA, y ellos se distinguen por la consideración del número de entradas (variables) de la red, se mensiona qui una de ellas.

Atluri et. al. [14] estima la humedad de suelo usando redes neurales a diferentes profundidades, y stima también la temperatura, además hace la comparación entre dos tipos de entrenamiento de redes neurales: backpropagation (BP) y Levenberg-Marquardt (LM), mostrando ser más rápido el segundo (LM).

Los datos de entrada usados son obtenidos de las superficies con gras y superficies suaves sin vegetación:

- Humedad de suelo (*sm*), temperatura de suelo, a diferentes profundidades (5, 10, 15, 30 ,60 y 90 *cm*), obtenidas mediante sensores portables.
- Radiación *infrarrojo cercano* (temperatura de brillantez) en las bandas *L* y *S* (TB-L, TB-S).

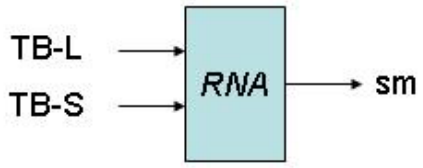

Fig. 2.2-1 RNA, *Atluri*

Jiang H. et. al. [13] considera una red neural llamado Counter Propagation Network (CPN), lo cual es una red basado en el modelo SOFM (Self Organization Feature Map) de entrenamiento no supervisado, y considera datos obtenidas desde satélite (de la Nacional Oceanic and Atmospheric Administration NOAA). Las variables de entrada son temperatura de brillantez (T_s), índice de vegetación (NDVI: Normalized Difference of Vegetation Index), la lluvia acumulada mensual, temperatura de aire a 2 *m* de altura por encima de la superficie (T_a), flujo radiactivo neto (R_{net}) y la humedad de suelo obtenidas de LDAS (Land Data Assimilation System, ver http://ldas.gsfc.nasa.gov/).

La red neural es entrenada considerando los siguientes parámetros:

$$N = \frac{NDVI - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}}$$

$$T_1 = \frac{T_s - T_a}{R_{net}}$$

$$T_2 = \frac{T_s}{R_{net}}$$

$$T_3 = T_s$$

$$T_4 = 2\frac{T_{s\max} - T_{s\min}}{T_{s\max} + T_{s\min}}$$

donde T_{smin} y T_{smax} son los correspondientes mínimo y máximo de T_s, lo mismo con NDVI, para seleccionar las mejores entradas, esto es, las variables de más alta

correlación con la humedad de suelo, en la siguiente lista se muestran las variables de entrada consideradas para la red, en cada uno de los 6 casos de prueba (por ejemplo el primer caso considera dos entradas N y T_1):

1. N T_1
2. N T_2
3. N T_3
4. N T_4
5. N T_4 DRR
6. N T_4 ARR

donde DRR es el promedio de la lluvia diaria y ARR es la lluvia acumulada de los 30 días previos.

En la tabla siguiente se muestra los valores de correlación entre las distintas variables y la humedad:

Entrada	Correlación
$T1$	0.0155
$T2$	-0.2374
$T3$	-0.1079
$T4$	-0.5449
N	0.4639
DRR	0.3068
ARR	0.6706

donde se observa que la lluvia acumulada está más altamente correlacionado con la humedad de suelo.

Esta experiencia es un buen argumento para la consideración de la lluvia acumulada en las estimaciones de humedad de suelo en otros modelos y particularmente en este trabajo de tesis.

2.3 Modelo de Balance de Agua

Este modelo es basado en el balance de agua del suelo [16]. Las componentes de balance de agua en el modelo son: precipitación (P), evaporación (E), runoff (R), y escurrimiento (G).

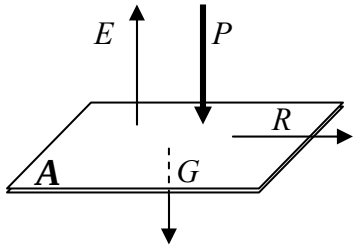

Fig. 2.3-1 Balance de agua

La dinámica del modelo de humedad de suelo es expresada como:

Aa

$$\frac{dw}{dt} = P(t) - E(t) - R(t) - G(t) \qquad \textbf{Ec. 2.3-1}$$

donde

$W(r)$ es el contenido de agua del suelo en el instante t

$P(t)$ es el promedio de la precipitación sobre el área A

E(t) es la media de la evapotranspiración sobre el área A

R(t) es la divergencia del flujo de agua (runoff)

G(t) es el escurrimiento hacia las profundidades del área A

La divergencia R consiste de una componente de rugosidad de superficie $S(t)$ y una componente de rugosidad de profundidad $B(t)$

$$R(t) = S(t) + B(t)$$

$$S(t) = P(t)\left(\frac{W(t)}{W_{\max}}\right)^m$$

$$B(t) = \frac{\alpha}{1+\mu} w(t)$$

donde $W_{\max}$ (mm) es una medida de la capacidad del suelo para retener el agua, $m > 1$, α y μ son constantes determinadas para cada tipo de suelo.

El flujo de pérdida a través de las profundidades (G) es estimado mediante

$$G(t) = \frac{\mu\alpha}{1+\mu} w(t)$$

La evapotranspiración $E(t)$ es estimada en este modelo mediante:

$$E(t) = E_p \frac{W}{W_{\max}}$$

donde E_p es la razón de evapotranspiración en mm por mes, dado de manera simplificada [16] como:

$$E_p = \begin{cases} 0 & , T < 0^\circ C \\ 16L\left(\frac{10T}{I}\right)^a & , 0 \le T < 26.5^\circ C \ldots (*) \\ -415.85 + 32.25T - 0.34T^2 & , T \ge 26.5^\circ C \end{cases}$$

donde

$$L = \frac{d}{30}\frac{h}{12}$$

$$I = \sum_{M=1}^{12} t\,, \quad t = \left(\frac{T_M}{5}\right)^{1.514}, \ \min(t) = 0\,, \text{ y}$$

$$a = 6.75\times10^{-7} I^3 - 1.71\times10^{-5} I^2 + 1.79\times10^{-2} I + 0.49$$

donde d es el número de días del mes, h es el número de horas de día, y T es la temperatura media mensual del aire.

El potencial de evapotranspiración (E_p) depende principalmente de la radiación neta por calentamiento de la superficie, sin embargo su estimación no es adecuado para grandes áreas, el trabajo en [16] usa un método de estimación de E_p basado sólo en temperatura (*), de modo que las componentes del sistema solo usan informaciones de temperatura y lluvia. Los parámetros fijos en el sistema son estimados mediante el método de *Simplex Search* [30].

3 Teoría y Obtención del Modelo

3.1 Sistemas Dinámicos

Supongamos que el par (X_t, Y_t) son observaciones obtenidas en el mismo instante en el tiempo, esto es $\Delta X_t = X_t - X_{t-1}$, de un sistema dinámico [33]-[34] con entrada X y salida Y como se muestra en la figura abajo (Fig. 3.1-1). Muchos sistemas pueden producir cambios continuos pero las observaciones son obtenidas en forma discreta como series de tiempo [1]-[4], además un modelo discreto puede explicar el comportamiento de un proceso continuo, trataremos más de cerca la construcción del modelo discreto.

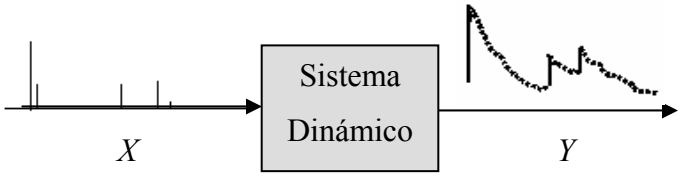

Fig. 3.1-1 Sistema dinámico

El sistema en mención podría ser el comportamiento de un proceso en la industria (sistemas de control, procesos de separación o filtración, aceleración de reacción) en la economía de una nación, etc.

3.1.1 Sistema Dinámico Continuo

Para obtener un sistema dinámico [1] de primer orden, considérese el tanque A de distribución de agua como se ilustra en la figura siguiente (Fig. 3.1-2)

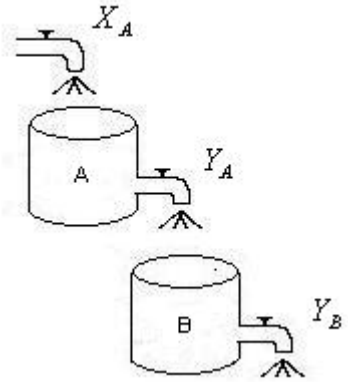

Fig. 3.1-2 Sistema de dos tanques

Este sistema consta de dos taques *A* y *B*, el tanque *A* posee entrada X_A y salida Y_A y el tanque *B* posee entrada *A* Y_A y salida Y_B, las cuales poseen un mecanismo de control de flujo de agua de modo que cualesquiera de las entradas o salidas pueden ajustarse, así por ejemplo la entrada del tanque *A* (flujo X_A) puede ser ajustada para obtener un flujo $g_A X_A$, puede observarse que ajustando (variando g_A) puede lograrse que el sistema *A* se encuentre en equilibrio ($g_A X_A = Y_A$), para una g_A arbitraria puede decirse que la razón de cambio de volumen es proporcional a $(g_A X_A - Y_A)/T_A$, esto es:

$$\frac{dY_A}{dt} = \frac{g_A X_A - Y_A}{T_A} \qquad \textbf{Ec. 3.1-1}$$

donde T_A es un tiempo fijo, g_A es una constante de proporcionalidad, y puede ser rescrita como:

$$(1 + T_A D)Y_A(t) = g_A X_A(t) \qquad \textbf{Ec. 3.1-2}$$

algunas veces el sistema podría tener efectos retardados; Supongamos por ejemplo que el tanque de agua esta vacía inicialmente, entonces una vez abierta *A* tomará un tiempo, digamos τ, para alcanzar el nivel de volumen V para luego responder en *B*. Demodo que podemos escribir la ecuación anterior (3.2-2) como:

$$(1+T_A D)Y_A(t) = gX_A(t-\tau)$$

Por otro lado puede también obtenerse la representación de un sistema con tiempo adelantado, aunque este es informalmente ya que es difícil encontrar sistemas de este tipo.

De la misma manera se obtiene la ecuación para el tanque *B* de la figura anterior (Fig. 3.1-2):

$$(1+T_B D)Y_B(t) = gY_A(t) \qquad \textbf{Ec. 3.1-3}$$

Combinando las ecuaciones anteriores (3.2-2 y 3.2-3) se tiene una representación de un sistema de segundo orden:

$$(1+(T_A+T_B)D+(T_A T_B)D^2)Y_B(t) = g_A \mathrm{g}_\mathrm{B}\ X_A(t)$$

en forma equivalente:

$$(1+\Xi_1 D+\Xi_2 D^2)Y_B(t) = g\ X_A(t) \qquad \textbf{Ec. 3.1-4}$$

donde $\Xi_1 = T_A + T_B$, $\Xi_2 = T_A T_B$, $g = g_A g_B$.

Muchos sistemas pueden ser aproximados por un sistema de segundo orden con retardo.

En general para un sistema más elaborado de *n* tanques con tiempo de retardo τ, y considerando las razones de cambio en la salida se puede obtener la siguiente expresión:

$$(1+\Xi_1 D+\Xi_2 D^2+...+\Xi_n D^n)Y_B(t)$$
$$=(\Omega_0+\Omega_1 D^1+\Omega_2 D^2+...+\Omega_n D^n)X_A(t-\tau)$$ **Ec. 3.1-5**

cuya solución puede escribirse como:

$$Y_A(t)=\int_0^\infty v(s)X_A(t-s)ds$$ **Ec. 3.1-6**

donde $v(s)$ es la función de respuesta al impulso (también llamado función de Green), más explícitamente las soluciones para los sistemas de uno y dos tanques se tiene:

- Caso 1 (un tanque):

  v(s) es definido por

  $$v(s)=g_A e^{-\frac{t}{T_A}}$$

  de modo que manteniendo constante $X(t)=1$, se tiene :

  $$Y(t)=g_A(1-e^{\frac{t}{T_A}})$$ **Ec. 3.1-7**

- Caso 2 (dos tanques):

  $$v(s)=\frac{g_A}{T_A-T_B}(e^{-\frac{s}{T_A}}-e^{-\frac{s}{T_B}})$$ **Ec. 3.1-8**

  de modo que manteniendo constante $X(t)=1$, se tiene :

  $$Y(t)=g_A g_B(1-\frac{T_A e^{-\frac{t}{T_A}}-T_B e^{-\frac{t}{T_B}}}{T_A-T_B})$$ **Ec. 3.1-9**

3.1.2 Sistema Dinámico Discreto

Desde la teoría infinitesimal se tiene que la *Derivada* de una función, sea *Y*, es definida por $\lim_{h\to 0}\dfrac{Y(t)-Y(t-h)}{h}=\dfrac{dY}{dt}=Y'(t)$, y algunas propiedades importantes de esta teoría son consideradas, así como el *teorema de valor medio* (ver por ejemplo [36]):

$$Y'(\xi)=\frac{Y(t)-Y(t-h)}{h}\text{, para algún } 0<\xi<h$$

o la *serie de Taylor*:

$$Y'(t)=\sum_{i=0}^{n}\frac{Y^{(i)}(t-h)}{i!}h^{i}+R_n(t)\text{, donde } R_n(t)=\frac{Y^{(n+1)}(\theta)h^{i+1}}{(n+1)!} \text{ y } 0<\theta<h$$

los cuales relacionan los sistemas dinámicos discretos y continuos.

Para ejemplificar considérese la ecuación diferencial $Y'(t)=aY(t)+X(t)$ donde *a* es una constante, usando el teorema del valor medio se tiene:

$\dfrac{Y(t)-Y(t-h)}{h}+aY(t)=X(t)$, de donde se obtiene: $Y(t)=\varepsilon Y(t-h)+\eta X(t)$, lo cual es un sistema de diferencias, también llamado sistema discreto.

De hecho los sistemas de diferencias dan una manera de aproximar soluciones de una ecuación diferencial (continua).

Realmente estas situaciones en las que se trata con observaciones discretas son también ampliamente estudiados en temas estadísticos, en este trabajo se optará por las notaciones dadas principalmente en la estadística, se mencionarán los modelos AR (Autoregressive) ARMA (Autoregressive Moving Average) [1]- [4] . La ecuación en diferencias (con *h*=1) obtenida anteriormente tiene la forma $Y_t=\varepsilon Y_{t-1}+\eta X_t$ el cual

es un AR(1), y en función del operador de retardo ∇ puede ser escrita como $(1+\varepsilon\nabla)Y_t = \eta X_t$

Sistemas más complejos pueden ser representados por la siguiente ecuación de diferencias (sistema discreto):

$$(1+\varepsilon_1\nabla+\varepsilon_2\nabla^2+...+\varepsilon_n\nabla^n)Y_t = g(1+\eta_1\nabla+\eta_2\nabla^2+...+\eta_m\nabla^m)X_{t-b} \qquad \textbf{Ec. 3.1-10}$$

el cual será referido como un modelo de función de transferencia de orden (*n*,*m*),

cuya solución tiene la forma $y_t = \sum_{i=0}^{\infty} v_i X_{t-i}$ donde v es también llamado *función de Green*.

Ésta puede ser también escrita en términos del operador de retardo *B* definida por:

$$B = 1-\nabla$$

de modo que el sistema anterior (3.2-10) puede expresarse como:

$$(1+\delta_1 B+\delta_2 B^2+...+\delta_r B^r)Y_t = g(1-\varpi_1 B-\varpi_2 B^2-...-\varpi_s B^s)X_{t-b} \qquad \textbf{Ec. 3.1-11}$$

de manera compacta esta última ecuación puede expresarse como:

$$\delta(B)Y_t = g\varpi(B)X_{t-b} = g\varpi(B)B^b X_t = \Omega(B)X_t$$

donde $\delta(B) = 1+\delta_1 B+\delta_2 B^2+...+\delta_r B^r$, $\delta(B) = 1-w_1 B-w_2 B^2-...-w_s B^s$, y $\Omega(B) = gw(B)B^b$

3.2 Funciones de Transferencia

Existen diversas maneras para resolver una ecuación diferencial ordinaria (continua), una de ellas es usando el operador *laplaciano* [34], aplicando este operador a la ecuación de la forma:

$$\begin{aligned}&(1+\Xi_1 D+\Xi_2 D^2+\ldots+\Xi_n D^n)y(t)\\&=(\Omega_0+\Omega_1 D^1+\Omega_2 D^2+\ldots+\Omega_m D^m)x(t-\tau)\end{aligned}$$

y considerando que los valores de frontera son todas cero:

$$\begin{matrix} D^i Y = 0 & , i \le n \\ D^j X = 0 & , j \le m \end{matrix}$$, considerando $\tau = 0$ se tiene:

$$a_0 Y(s) + a_0 s Y(s) \cdots + a_n s^n Y(s) = b_0 X(s) + b_0 s X(s) \cdots + a_m s^m X(s)$$

el cual puede ser expresado como:

$$Y(s) = \frac{b_0 + b_0 s \cdots + a_m s^m}{a_0 + a_0 s \cdots + a_n s^n} X(s) \text{, ó } Y(s) = \frac{N(s)}{D(s)} X(s)$$

de manera compacta $Y(s) = G(s)X(s)$

$G(s)$ es definida como la función de transferencia del sistema. La siguiente figura será referida como diagrama de bloque.

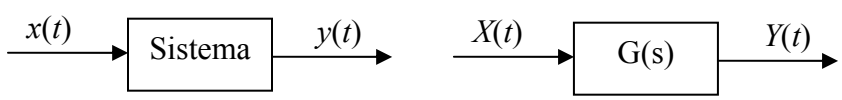

Fig. 3.2-1 sistema general (izq.), sistema más concreto (der.)

Puede verse que $G(s)$ tiene las propiedades de una transformación lineal (satisface las dos propiedades a y b como se muestra en seguida), para ello

considérese las entradas $Y_1(s)$, $Y_2(s)$ y las salidas $X_1(s)$, $X_2(s)$ y k una constante, entonces:

a. $G(s)(kX_1(s)) = kG(s)X_1(s) = kY_1(s)$

b. $G(s)(X_1(s) + X_2(s)) = G(s)X_1(s) + G(s)X_2(s) = Y_1(s) + Y_2(s)$

mediante los diagramas de bloque se puede representar mejor las funciones de transferencia y simplificarlas como es usual en sistemas eléctricos (paralelos y en serie).

3.2.1 Función de Transferencia en Paralelo

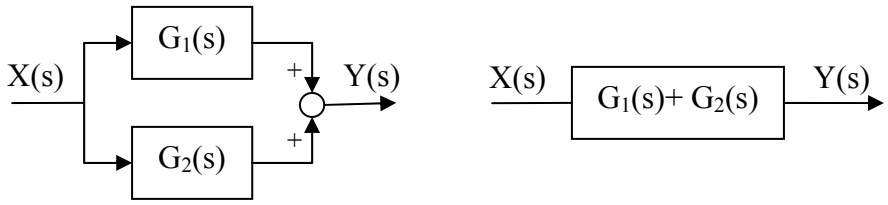

Fig. 3.2-2 Función de transferencia en paralelo

La función equivalente puede ser determinada como sigue:

$$Y(s) = G_1(s)X(s) + G_2(s)X(s) = (G_1(s) + G_2(s))X(s)$$

3.2.2 Función de Trasferencia en Serie

El diagrama siguiente muestra dos sistemas dinámicos en serie y su correspondiente sistema equivalente.

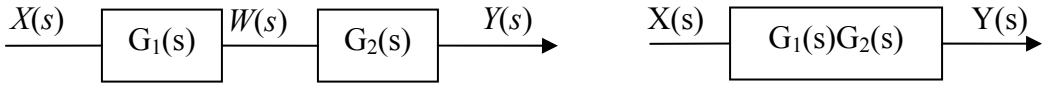

Fig. 3.2-3 Función de transferencia en serie

La función equivalente puede ser determinada como sigue:

$$Y(s) = G_2(s)W(s)$$
$$W(s) = G_1(s)X(s)$$

$$Y(s) = G_2(s)(G_1(s)X(s)) = G_1(s)G_2(s)X(s)$$

3.2.3 Función de Transferencia con Realimentación

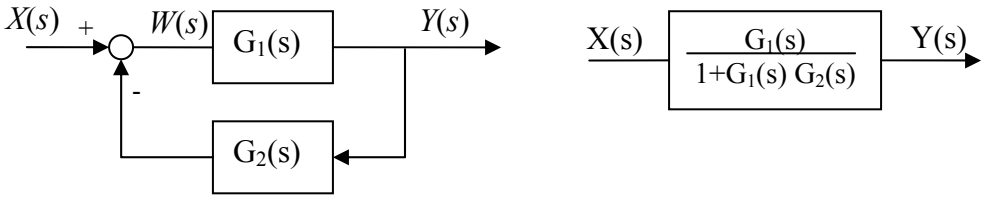

Fig. 3.2-4 funciones de transferencia con realimentación

La función equivalente puede ser determinada como sigue:

$$Y(s) = G_1(s)W(s)$$
$$W(s) = X(s) - G_2(s)Y(s)$$

de este sistema de ecuaciones se tiene:

$$Y(s) = \frac{G_1(s)}{1 + G_1(s)G_2(s)} X(s)$$

3.3 Modelos de Función de Transferencia

Los modelos que se tratarán aquí son basados en la construcción de funciones de transferencia con variables de intervención externas. Ésta será formulada en dos modelos, en su forma más general considerando las variables de intervención naturales externos, estos son lluvia acumulada, temperatura de aire principalmente (Modelo de Función de Transferencia de Múltiples Entradas), y en su forma más simple considerando solamente la variable lluvia acumulada (Modelo de Función de Transferencia Autoregresivo). Para tener una idea de las relaciones entre los parámetros en mención obsérvese las series de tiempo de los parámetros humedad de suelo, lluvia acumulada y temperatura, en un período de 2 meses (Nov. y Dic. 2003) en las siguientes figuras con datos de la estación Monte Gilarte (Tabla 4.3-1).

- Series de tiempo obtenidas de la estación Monte Gilarte

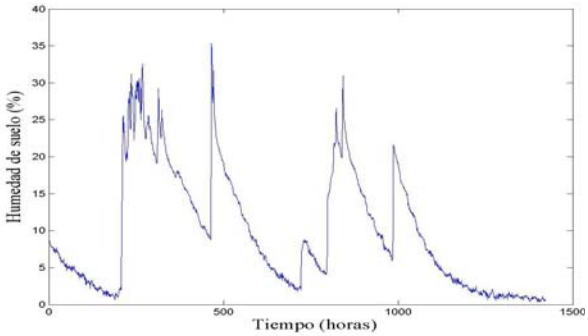

Fig. 3.3-1 Humedad de suelo

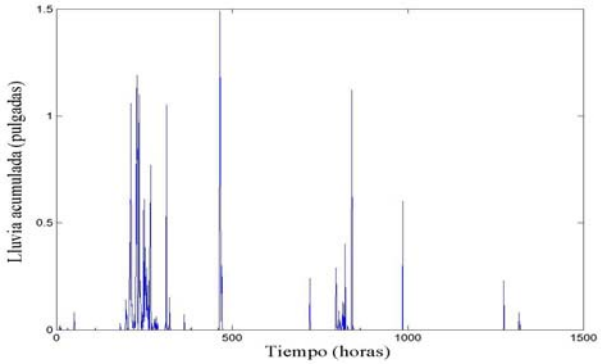

Fig. 3.3-2 Lluvia acumulada

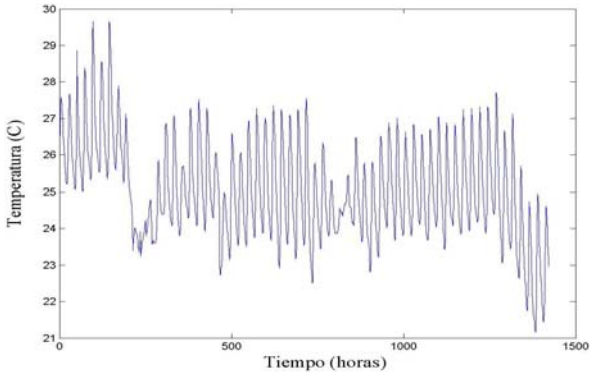

Fig. 3.3-3 Temperatura de aire

Las figuras muestran las series de tiempo para los parámetros de humedad de suelo, lluvia acumulada y temperatura (en ese orden), tomadas en intervalos de tiempo de 1 hora durante 2 meses, donde puede observarse ciertas relaciones entre estos parámetros, así como los cambios en el nivel de humedad con los eventos de lluvia y que las caídas en la humedad con los periodos de sequía.

Para mostrar que los comportamientos de los parámetros con el cambio de lugar (otra estación, otro tipo de suelo), obsérvese las siguientes series de tiempo para el mismo intervalo de tiempo anterior con datos de otras dos estaciones diferentes Maricao e Isabela.

- Series de tiempo obtenidas de la estación Maricao

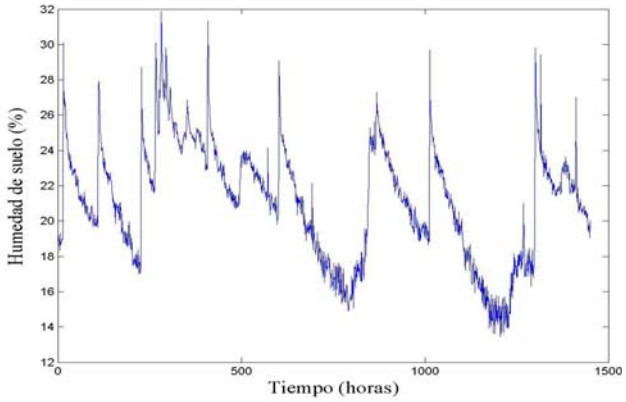

Fig. 3.3-4 Humedad de suelo (Maricao)

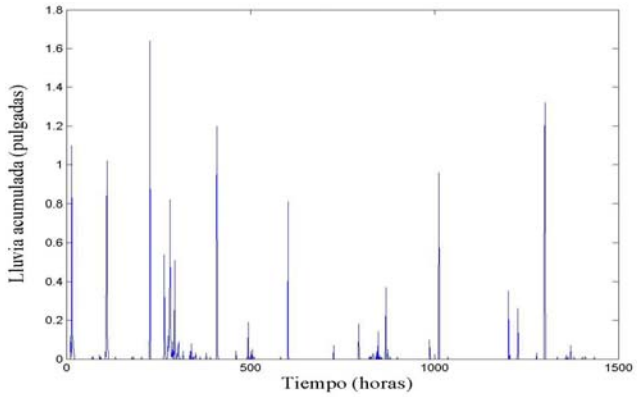

Fig. 3.3-5 Lluvia acumulada (Maricao)

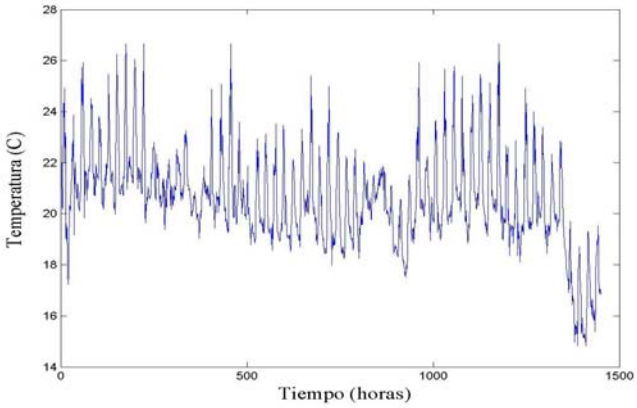

Fig. 3.3-6 Temperatura de aire (Maricao)

- Series de tiempo obtenidas de la estación Isabela

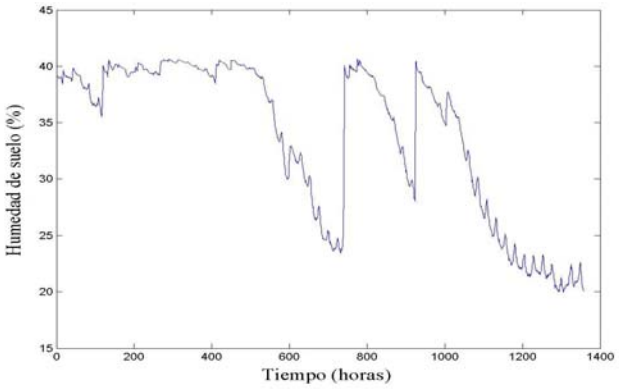

Fig. 3.3-7 Humedad de suelo (Isabela)

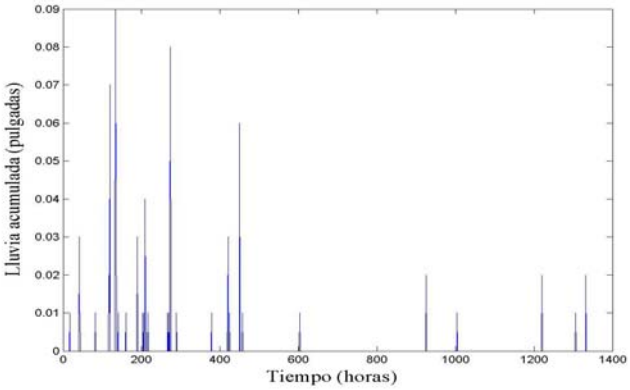

Fig. 3.3-8 Lluvia acumulada (Isabela)

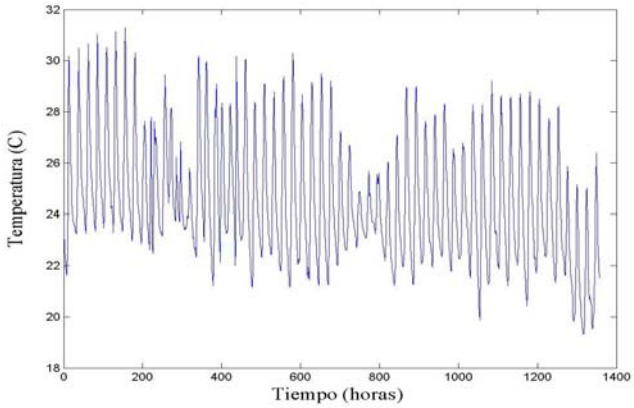

Fig. 3.3-9 Temperatura de aire (Isabela)

Puede obtenerse algunas conclusiones a partir de los datos mostrados en las figuras anteriores para las tres estaciones mencionadas:

La respuesta del nivel de humedad de suelo durante los periodos de lluvia y las pendientes de las caídas de humead con los periodos de sequía caracterizan un punto (o una región) de la superficie en observación.

Puede notarse que las series de tiempo tomadas de la estación Isabela, aunque estos datos no fueron tomados en cuenta para la prueba del modelo debido a que en el intervalo de 60-800 hay un evento de lluvia que no aparece (puede observarse que hay cambio de humedad pero en la serie de lluvia no se refleja) , revelan aún mejor la relación de la humedad con y la temperatura, aquí puede verse una componente sinusoidal en la humedad al considerarse datos con periodos de 24 horas, esto implica que para realizar un modelamiento por horas (por ejemplo 1 día o menos) mediante una función de transferencia (FT) es justo tener que considerar el parámetro de temperatura, de otra manera si se consideran el modelamiento con periodos largos de datos conviene mas bien eliminar esa "pequeña" componente periódica de la humedad y así no tener que considerar el parámetro temperatura como variable de la FT, esas componentes sinusoidales pueden ser eliminadas mediante filtros [3].

Nota. Debido a la forma de de la gráfica de la serie correspondiente a la temperatura (forma sinusoidal), ésta puede ser aproximada conociendo sus valores mínimos (*MinTemp*) y máximo (*MaxTemp*) con la función:

$$f(t) = a\,(MaxTemp - MinTem)\ Sin(2\pi t/24 + b) + c\,,\ \ t \in Z \qquad \textbf{Ec. 3.3-1}$$

encontrándose las constantes *a*, *b* y *c*, donde ***Z*** es el conjunto de números enteros, esto es posible debido a su periodicidad (de periodo 24 horas), éste echo hace que si se opta por considerar la temperatura, solo bastará tomar los mínimos y máximos de cada 24 horas de temperatura para obtener una serie tiempo aproximada por horas, puede verse a esto como una forma de interpolación temporal de la temperatura.

3.3.1 Modelos de Intervención

Una serie de tiempo esta afectado por ciertos eventos externos, por ejemplo si de negocios se trata, las ventas, las ofertas, etc., en el caso de humedad de suelo las lluvias, la temperatura, esos eventos externos son llamados *intervención* [2]. Existen dos tipos de variables de intervención comunes, la intervención *paso* o permanente

$$S_t^T = \begin{cases} 0 & ,t < T \\ 1 & ,t \geq T \end{cases}$$

donde T es el instante en que ocurre la intervención, la intervención *pulso* se puede expresar como:

$$P_t^T = \begin{cases} 0 & ,t \neq T \\ 1 & ,t = T \end{cases}$$

Además, se deduce fácilmente la relación entre éstas dos últimas funciones: $P_t^T = S_t^T - S_{t-1}^T$, y en términos del operador de retrazo se tiene:

$$P_t^T = S_t^T - BS_t^T = (1-B)S_t^T$$

Se pueden expresar funciones de respuestas a las intervenciones de paso o impulso, algunas formas conocidas son:

1. Un impacto de una intervención con efecto retardado *b* puede tener la forma siguiente:

   $\alpha B^b S_t^T$, o $\alpha B^b P_t^T$

2. Un impacto de una intervención con efecto retardado y respuesta gradual puede tener la forma siguiente:

   $\frac{\alpha B^b}{1-\delta B} S_t^T$, o $\frac{\alpha B^b}{1-\delta B} P_t^T$, con $|\delta| < 1$.

En forma más general varias respuestas pueden ser producidas por combinaciones de entradas de *pasos* y *pulsos* como el siguiente:

$$\frac{\alpha_0 B^b}{1-\delta B} S_t^T + \alpha_1 B S_t^T \qquad \textbf{Ec. 3.3-2}$$

y una respuesta puede ser representada por la fracción polinomial:

$$\frac{\alpha(B) B^b}{\delta(B)} \qquad \textbf{Ec. 3.3-3}$$

donde $\alpha(B) = \alpha_0 + \alpha_1 B + ... + \alpha_s B^s$ y $\delta(B) = 1 + \delta_1 B + ... + \delta_r B^r$ son polinomios en B. *b* es el tiempo de retardo para el efecto de intervención. Si las raíces del polinomio $\delta(B)$ están dentro del disco unitario ($|B| < 1$) entonces el sistema es estable y casual, si se encuentran fuera del disco unitario ($|B| > 1$) el proceso no es casual y representan un fenómeno de respuesta gradual y no es útil para representar cualquier fenómeno real, y si las raíces están sobre del círculo unitario ($|B| = 1$) el proceso no es estacionario y representan un impacto creciente linealmente.

3.3.2 Modelo de Función de Transferencia de Múltiples Entradas

Con las observaciones hechas puede hacerse la siguiente formulación

$$H_t = \frac{\alpha_{11} + \alpha_{12} B}{1 - \delta_1 B} R_t + \frac{\alpha_{21} + \alpha_{22} B}{1 - \delta_2 B} T_t \qquad \textbf{Ec. 3.3-4}$$

donde H_t es la estimación de la humedad, R_t datos de lluvia, T_t datos de temperatura en el instante *t*, las fracciones de los sumandos pueden ser aproximadas de la siguiente manera:

$$V_i(B) = (\frac{\alpha_{i1} - \alpha_{i2} B}{1 - \delta_i B}) \cong v_{0,i} + v_{1,i} B + ... + v_{n,i} B^n \text{ , } i = 1,2 \qquad \textbf{Ec. 3.3-5}$$

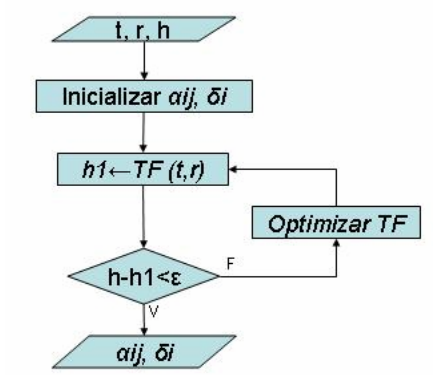

Fig. 3.3-10 Algoritmo Optimización de la FT

La figura que se muestra arriba (Fig. 3.3-10) muestra el diagrama de flujo para la optimización de la función de transferencia; esta optimización consiste en encontrar los mejores parámetros (aquellos que minimizan la función) α_{ij} y δ_i de la función H_t, para ello se consideró el método llamado *Simplex Search*, el cual es aplicado también para minimizar una función de dominio discreto.

3.3.3 Modelo de Función de Transferencia Autoregresivo

De acuerdo a las observaciones echas al principio puede verse que la serie de humedad de suelo puede expresarse en función de una componente periódica f :

$$H = \overline{H} + f(t) \qquad \textbf{Ec. 3.3-6}$$

donde $\overline{H}$ es la serie de humead sin la componente periódica, f es la componente periódica. Aquí se prefiere tener que modelar $\overline{H}$ (en adelante se escribirá simplemente H en lugar de $\overline{H}$) para evitar cargar toda la responsabilidad de periodicidad a los coeficientes de las variables de intervención.

Por otra parte uno de los problemas con los modelos de estimación de humedad de suelo cuando se llevan a la práctica es el número de parámetros de entrada que se requieren, el modelo que propone esta tesis sólo usa unametro de entrada (datos de lluvia). El modelo es ideado en base a pruebas y las experiencias obtenidas de los modelos anteriores mencionados en la sección 2. Una de las bases fue el Modelo de Balance de Agua, el cual puede ser expresado en forma aproximada como:

$$\frac{dw}{dt} = aP(t) - \lambda w(t) \qquad \textbf{Ec. 3.3-7}$$

el cual también puede ser vista como una ecuación de diferencias (ver sección 3.1.2), tomando la forma:

$$w_t = \alpha w_{t-1} + \beta P_t$$

lo cual tiene la forma autoregresiva, que finalmente usando el operador de retardo B puede ser expresada como:

$$w_t = \frac{\alpha_1(B)}{\delta_1(B)} P_t + \frac{\alpha_2(B)}{\delta_2(B)} a_t$$

donde a_t es el ruido (con las notaciones usadas en la sección 2), ya en esta transformación puede verse el camino que conduciría a hacer la formulación que se muestra en la siguiente ecuación.

Como se ve diferencia del primer modelo ésta toma en consideración sólo un parámetro que puede ser suficiente para modelar la humedad con periodos largos de datos (como 24 horas, 1 mes, o más), ya que se usa directamente los datos de humedad de tiempos atrasados

$$H_t = \frac{\alpha_{11} + \alpha_{12} B}{1 - \delta_1 B} H_{t-1} + \frac{1}{1 - \delta_2 B} R_t^{\lambda} \qquad \lambda > 0 \qquad \textbf{Ec. 3.3-8}$$

donde R y H son series de tiempo, $R_{t,n}^{\lambda} = \left(\cdots, r_{t-m}^{\lambda}, \cdots, r_{t-1}^{\lambda}, r_t^{\lambda}\right)$ donde r_t es la observación de lluvia acumulada en el tiempo t y $H_{t-1} = \left(\cdots, h_{t-n}, \cdots, h_{t-2}, h_{t-1}\right)$ donde h_t es la observación de humedad en el tiempo t.

Puede observarse que la mayor responsabilidad de mantener el nivel de humedad está en los coeficientes de H debido principalmente a α_{11} puesto que en periodos secos se tiene $R_t = 0$. La responsabilidad del cambio en la humedad con la presencia de lluvia es atribuida al coeficiente δ_2 de R. obsérvese además que el numerador de este último es 1, esto además de hacer más rápido la optimización de los coeficientes de la función de transferencia, se limita prácticamente a las respuestas con los eventos de lluvia.

La constante λ cuya función es la "calibración" de los efectos de las lluvias en los cambios de niveles de humedad (calibrar los impulsos). Los valores de λ pueden variar por regiones, tipos de suelos, vegetación, etc. como se verá en el siguiente capítulo.

La búsqueda de los parámetros de esta función puede ser realizada por el mismo algoritmo usado en el modelo anterior. A continuación se trata de uno de los algoritmos de optimización llamado *Simplex Search*, y el *algoritmo genético*.

Para asegurarse de que el sistema es estable solo hace falta que las constantes δ_i cumplan con la condición $|\delta_i| < 1, \quad \forall i$, en caso $\exists i / |\delta_i| > 1$ el sistema no es estable y en el caso $\exists i / |\delta_i| = 1$ no se asegura su estabilidad. Este análisis es basado en la convergencia de la serie geométrica, ya que $\frac{\alpha_1(B)}{1-\delta B}$ puede ser expresado como una suma $\alpha_1(B)\sum_{j=1}^{\infty}\phi_i^{\,j} B^j$.

Observando las series, por cada eventeo de lluvia los niveles de humedad cambian de manera abrupta (inpulsos), y esta función de transferencia es especializada en detectar esos cambios. Como se puede imaginar, pueden darse los casos en que se registre poca lluvia sin enbargo los niveles de humedad del suelo sean altos, o viceversa. Estos cambios podrían no ser detectados por el modelo, y si los coeficientes de la función son ajustados en esas condiciones adversas de clima podrían dar malas estimaciones de humedad. Caso típico de estas condiciones atmosféricas puede ser en las temporadas de huracanes u otros fenómenos. Estos casos serán tratados en trabajos futuros.

3.4 Algoritmos de Optimización

La función de transferencia en consideración es un caso de función de dominio discreta, del cual se requieren los coeficientes más óptimos. Los métodos de regresión no es adecuado en este caso debido a la forma en que aparecen los coeficientes ($\delta's, \alpha's$). Se tienen diversas alternativas de solución para este problema, considerando aquellos métodos aplicables a funciones de dominio discreto, entre ellos se tiene el método del *simplex search*, los *algoritmos genéticos*. Ya que el interés principal es optimizar la función y no el tiempo de cómputo, cualesquiera de estos métodos pueden ser usados, para otros casos (funciones continuas) se trata de problemas de máximo o mínimo donde quizás sea conveniente aplicar teorías infinitesimales directamente (como derivadas o gradientes de funciones).

3.4.1 Simplex Serach

El algoritmo de optimización *Simplex Search* [30] es aplicado con frecuencia para encontrar el mínimo (o máximo) de una función de *n* variables, un *simplex* es una figura geométrica de *n* dimensiones, con *n*+1 puntos y si se trata de 2 variables (2 puntos) que se puede ver como un triángulo, 3 variables un tetraedro, así sucesivamente, se describe el método como sigue:

Sea $f : R^n \to R$ una función real de variable vectorial a minimizar (o maximizar) y $x^{(0)} \in R^n$ un punto inicial de dimensión *n*, se eligen *n* puntos adicionales utilizando la siguiente regla, para cada $i, j = 1,...n$:

$$x^{(i)} = \begin{cases} x_j^{(0)} + \delta_1 \ , i \neq j \\ x_j^{(0)} + \delta_2 \ \ i = j \end{cases} \qquad \textbf{Ec. 3.4-1}$$

donde $\delta_1 = \left(\frac{\sqrt{n+1} + n - 1}{n\sqrt{2}} \right) \alpha$, $\delta_2 = \left(\frac{\sqrt{n+1} - 1}{n\sqrt{2}} \right) \alpha$, α es elegido en forma arbitraria.

Sea S el conjunto de $n+1$ los vértices $x^{(i)}$ de un simplex, calculando los parámetros $\delta' s$ calculados y fijando el valor de α seguir los siguientes pasos:

1 $f(x') = \max\{f(s) : s \in S\}$ (vértice a ser reflejado ver Fig. 3.4-1)

2 Calcular el centroide x_c, $x_c = \frac{1}{n} \sum_{\substack{x \in S \\ x \neq x'}}^{n} x$, para obtener $x_{new} = 2x_c - x'$

3 Comparar $f(x_{new})$ y $f(x')$

3.1 Si $f(x_{new}) < f(x')$, entonces $S := \{S - \{x'\}\} \cup \{x_{new}\}$ y repetir los pasos (1)-(3).

3.2 Si $f(x_{new}) \geq f(x')$ reducir tamaño de del simplex $\alpha := \frac{\alpha}{2}$ repetir los pasos (1)-(3).

4 El proceso terminará si $S_f < \varepsilon$, donde $S_f = \sqrt{\frac{1}{n-1} \sum_{i=0}^{n} (f_i - \overline{f})^2}$, con $f_i = f(x^{(i)})$, $\overline{f} = \frac{1}{n} \sum_{i=1}^{n} f_i$.

El proceso de de iteración también puede ser controlado de otra manera, comparando por ejemplo los dos últimos términos obtenidos: $f(x')$ y $f(x_{new})$.

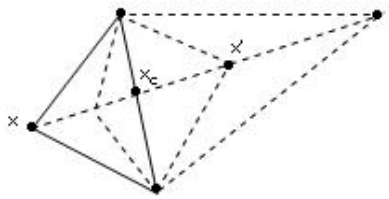

Fig. 3.4-1 Un simplex, caso *n*=2

Esta última figura muestra un simplex de línea continua, y el posible vértice del nuevo simplex formada con las líneas discontinuas, el punto x (el punto a ser reemplazado, o el peor punto) puede moverse a lo largo de la recta que contiene los puntos x, x', el punto x' es el caso de una reflexión simétrica, x_c es el centroide de la arista opuesta a la vértice x, el punto dentro del triángulo es el caso de una contracción.

3.4.2 Algoritmos Genéticos

Una alternativa de resolver problemas complejos de optimización, entre elos minimización o maximización de funciones, es usando *algoritmos genéticos* [31]. Este algoritmo usa la idea del proceso evolutivo biológico (población, reproducción y mutación), cada individuo en la población se llama *cromosoma* que representa una solución del problema, un cromosoma es una cadena de *genes* (bits), aunque no siempre.

Existen solamente dos tipos de operaciones en algoritmos genéticos:

1. Operaciones de Genética: crossover y mutación
2. Operación de evolución: Selección

Crossover es considerado como el principal operador genético, este opera sobre dos cromosomas para generar descendientes.

Mutación es el operador con el cual se altera uno o más genes.

Selección existen tres básicos temas involucrados en la fase de selección: espacio de muestreo, mecanismo de muestreo y probabilidad de selección, la selección, y la consiguiente reproducción se puede hacer en las siguientes formas principales: basado en el rango, rueda de ruleta, selección de torneo.

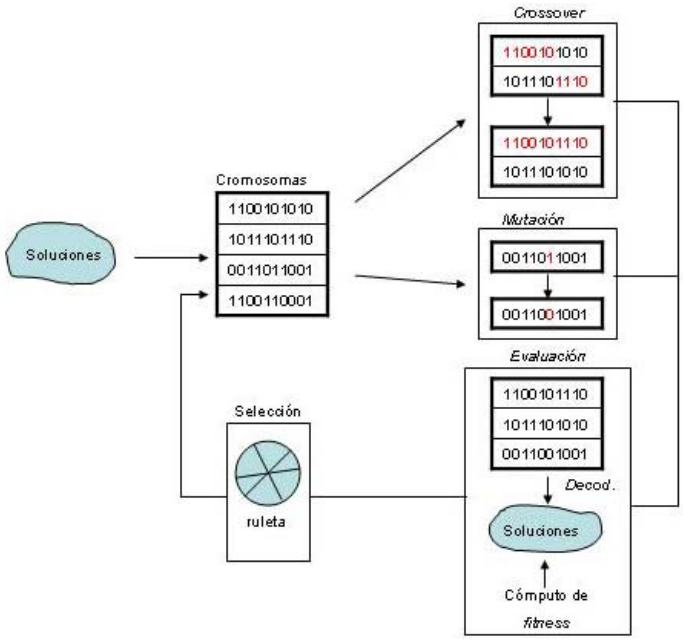

Fig. 3.4-2 Estructura general de algoritmos genéticos

Sean $P(t)$ y $C(t)$ parientes y descendientes en la actual generación t; esta estructura general de algoritmos genéticos se describe en pseudo código siguiente:

```
Program Algoritmos Geneticos
  T←0;
  inicializar P(t)
  evaluar P(t)
  While ~(condicion de finalizar) do
      reconbinar P(t) para generar C(t)
      evaluar C(t)
      seleccionar P(t+1) de P(t) y C9t)
      T← t+1
  End
End
```

Fig. 3.4-3 Pseudo código de algoritmos genéticos

Los algoritmos han mostrado ser versátiles y efectivos para resolver problemas de optimización, aunque en muchas situaciones el algoritmo genético simple no funciona bien, y otros métodos de hibridización han sido propuestos.

Los algoritmos genéticos pueden tratar con cualquier tipo de funciones objetivo y sobre espacios discretos o continuos. Este método se diferencia de los otros métodos de búsqueda en que usa un conjunto de soluciones iniciales llamados población elegidas aleatoriamente.

3.5 Métodos de Interpolación

Algunos métodos de interpolación son apropiados para completar datos faltantes, así por ejemplo la variable de temperatura puede ser estimada en una región a partir de algunos valores estimados dentro de dicha región. Así por ejemplo la temperatura en cualquier punto de una ciudad entera puede ser estimada (interpolada)

conociendo las temperaturas en pocos puntos (estratégicamente ubicadas) dentro de la ciudad.

En la próxima sección se describen los tres métodos de interpolación, usando la regla de Kriging, vecinos más cercanos (KNN), y mediante redes neurales artificiales (RNA.)

3.5.1 Método de Kriging

Básicamente se verá el método "Kriging Ordinario" [5], este método esta basado en la minimización de la varianza, estima un valor desconocido (x_0) usando datos de los vecinos cercanos (x_α) a este, como se muestra en la siguiente figura (Fig. 3.5-1).

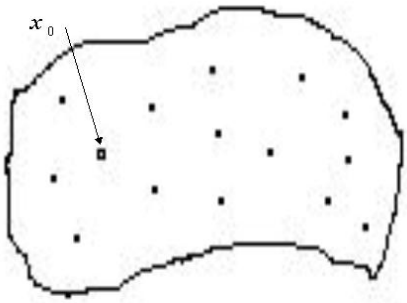

Fig. 3.5-1 Interpolación Kriging

El valor del punto a conocer, $Z(x_o)$ es representada como una combinación lineal de los valores de los n vecinos

$$Z^*(x_0) = \sum_{\alpha=1}^{n} w_\alpha Z(x_\alpha)\text{, con } \sum_{\alpha=1}^{n} w_\alpha = 1 \qquad \textbf{Ec. 3.5-1}$$

Asumir que los datos pertenecen a una función aleatoria $Z(x)$ con variograma $\gamma(h)$ [5], obsérvese el esperado de los incrementos

$$E[Z^*(x_0)-Z(x_0)] = E[\sum_{\alpha=1}^{n} w_\alpha Z(x_\alpha) - Z(x_0)\sum_{\alpha=1}^{n} w_\alpha]$$
$$= \sum_{\alpha=1}^{n} w_\alpha E[Z(x_\alpha)-Z(x_0)] = 0 \qquad \textbf{Ec. 3.5-2}$$

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

$$Z^*(x_0)-Z(x_0) = \sum_{\alpha=1}^{n} w_\alpha Z(x_\alpha) + (-1)Z(x_0) = \sum_{\alpha=0}^{n} w_\alpha Z(x_\alpha) \qquad \textbf{Ec. 3.5-3}$$

donde $\sum_{\alpha=0}^{n} w_\alpha = 0$, con $w_0 = -1$

Y la estimación de la varianza resulta

$$\sigma_E^2 = E[(Z^*(x_0)-Z(x_0))^2]$$
$$= -\gamma(x_0-x_0) - \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} w_\alpha w_\beta \gamma(x_\alpha - x_\beta) + 2\sum_{\alpha=1}^{n} w_\alpha \gamma(x_\alpha - x_0) \qquad \textbf{Ec. 3.5-4}$$

y minimizando (derivando respecto de los pesos) la varianza estimada se obtiene el sistema Kriging Ordinario (*KO*):

$$\begin{pmatrix} \gamma(x_1-x_1) & \cdots & \gamma(x_1-x_n) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma(x_n-x_1) & \cdots & \gamma(x_n-x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} w_1^{KO} \\ \vdots \\ w_n^{KO} \\ \mu_{KO} \end{pmatrix} = \begin{pmatrix} \gamma(x_1-x_0) \\ \vdots \\ \gamma(x_n-x_0) \\ 1 \end{pmatrix} \qquad \textbf{Ec. 3.5-5}$$

donde los w_α^{KO} son los pesos a ser asignados a los valores de los datos y μ_{KO} es el parámetro de Lagrange. Note que el parámetro de Lagrange resulta cuando se convierte el problema con restricciones en un problema sin restricciones

El sistema anterior también puede ser reescrito como

$$\begin{cases} \sum_{\alpha=1}^{n} w_\beta^{KO} \gamma(x_\alpha - x_\beta) + \mu_{KO} = \gamma(x_\alpha - x_0) \\ \sum_{\beta=1}^{n} w_\beta^{KO} = 1 \end{cases}, \text{ para } \alpha = 1,...,n \qquad \textbf{Ec. 3.5-6}$$

La estimación de Kriging Ordinario es

$$\sigma_{KO}^2 = \mu_{KO} - \gamma(x_0 - x_0) + \sum_{\alpha=1}^{n} w_\alpha^{KO} \gamma(x_\alpha - x_0) \qquad \textbf{Ec. 3.5-7}$$

Kriging Ordinario es un interpolador exacto en el sentido de que el error es cero en los puntos conocidos, esto es

$Z^*(x_0) = Z(x_0)$, si $x_0 = x_\alpha$

3.5.2 Método de Vecinos Más Cercanos

El método de vecinos más cercanos (conocido como KNN, K-Nearest Neiborhood) es muy aplicado para diferentes propósitos, existiendo diferentes formulaciones e interpretaciones, por ejemplo cuando se usa para reconocimiento de patrones en imágenes multiespectrales se trata de ordenar los *píxeles* (unidad más pequeña de imagen) vecinos de acuerdo a una métrica de distancia aplicada a cada uno de los píxeles que componen la imagen multiespectral [6].

El método que trataremos aquí esta basado en la métrica de distancia (euclidiana, máximo, supremo, LP, etc.), y por facilidades de interpretación usaremos

la métrica euclidiana, aplicado a las muestras vecinas distribuidas geo-espacialmente, el criterio para obtener un valor pueden ser diversos, puede estar dado por la mayoría, o por el promedio, etc. de los vecinos más cercanos.

Este último es muy usual en manipulaciones de imágenes, así como rotación, variación del número de píxeles (esta última operación es también conocido como *resampling*), entre otros, debido que aquí esta técnica tiene la capacidad de mantener las características de los píxeles de la imagen. Existen otros métodos y técnicas de interpolación, sin embargo para interpolación de datos distribuidos geo-espacialmente se prefiere este método que de una manera más formal se explica como sigue.

Sean $x_1,...,x_n$ datos distribuidas espacialmente a distancias $d_1,...,d_n$ de x_0, sin pérdida de generalidad puede considerarse ordenadas de menor a mayor, $d_1 \leq,...,\leq d_n$, el método consiste en tomar los k primeros datos y aplicarle una función *f* adecuada (dependerá de los tipos de datos) tal *f* puede consistir por ejemplo en tomar el máximo de los datos, la media de los datos, o el valor que más se repite (mayoría), etc.

$x_0 = f(x_1,...,x_k)$, donde *f* puede ser $f = \max, mean, \ ...$

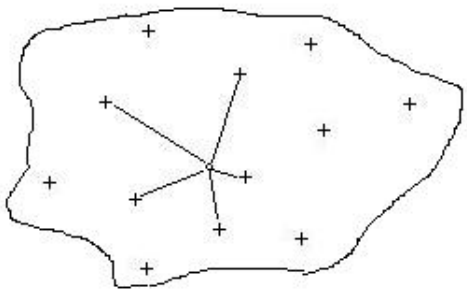

Fig. 3.5-2 Vecinos más cercanos

En la en la figura anterior (Fig. 3.5-1) el símbolo ‘o’ es el valor por estimar, considerando 5 vecinos más cercanos señaladas con el símbolo ‘+’.

3.6 Redes Neurales Artificiales

Una red neural artificial (RNA o simplemente Red) es un sistema para resolver problemas de manera similar al cerebro humano, puede enunciarse como sigue:

Una RNA es un procesador distribuido paralelo formado por unidades de procesamiento simples, el cual tiene la propiedad natural de adquirir conocimiento y hacer disponible para su uso, se asemeja al cerebro humano en dos aspectos:

1. El conocimiento es adquirido por la red a través de un proceso de aprendizaje
2. La interconexión, conocido como peso sinóptico, son usados para guardar el conocimiento adquirido

Cada unidad básica, llamada neurona, tiene la capacidad de aprender. Una forma básica de una red neural se muestra en la siguiente figura (Fig. 3.6-1)

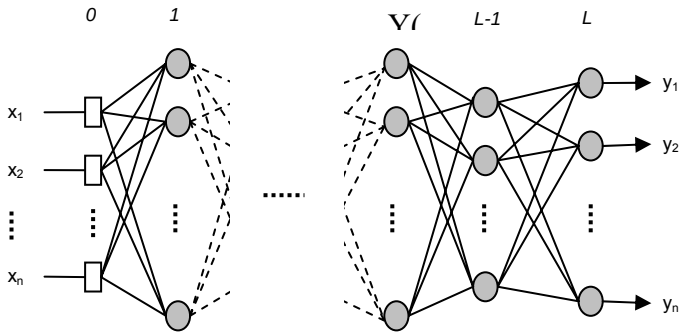

Fig. 3.6-1 RNA de *L*+1 capas

Esta red consta de n+1 capas donde una capa l posee $|N_l|$ ($0 \le l \le L$) nodos, donde N_l es el conjunto de los índices de los nodos en la capa l, la capa 0 es la capa de entrada el cual no posee neuronas, considerándose que la entrada y la salida de la capa 0 es la misma.

Los pesos que conectan las capas l-1 y l pertenecerán a la capa l, de modo que los pesos en la capa l son denotados por $w^l{}_{ij}$ para $i \in N_{l-1}$, $j \in N_l$

3.6.1 Elementos de Una Neurona

Una neurona (Fig. 3.6-2) como elemento procesante de información tiene tres elementos básicos:

1. Un conjunto de conexiones (w_i) caracterizado por *pesos*, específicamente cada señal x_i es conectado a la neurona mediante la multiplicación por su respectivo peso.
2. Un sumador (∑), cada producto de *señal* por *peso* es sumado.
3. Una constante (b llamado *bias*), el cual es conectada a la neurona mediante multiplicación por un peso inicial igual a 1.
4. Una función de activación (f), con el que se controla la amplitud de la salida de la neurona.

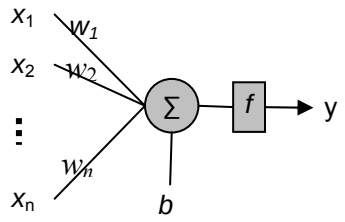

Fig. 3.6-2 Una Neurona y sus elementos

$$y = f(w_1x_1 + w_2x_2 + \cdots + w_nx_n + b) = f(\sum_i w_ix_i + b)$$ **Ec. 3.6-1**

para la *j*-ésima neurona de la capa *l* se tiene la relación siguiente:

$$y_k = f(s) \text{ , donde } s = \sum_{j \in N_{l-1}} w_{ki}x_i + b_k$$ **Ec. 3.6-2**

Obsérvese al considerar dos capas internas de la red, las salidas y_k^l de la capa *l* en función de las salidas y_k^{l-1} de la capa anterior *l*-1, puede ser escrita como:

$$s_k^l = \sum_{j \in N_{l-1}} w_{kj}^l y_j^{l-1} + b_k^l, \quad y_k^l = f(s_k^l)$$

donde los b_k^l que generalmente son inicializados con 1.

El sistema anterior (3.8-2) se puede expresar en términos de matrices o vectores como sigue:

$$s = \begin{bmatrix} w_1 & \cdots & w_n & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ b \end{bmatrix} \text{ o } s = \begin{pmatrix} w_1 & \cdots & w_n & 1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & \cdots & x_n & b \end{pmatrix}$$

donde "." significa producto de vectores (producto interior)

3.6.2 Tipos de Funciones de Activación

Existen algunos tipos de funciones de activación conocidos, éstas por lo general controlan la salida de la neurona a un intervalo [0,1], [-1,1], algunos de ellos son:

b) La función *Sigmoide*

$$f(x) = \frac{1}{e^{-ax} - 1}, \; a > 0$$ **Ec. 3.6-3**

c) La función *Tangencial*

$$f(x) = a\,Tan^{-1}(x)$$

d) La función *escalón unitario*

$$f(x) = \begin{cases} 1 & , x \geq 0 \\ 0 & , x < 0 \end{cases}$$

e) La función *lineal por partes*

$$f(x) = \begin{cases} 0 & , \quad x \leq 1/2 \\ x + 1/2 , & -1/2 < x < 1/2 \\ 1/2 & , \quad x \geq 1/2 \end{cases}$$

5. La función *tangente hiperbólica*

$$f(x) = \tanh(x)$$

3.6.3 Algoritmo "Back Propagation"

Consideremos una red con L+1 capas (capa $L = 0$ es la capa de entrada y la capa L es la capa de salida) y la función de transferencia sigmoide $f = 1/(e^{-x} - 1)$ con $a = 1$, en la n-ésima (n>0) iteración los pesos de la capa l ($1 \leq l \leq L$) con $|N_l|$ nodos son actualizados de la siguiente manera:

- Para la capa L (capa de salida):

  Los pesos en la n-ésima iteración son actualizados mediante

$$w_{ij}^{L}(n+1) = w_{ij}^{L}(n) + \Delta w_{ij}^{L}(n) \quad , n > 0 , j \in N_{L-1}, k \in N_L \qquad \textbf{Ec. 3.6-4}$$

  La medida del error para la red es dada por la siguiente suma de cuadrados

$$E = \frac{1}{2}\sum_{k \in N_L}(t_k - y_k^L) \qquad \textbf{Ec. 3.6-5}$$

ya que el objetivo es minimizar el error de salida, entonces se consideran las derivadas del error respecto de los pesos

$$\Delta w_{ij}^L(n) = -\eta \frac{\partial E}{\partial w_{lij}^L} \qquad \textbf{Ec. 3.6-6}$$

donde η es el factor de aprendizaje de la red, haciendo uso de la regla de la cadena el factor derivada parcial (respecto de los pesos) del error E puede ser expresado como sigue:

$$\frac{\partial E}{\partial w_{kj}^L} = \frac{\partial E}{\partial y_k^L}\frac{\partial y_k^L}{\partial s_k^L}\frac{\partial s_k^L}{\partial w_{kj}^L} \qquad \textbf{Ec. 3.6-7}$$

donde $s_k^L = \sum w_k^L y_i^{L-1} + b_k^L$, cada uno de los factores de esta última expresión (ec. 3.8-7) pueden ser expresadas como sigue:

$$\frac{\partial y_k^L}{\partial s_k^L} = f'(s_k) = (1 - y_k^L)y_k^L \qquad \textbf{Ec. 3.6-8}$$

$$\frac{\partial s_k^L}{\partial w_{kj}} = y_j^{L-1} \qquad \textbf{Ec. 3.6-9}$$

reemplazando éstas tres últimas expresiones en la expresión anterior (3.6-7) se tiene:

$$\Delta w_{kj} = \eta(t_k - y_k^L)(1 - y_k^L)y_k^L y_j^{L-1}, \;\; j \in N_{L-1}, k \in N_L \qquad \textbf{Ec. 3.6-10}$$

- Para la capa L-1:

  De manera análoga como en la capa L, podemos escribir Δw_{ij}^{L-1} como:

$$\Delta w_{ji}^{L-1} = -\eta \frac{\partial E}{\partial y_j^{L-1}} \frac{\partial y_j^{L-1}}{\partial s_j^{L-1}} \frac{\partial s_j^{L-1}}{\partial w_{ji}^{L-1}}, \; i \in N_{L-2}, \; j \in N_{L-1} \qquad \textbf{Ec. 3.6-11}$$

$$\Delta w_{ji}^{L-1} = -\eta \frac{\partial E}{\partial y_j^{L-1}} (1 - y_j^{L-1}) y_j^{L-1} y_i^{L-2} \qquad \textbf{Ec. 3.6-12}$$

$$\frac{\partial E}{\partial y_j^{L-1}} = \sum_{k \in N_L} \frac{\partial E}{\partial s_k^L} \frac{\partial s_k^L}{\partial y_j^{L-1}} = \sum_k \frac{\partial E}{\partial s_k^L} w_{kj}^L \qquad \textbf{Ec. 3.6-13}$$

nuevamente como en el paso anterior se tiene:

$$\frac{\partial E}{\partial s_k^L} = \frac{\partial E}{\partial y_k^L} \frac{y_k^L}{s_k^L} = -(t_k - y_k^L)(1 - y_k^L) y_k^L \qquad \textbf{Ec. 3.6-14}$$

reemplazando esta última expresión (3.6-14) en la ecuación anterior (3.6-13) y la ecuación (3.6-13) en (3.6-11) se tiene:

$$\Delta w_{ij} = \eta (1 - y_j^{L-1}) y_j^{L-1} y_i^{L-1} \sum_{k \in N_L} (t_k - y_k^L)(1 - y_k^L) y_k^L w_{kj}^L \qquad \textbf{Ec. 3.6-15}$$

Por fines prácticos se da la siguiente notación:

$$\delta_k^L = (t_k - y_k^L)(1 - y_k^L) y_k^L \qquad \textbf{Ec. 3.6-16}$$

$$\delta_j^{L-1} = (1 - y_j^{L-1}) y_i^{L-1} \sum_{k \in N_L} (t_k - y_k^L)(1 - y_k^L) y_k^L w_{kj}^L = (1 - y_j^{L-1}) y_j^{L-1} \sum_{k \in N_L} \delta_k^L w_{kj}^L \qquad \textbf{Ec. 3.6-17}$$

de modo que

$$\Delta w_{kj}^L(n) = \eta \delta_k^L y_j^{L-1} \qquad \textbf{Ec. 3.6-18}$$

$$\Delta w_{kj}^{L-1}(n) = \eta \delta_k^{L-1} y_j^{L-2} \qquad \textbf{Ec. 3.6-19}$$

Este proceso puede aplicarse nuevamente para la capa siguiente (L-2) y así sucesivamente hasta la capa 1.

Una modificación de este algoritmo es el *backpropagation* con *momentum*, donde los pesos se corrigen de la siguiente manera:

$$\Delta w_{ij}^{l}(n) = \eta \delta_i^l o_j^{l-1} + \mu \Delta w_{ij}^{l}(n-1) \qquad \textbf{Ec. 3.6-20}$$

para $i \in N_l$, $j \in N_{l-1}$; η, μ son la razón de aprendizaje y momentum respectivamente

En resumen, los δ's son calculados para cada capa de la siguiente manera:

a. Para la capa de salida (l=L):

$$\delta_i^L = y_i^L (1 - y_i^L)(t_i - y_i^L), \quad i \in N_L \qquad \textbf{Ec. 3.6-21}$$

b. Para las capas internas ($1 \le l \le L-1$):

$$\delta_i^l = y_i^l (1 - y_i^l) \sum_{k \in N_l} \delta_k^{l+1} w_{kj}^{l+1}, \quad j \in N_l \qquad \textbf{Ec. 3.6-22}$$

3.7 Regla de Kohonen

Como se mencionó antes, una red es entrenada para reconocer (o clasificar) características de objetos, una clasificación puede ser hecha de dos maneras: conociendo las muestras de entrenamiento (clasificación supervisada) o sin conocerlo (clasificación no supervisada), en el caso de conocerse las muestras de las clases puede usarse una RNA con entrenamiento BP o LM (más sofisticado que BP), si no se disponen de muestras de entrenamiento ya no es factible usar el método anterior, sin embargo existen diversos métodos para realizar una clasificación no supervisada así como K-mean, Kohonen Self Organizing Feature Map (SOFM) [12], entre otros, este último es basado en una RNA de una capa el cual funciona como sigue:

a. Asumir que los nodos de salida están conectadas a un arreglo (de 1 o 2 dimensiones)

b. Asumir que la red es completamente conectada, todos los nodos de la capa de entrada están conectados a todos los nodos en la capa de salida

c. Usar el algoritmo competitivo de aprendizaje como sigue:

- Escoger aleatoriamente un vector de entrada x
- Determinar el ganador nodo de salida ganador i, con peso w_i conectando las entradas a los nodos de salida

$$|w_i - x| \leq |w_k - x|, \quad \forall k \qquad \textbf{Ec. 3.7-1}$$

- Dado el nodo ganador (i), los pesos son actualizados:

$$w_k(\text{nuevo}) = w_l(\text{antiguo}) + \mu\aleph(i,k)(x - w_k) \qquad \textbf{Ec. 3.7-2}$$

donde $\aleph(i,k)$ es llamado la función de vecindad que tiene valor 1 cuando $i = k$, y cae dentro de las unidades del arreglo de salida entre i y k con la norma valor absoluto $|r_k - r_i|$, $r \in R$. Los pesos asociados con los otros nodos de salida lejanos no cambian significativamente.

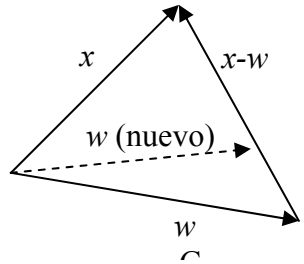

Fig. 3.7-1 Regla de Kohonen

Un ejemplo de función vecindad es $\aleph(i,k) = e^{-|r_k - r_i|^2/(2\sigma^2)}$, donde el parámetro σ disminuye gradualmente con el tiempo.

En este algoritmo existe una estructura topológica impuesta sobre los nodos de la red, este proyecta un mapa que preserva las relaciones de vecindad, y los nodos de salidas son arreglos geométricos.

3.8 Algoritmos de Estimación de Temperatura de Superficie e Índice de Vegetación

Los datos de observación de MODIS están disponibles vía FTP desde su página Web (http://modis.gsfc.nasa.gov), éstos datos están guardados reescalados a entero (SI: scaled integer) y en formato científico HDF (Hierarchical Data Format), para obtener los datos en unidades físicas debe hacerse una transformación con las constantes *scale* y *offset* que se encuentran en el conjunto de datos mediante las siguientes ecuaciones:

$$rad = rad_scale(SI - rad_offset)$$

$$refl = refl_scale(SI - refl_offset)$$

donde *rad_scale* y *rad_offset* son las constantes de transformación de los datos en SI a radianza, constantes análogos en la segunda ecuación para la transformación a reflectancia.

Esta reflectancia ($\rho *$) es obtenida por el sensor y su valor varía con el ángulo solar θ, este efecto puede ser corregido mediante

$$\rho = \frac{\rho *}{\cos(\theta)} \qquad \textbf{Ec. 3.8-1}$$

donde ρ es la radianza neta de la superficie.

La reflectancia, L, obtenida es convertido a temperatura de brillantez (TB) mediante la ecuación de Plank

$$TB = \frac{hc(\lambda k)^{-1}}{\log(1 + 2hc^2 L^{-1})} \qquad \textbf{Ec. 3.8-2}$$

donde L es la reflectancia captada por el sensor, h, k, c son constantes cuyos valores son:

c = 299 792 458 m/s (velocidad de la luz)

$h = 6.6261 \times 10^{-34}$ Js (constante de Planck)

$k = 1.380\ 6503 \times 10^{-23}$ J K^{-1} (constante de Boltzmann)

λ: Longitud de onda de observación

3.8.1 Selección de Píxeles, Máscara

Debido a la presencia de nubes en la atmósfera, las observaciones en el espectro visible y termal son afectadas, existen algoritmos que permiten corregir en cierta medida estas perturbaciones causadas por nubes de agua, aerosoles u otras partículas presentes en la atmósfera; en este trabajo no se consideran los algoritmos de corrección atmosférica sino se separara las regiones nubladas mediante la construcción de una máscara (máscara de nube), esta máscara está basada en las diferencias de temperaturas de brillantez (TB) de los canales 20 y 31 del sensor MODIS, el algoritmo es como sigue:

- Para superficie de suelo:
    - $TB_{20} - TB_{31} > 6$, implica la presencia de nube cirrus denso. Asignar el valor 1 a la máscara
    - $TB_{20} < 270$, implica la presencia de nube delgada. Asignar el valor 2 a la máscara
    - Si ninguno de los casos anteriores se satisface, implica la no presencia de nubes. Asignar valor 0 a la máscara.
- Para superficie del Mar
    - $TB_{20} - TB_{31} > 10$, implica la presencia de nube cirrus denso. Asignar el valor 11 a la máscara
    - $TB_{20} < 273$, implica la presencia de nube delgada. Asignar el valor 12 a la máscara
    - Si ninguno de los casos anteriores se satisface, implica la no presencia de nubes. Asignar valor 10 a la máscara.

3.8.2 Temperatura de Superficie (LST)

La temperatura de superficie conocido como LST (Land Surface Temperature) puede también ser estimada usando las imágenes multiespectrales obtenidas desde el satélite TERRA/MODIS, específicamente por la utilización de las bandas 31 y 32, éstas bandas son equivalentes a las bandas 4 y 5 del satélite NOAA/AVHRR aunque con diferente ancho de banda.

Experimentalmente [7] se observa la relación lineal entre $T_s - T_5$ y $T_s - T_4$, de modo que

$$T_s - T_5 = a(T_s - T_4) + b \qquad \textbf{Ec. 3.8-3}$$

donde T_4, T_5 son las temperaturas de brillantez de los canales 4 y 5, y T_s es la temperatura de superficie.

$$T_s = T_4 + \frac{1}{a-1}(T_4 - T_5) - \frac{b}{a-1} \qquad \textbf{Ec. 3.8-4}$$

las constantes a y b son obtenidas mediante regresión [7], de modo que se tiene la relación:

$$T_s = 1.0346T_4 + 2.5779(T_s - T_5) - 10.05 \qquad \textbf{Ec. 3.8-5}$$

Nota. En la práctica la temperatura de superficie puede aproximarse con la temperatura de aire.

3.8.3 Índice de Vegetación (NDVI)

El índice de vegetación es conocido como NDVI (Normalized Difference of Vegetation Index) es un indicador de salubridad de la vegetación. Existen otros indicadores como LAI (Leaf Area Index) [9] el cual considera la geometría de la planta, para estimar el NDVI se aprovecha del principio de fotosíntesis de la planta, las plantas en éste proceso absorben la energía solar como puede observarse en los niveles de la reflectancia en el espectro que se muestra en la figura siguiente (Fig. 3.8-1), obsérvese principalmente en el rango [0.4, 1.1] μm.

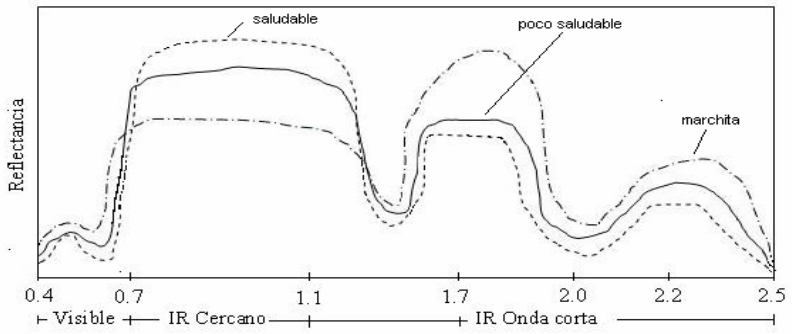

Fig. 3.8-1 Reflectancia de la vegetación

Este parámetro es posible obtenerlo desde las imágenes multiespectrales, aquellas con rango de frecuencia *rojo* e *infrarrojo cercano* (en el caso de los canales de MODIS $Ch1$: 0.620-0.670 μm. y $Ch2$: 0.841-0.876 μm). Usamos las imágenes del satélite TERRA (sensor MODIS) correspondientes a los canales 1 y 2 ($Ch1$ y $Ch2$) para obtener el índice de vegetación mediante la siguiente fórmula:

$$NDVI = \frac{Ch2 - Ch1}{Ch2 + Ch1} \qquad \textbf{Ec. 3.8-6}$$

los mismos canales 1 (rojo) y 2 (infrarrojo cercano) del sensor AVHRR también pueden ser usados [8]. La resolución espacial de este satélite es aproximadamente 1Km, y el sensor MODIS en las mismas bandas es 250 *m* (ver Apéndice C).

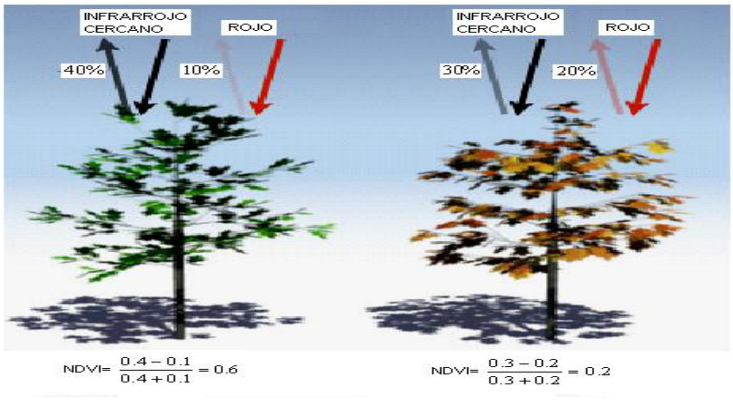

Fig. 3.8-2 Ejemplo de NDVI

En la figura arriba (Fig. 3.8-2) se muestran dos plantas con estados de salubridad diferentes, el de la izquierda se muestra más saludable con índice *NDVI* = 0.6. Realmente el índice NDVI no siempre indica el estado de salubridad de una planta sino que también puede indicar la abundancia de vegetación, sin embargo nos da una idea de la existencia o estado de la vegetación.

Una de las utilidades del índice NDVI está en el monitoreo de la vegetación para detectar por ejemplo incendios o el cambio de su estado con el tiempo, también puede ser usado para distinguir la Tierra y el Mar (o trazar las costas de un continente), en correcciones atmosféricas, etc.

3.9 NEXRAD (Next generation Radar)

Ya que uno de los objetivos de este trabajo de tesis es también estimar no solo en forma puntual sino también en forma espacial, para ello es necesario obtener datos de lluvia en toda el área en consideración (Puerto Rico). Afortunadamente, se dispone

de las observaciones del radar NEXRAD (Next generation Radar) localizado a *Latitud*= 18. 1156 y *Longitud*=-66. 0781 (cerca de San Juan-Puerto Rico).

Fig. 3.9-1 NEXRAD

La figura (Fig. 3.9-1) muestra un de las instalaciones de una de las estaciones del sistema NEXRAD, ésta puede detectar lluvia alrededor de 80 millas náuticas, este radar puede observar y cuantificar la lluvia y velocidad de viento. La reflectancia de las partículas de lluvia traducidas a niveles de precipitación son obtenidas en intervalos de tiempo cortos, cada 5 *min*.

El sistema general de un radar Doppler [29] es como muestra la figura siguiente:

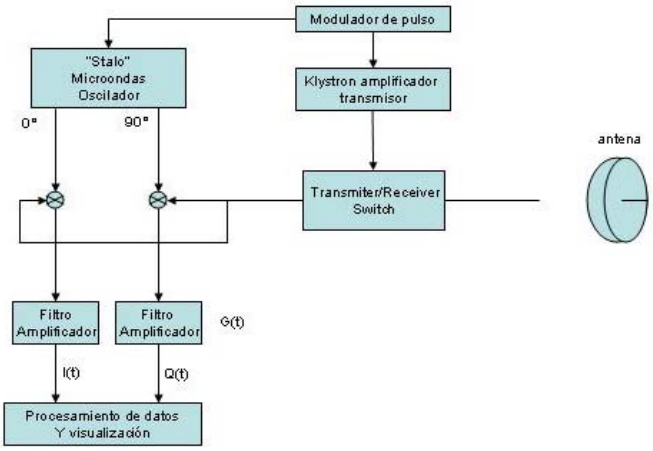

Figura 3.9-1 Diagrama simplificado de radar Doppler

Ecuación ideal de radar, sin considerar pérdidas:

$$P_r = \frac{P_t G_t G_r \lambda^2 \sigma}{(4\pi)^3 R^4}$$

donde P_t (en wats) es la potencia recibida, G_t es la ganancia del transmisor, G_r es la ganancia del receptor, λ es la longitud de onda, R es el rango del radar (alcance máximo), σ (radar cross-section) depende de varios factores: material, frecuencia, geometría, ángulo de incidencia, polarización.

Las observaciones de Radar pueden diferir de las observaciones de Raingauge en considerable magnitud debido a la pobre resolución espacial y otras características naturales de diseño del radar.

Estudios hechos [27] para comparar ambas fuentes demuestran una sobreestimación de la lluvia acumulada obtenidas por el radar, la razón G/R es 1.36 cuando se considera toda el área de barrido del radar, esta razón es definida por:

$$\frac{G}{R} = \frac{\sum_{i=1}^{n} G_i}{\sum_{i=1}^{n} R_i}$$

donde G_i es la i-ésima muestra obtenida de raingauge, y R_i es la i-ésima muestra obtenida de radar, valores G/R cercanos a 1 indicarían buena estimación de lluvia mediante radar.

Con el mismo trabajo en [27] se demuestra que si las muestras se toman haciendo una previa clasificación, en regiones con gradiente de reflectividad alta

(HRG) y los de gradiente de reflectividad baja (LRG), resulta ~0.88 para el primer caso y ~2.23 para el segundo caso en promedio, este último indica una subestimación.

G/R depende también de la estación del año, Junio-Agosto: ~0.79, Octubre-Enero: ~1.82.

G/R depende de la distancia también, las comparaciones hechas dividiendo en cuatro intervalos a lo largo del radio de alcance del radar: 0-0.49.9, 50-99.9, 100-149.9 y 150-230 in Km. como se muestran en la siguiente tabla.

Tabla 3.9-1 Variaciones con la distancia del radar

G/R	0-49.9	50-99.9	100-149.9	150-230
HRG	1.18	0.79	0.79	0.97
LRG	2.29	1.57	2.33	2.88
N muestras	101(13)	239(16)	256(23)	244 (22)
N muestras	143 (15)	197(15)	259(14)	200 (15)

En la tabla los valores entre paréntesis indican el número de eventos de lluvia colectadas, y el número total de muestras fueron 43 en todo el área seleccionado.

Los valores menores que 1 indican sobrestimación, y valores mayores que 1 indican subestimación de radar. Para los casos de alta reflectividad (HRG), la tabla exhibe una subestimación en distancias cercana, y una sobrestimación a distancias medias.

4 Materiales y Procedimiento

Para la implementación de los algoritmos, almacenamiento y manejo de los datos, se cuenta con los laboratorios dentro de la Universidad de Puerto Rico, y como fuente de datos a diversos sistemas de almacenamiento y distribución de datos de los Estados Unidos.

Los algoritmos que se mencionan aquí son implementados usando el software MATLAB principalmente, y los software C/C++, FORTRAN como herramienta para el procesamiento de los datos crudos, y datos empaquetados de tamaños y cantidades relativamente grandes.

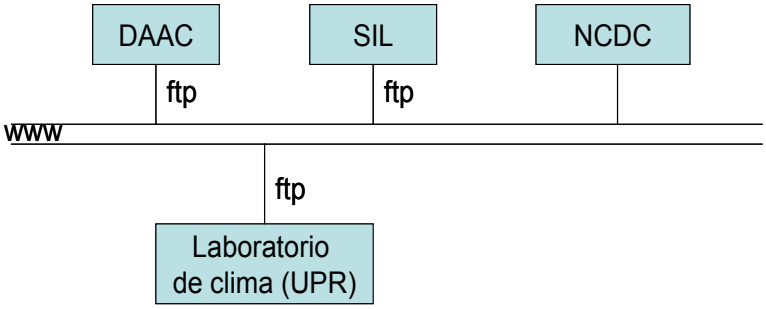

Fig. 4.1-0 Fuentes de datos

A continuación se describen los componentes del sistema que se muestra en la figura arriba (Fig. 4.1 0) de las fuentes y procesamiento de datos usados en el desarrollo de este trabajo.

4.1 El Laboratorio de Climatología

El laboratorio de Climatología está ubicado en la Facultad de Ingeniería Mecánica de la Universidad de Puerto Rico, consta de un sistema servidor (Climatemaster) SUN Solaris SO UNIX, y 6 terminales en sistema UNIX provistas de los software's C/C++, FORTRAN, y MATLAB, entre otras aplicaciones de utilidad para el análisis y la visualización de los datos, además 2 terminales de la red INME (Red de servidor de Ingeniería Mecánica) con sistema Windows XP.

4.2 El Laboratorio PaSCoR

El laboratorio PaSCoR (Partnership for Spatial and Computational Research) es de los laboratorios para el procesamiento y visualización de datos georeferenciados, esta posee software tales como ERDAS, ArcView/ArcMap entre otros en sistema Windows XP útiles para el análisis y clasificación de imágenes, posee fotografías aéreas del área de Puerto Rico, y sistemas de mediciones y georeferenciación GPS (Global Positioning System) de alta resolución (~1m), y su respectivo software de procesamiento y visualización de datos de GPS.

4.3 Fuentes de Datos

Los datos que se usan en este trabajo provienen de diferentes fuentes de información tales como:

4.3.1 DAAC

El centro de procesamiento y distribución EOS **(Earth Observing System Data Gateway)** de los Estados Unidos de Norte América es un sistema de almacenamiento y distribución de datos geofísicos, es uno de 8 empresas **DAAC**s (Data Active Archive Center) dedicadas al almacenamiento y distribución de datos para el público en general http://daac.gsfc.nasa.gov/ .

4.3.2 SIL

El laboratorio SIL (Space Information Laboratory) es un de los componentes del TCESS (Tropical Center for Earth and Space Studies) ubicado en la Universidad de Puerto Rico es encargado de la recepción procesamiento y distribución de imágenes satelitales tales como NOAA, LANDSAT, MODIS y RADARSAT para fines de investigación científica, En la figura siguiente (Fig. 4.3-1) se muestran las instalaciones de las antenas receptoras. http://tcess.uprm.edu.

Fig. 4.3-1 SIL

4.3.3 WCC

Los datos de muestreo en el área de Puerto Rico son proveídos actualmente por Natural Resources Conservation Service (NRCS), a través de su unidad "National Water Climate Center" (WCC) del departamento de agricultura de los Estados Unidos. http://www.wcc.nrcs.usda.gov/scan/.

Las estaciones de humedad de suelo (Fig. 4.3-2) están dotados de sensores que miden temperatura de aire y suelo por niveles de profundidad (por capas), velocidad de viento, humedad relativa del aire, humedad de suelo por niveles de profundidad, lluvia acumulada por horas y días, radiación solar, entre otros. En la siguiente tabla se muestran algunas de las características de suelo y su geolocalización (- significa que no hay información) del lugar de las estaciones experimentales principalmente su ubicación y tipo de suelo.

Tabla 4.3-1 Localización y características de las estaciones experimentales

Estaciones	**Coordenadas (Grados)**	**Elevación (pies)**	**Profundidad (cm)**	**Arena (%)**	**Barro (%)**	**Densidad (gr/cc)**
Maricao	18° 09´ 26.5´´ N 67° 00´ 14.4´´ W	2450	0-5	31.4	42	1.5
Guilarte	18° 08´ 47.6´´ N 66° 46´ 02.8´´ W	3345	0-5	10.3	57.7	1.09
Guánica	17° 58´ 20.1´´ N 66° 52´ 05.4´´ W	541	0-5	25	55	1.5
Combate	17° 58´ 45.9´´ N 67° 10´ 07.9´´ W	33	0-5	81.8	11.9	1.59
Isabela	18° 17' N 67° 02' W	50	-	-	-	-

Debido a la carencia de datos continuos (suficientes se consideran para la prueba del modelo solo dos de ellos, los datos de algunas de las estaciones son discontinuos por problemas de instrumentación o su sistema de transmisión de datos.

Fig. 4.3-2 Cinco estaciones WCC

4.3.4 USGS

United States Geological Survey (USGS) con más de 120 estaciones experimentales de lluvia y flujo de agua en los ríos, es la institución con mayor número de estaciones en cuanto al parámetro de lluvia, estas estaciones proveen medidas de lluvia acumulada por cada 15 minutos. http://pr.water.usgs.gov/

El área de Puerto Rico posee también una gamma de tipos de suelos, los datos de tipos de suelos que se muestra en la siguiente figura (Fig. 4.3-3) son administrados por la USGS.

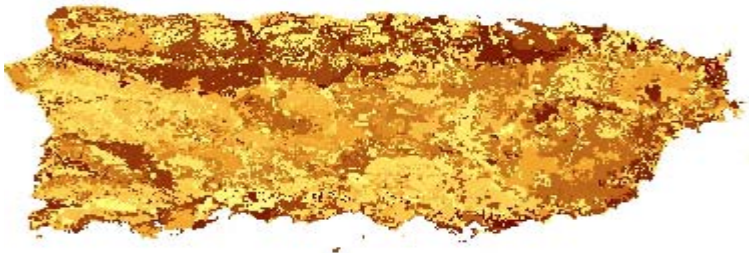

Fig. 4.3-3 Clases de suelos

El mapa muestra más de 400 clases de suelos y puede ser visualizado haciendo uso del software ArcView/ArcMap, el cual es exportado a un formato ASCII para su utilización en Matlab manteniendo la identificación de las clases y su geolocalización. Ya que los datos requeridos para en este trabajo de tesis es a resolución de 1 *Km*, aunque con la operación de cambio de formato y resolución se pierden algunas clases debido a la reducción de la resolución espacial (de ~250*m* a 1*Km*) como es requerida, originalmente algunas clases se encuentran en áreas de menos de 1x1Km^2.

4.3.5 NCDC

El Nacional Climate Data Center, uno de los mas grandes centros de recopilación y distribución de parámetros geofísicos, esta provee al público en general datos como temperatura, lluvia, humedad, entre otros incluyendo datos captados por satélites y de radares. http://www.ncdc.noaa.gov/oa/ncdc.html.

El usuario puede hacer un requerimiento vía Internet a través de su Web servidor, la figura siguiente (Fig. 4.3-4) muestra un mapa con las localizaciones de

los sensores disponibles para la región de Puerto Rico, el cual además posee herramientas gráficas de selección de datos.

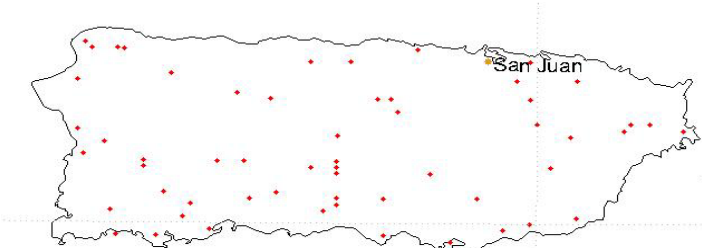

Fig. 4.3-4 Localización de las estaciones NCDC

4.4 Otras Fuentes de Datos

Otras fuentes de datos que hacen posible este trabajo son las publicaciones y la gentileza de aquellos que facilitaron sus trabajos en forma digital, como es el siguiente.

Clases de vegetación, El área de Puerto Rico está cubierta por una variedad compleja de vegetación, debido a la ubicación geográfica lo cual lo convierte en una región tropical, con abundante vegetación y lluvias. Los tipos de vegetación más abundantes se muestran en la figura siguiente (

Fig. **4.4-1**) donde la clasificación fue realizada tomando imágenes de satélite (Landsat) [10] basándose en el índice de vegetación y previa clasificación en regiones con características similares, también se consideran dentro de las clases las zonas urbanas y áreas de cultivo.

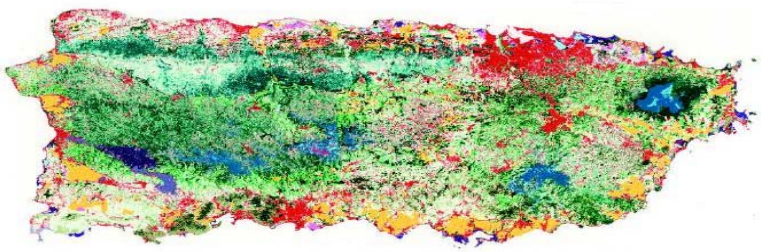

Fig. 4.4-1 Clases de vegetación

Fig. 4.4-2 Colores asignados a las clases

Este mapa está generado con el software ERDAS (disponible en PaSCoR), y con el propósito de su utilización en este trabajo también ésta es exportado al formato ASCII cambiando la resolución espacial a 1*Km,* aunque en el proceso hay perdidas de dos de las clases, aquellas clases de áreas pequeñas, recordándose que en este caso aparecen regiones con áreas pequeñas debido a la resolución del sensor que porta el satélite Landsat es de 30*m* el cual distingue regiones con áreas de menos de 1x1Km^2.

Debe notarse la existencia de una aparente similitud con el mapa se suelos (Fig. 4.3-3)

4.5 Regiones Climatológicas

Con la finalidad de disminuir el número de funciones de transferencias a construir se determinaron cinco regiones mediante una clasificación no supervisada, este método es basado en RNA (y la regla de Kohonen) [12]. Se realizó una clasificación con datos históricos (40 años) de temperatura, lluvia, humedad de suelo, vegetación, elevaciones y rugosidad de la superficie, sobre todo el área PR.

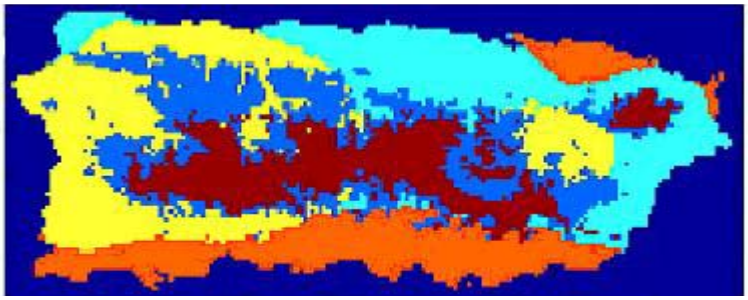

Fig. 4.5-1 Cinco regiones climatológicas

Se determinaron 5 regiones (Fig. 4.5-1) con los datos históricos disponibles mencionados, aparentemente las 5 regiones determinadas aquí sugieren la construcción de solo 5 funciones de transferencia aunque según los datos de muestreo obtenidos sugieren la necesidad de más clases, debido posiblemente a la existencia de diversidad de suelos y vegetación en cada región los cuales se reflejan en las series de tiempo de la humedad obtenidas de 5 (dos de ellas se muestran en las sección 3.3) diferentes estaciones experimentales del área oeste de la Isla.

La determinación de las regiones es con el propósito de minimizar el número de estaciones experimentales de muestreo.

Notación.

Usaremos a veces la notación conjuntista (denotando regiones o áreas o clases como conjuntos)

$$A = A_1 \cup A_2 \cup \ldots \cup A_n$$

para representar la descomposición de una región A en subregiones (o clases) A_i disjuntas, en caso de 5 regiones $n = 5$, en caso se las clase de vegetación se escribirá $V = V_1 \cup V_2 \cup \ldots \cup V_{30}$ para significar 30 clases, con esto verá al área de Puerto Rico como una unión de de 30 clases .

Se escribirá $g \in V_i$ para referirse a que la geolocalización g corresponde a la i-ésima región (o clase).

5 Interpolación Espacial de Parámetros Geofísicos

5.1 Interpolación de NDVI

Los cambios en los tipos de vegetación no es muy frecuente, donde había bosque de árboles sigue habiendo los mismos árboles, salvo los cambios causados por las crecientes urbanizaciones, las talas forestales o incendios. Bajo estas consideraciones, para interpolar los valores NDVI se afirma que: Si un píxel tiene buena información entonces los otros píxeles de regiones cercanas con la misma clase de vegetación que carecen de buena información reemplazando mediante el uso de algún criterio (o función).

En un primer intento por interpolar, el método más apropiado que puede dar un resultado es el método de vecinos más cercanos (KNN, K-Nearest Neiborhood), que en este caso la métrica es aplicada a las distancias espaciales (en el caso de clasificación de imágenes la métrica es aplicada a los datos espectrales) pudiéndose llamar distancias espectrales. El problema con este método es que hay más inexactitud al ser aplicado debido a que los datos de las fronteras no siempre dan buena información de vegetación, debido a que la vegetación no está influenciada por presencia de nubes en su atmósfera ni por la vegetación en áreas vecinas. El método de Kriging también presenta deficiencias desde que esta funciona asignando pesos a los datos circundantes a los faltantes. Ambos métodos pueden dar una buena aproximación si los datos obtenidos del satélite son al 100% puro y el muestreo son suficientes.

La estrategia señalada para este caso es usar el mapa de clases de vegetación, el diagrama de la figura siguiente (Fig. 5.1-1) muestra el proceso de obtención de NDVI usando el mapa de vegetación.

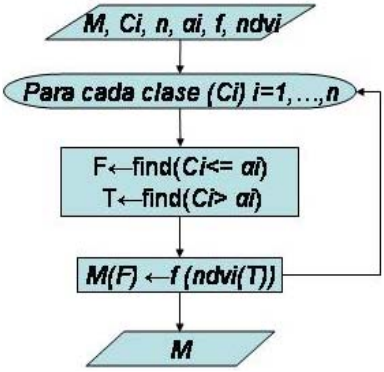

Fig. 5.1-1 Algoritmo de interpolación de NDVI

Aquí las entradas son los índices de las clases de vegetación (*Ci*), *n* el número de clases, NDVI imagen obtenida de MODIS, α es un parámetro que depende de cada clase *i* (αi, threshold) el cual divide los píxeles de la clase en buenos y no buenos, este valor α es característico de cada clase de vegetación, asumiendo por ejemplo para el caso de una clase de vegetación cuyos valores de NDVI esta en [*0.5*, *0.7*] entonces $\alpha=0.5$, valores menores de α indican una subestimación del índice NDVI debido posiblemente a la presencia de nubes tenues, puede asumirse que casi siempre ocurren subestimaciones. M es el resultado de reemplazar los píxeles no buenos por una operación elegida por algún criterio *f*, por ejemplo *f* puede ser la media o quizás la función máximo.

Fig. 5.1-2 Composición de 16 días de NDVI, Nov. 2003

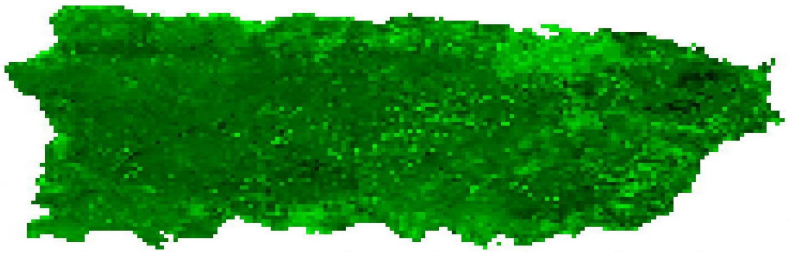

Fig. 5.1-3 Composición de 16 días de NDVI, Interpolación, Nov. 2003

La figura que se muestra arriba (Fig. 5.1-3) muestra el resultado de aplicar el algoritmo anterior, verde mas claro indica menor valor de NDVI. En el caso de la región ubicado en el Yunque, casi siempre hay presencia de nubes de manera que se hace difícil obtener una imagen que represente un valor real de NDVI, incluso es difícil obtener la imagen compuesta por varios días como es el caso que se muestra en la figura anterior (Fig. 5.1-2) el cual es una composición de 16 días consecutivos de imágenes. Debe tenerse presente que este método no puede ser usado para monitorear la vegetación, sino aplicado en condiciones normales para propósitos específicos.

5.2 Interpolación de Temperatura Usando Kriging

Es razonable interpolar por este método puesto que donde hay un píxel oscuro (píxel sin dato de temperatura) se toman los datos de los píxeles vecinos para la interpolación, para interpolar este parámetro también puede usarse KNN.

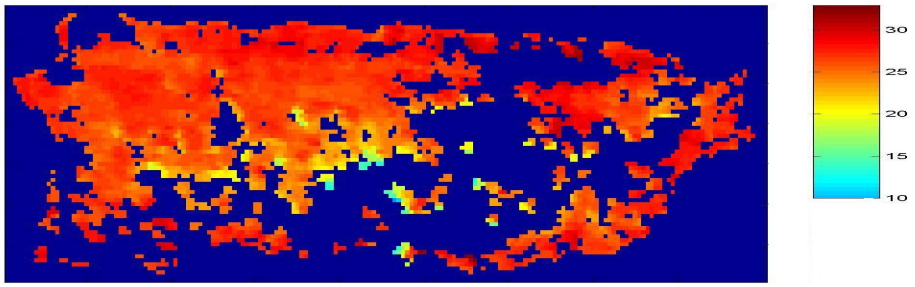

Fig. 5.2-1 Temperatura °C de superficie, 1 Nov. 2003

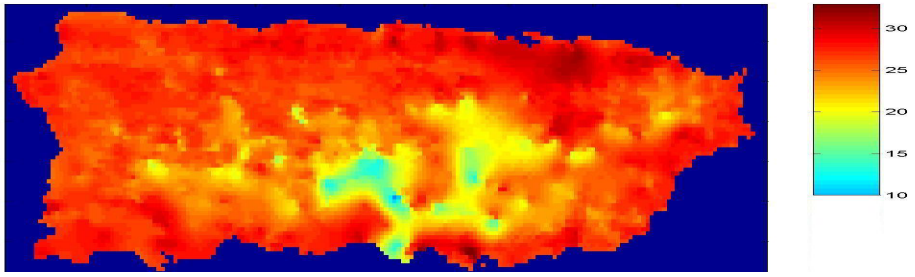

Fig. 5.2-2 Temperatura °C de superficie, 1 Nov. 2003

La figura que ese muestra arriba (Fig. 5.2-1) muestra la imagen de temperatura de superficie obtenida por el sensor MODIS, en la que se observan regiones con píxeles sin datos de temperatura, para su interpolación (Fig. 5.2-2) se asume que las perturbaciones separadas por la máscara son netamente las causadas por la presencia de nubes (no causadas por averías en el sensor u otros factores), de este modo se

puede pensar que las fronteras de las regiones sin datos de temperatura pueden dar información en las regiones cubiertas por nubes.

En la figura (Fig. 5.2-1) puede observarse que en algunas regiones oscuras (sin datos de temperatura) se presentan bordes con niveles de temperaturas bajas (ver la leyenda de imagen), y debe notarse además que las temperaturas bajas no son solo debido a la elevación del lugar sino que esto es debido al estado real del tiempo en esa región (regiones cubiertas por nubes o con lluvias presentan menor temperatura).

5.3 Interpolación de Temperatura Usando RNA

La temperatura de la atmósfera dentro de la capa de la troposfera disminuye aproximadamente ~1.2 ° C por cada 100 metros de altura. Debido a esta dependencia, el parámetro de elevaciones es considerado para efectuar la interpolación con redes neuronales, de otro lado la vegetación también contribuye en la variación de la temperatura (más vegetación menos variación de temperatura).

Este método de interpolación consiste en tres pasos como se muestran en la figura siguiente (Fig. 5.3-1) usando redes neurales: Primero una previa clasificación no supervisada de los datos mediante una red SOFM (Self Organizing Features Map) para determinar las clases o regiones similares relacionadas a la temperatura y las elevaciones, Segundo extraer un numero igual de muestras más representativas (seleccionada por el método de distancia Euclidiana u otro método) de cada clase, y por último entrenar una RNA (con el algoritmo BP o L M) con todo el conjunto de clases y sus correspondientes valores de temperatura de superficie.

Éste método tiene sentido también debido a que los niveles de temperatura están influenciados por las altidudes, también por la vegetación.

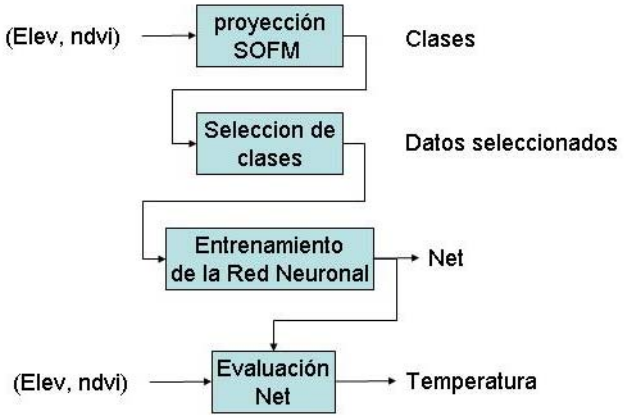

Fig. 5.3-1 Algoritmo RNA de obtención de temperatura de superficie

La imagen que se muestra en la figura siguiente (Fig. 5.3-2) es el resultado de aplicar el sistema que se muestra en la figura anterior (Fig. 5.3-1) a la misma imagen que muestra la figura de la anterior sección (Fig. 5.2-1). Debe notarse que los efectos de temperatura de los bordes de las regiones sin datos son eliminados por la clasificación previa, de modo que la imagen de temperatura en este caso representa lo que seria la temperatura en condiciones atmosféricas claras (sin nubes)

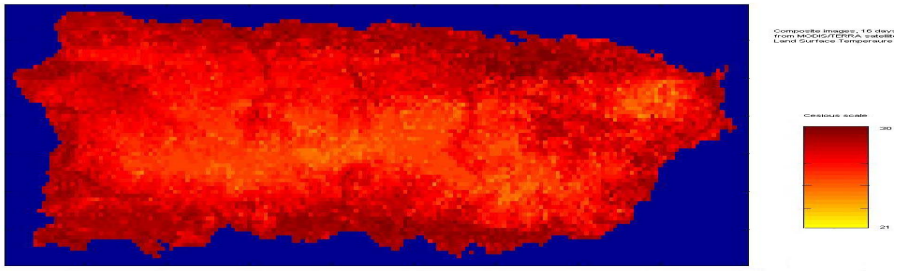

Fig. 5.3-2 Temperatura (°C) de superficie, Nov 2003

Esta imagen reflejaría entonces un día sin la presencia de nubes en el área de Puerto Rico.

5.4 Interpolación de Observaciones de Radar

Los datos observados por el Radar están disponibles y pueden ser obtenidas vía FTP desde los servidores de la NCDC con varios niveles de procesamiento (nivel I, II y III), uno de los productos del radar NEXRAD es la medida de lluvia acumulada el cual es obtenida convirtiendo las observaciones de reflectancia a lluvia acumulada mediante la siguiente ecuación:

$$dB = -6.125 + 0.125\rho$$
$$r = 10^{\log(dB)/10}$$

donde ρ es la reflectancia y r es la lluvia acumulada, esta operación se considera en la segunda fase del proceso del sistema que se muestra en la figura (Fig. 5.4-1). La resolución de esta es aproximadamente de $5Km$ [26] y para este trabajo se requiere llevar a $1Km$ de resolución lo cual es realizada usando el método de interpolación Kriging.

Para una cantidad numerosa de datos conlleva a trabajos tediosos de procesamiento si ésta es realizada de una en una puesto que se tienen observaciones por cada hora, así que esto puede resolverse realizando un Script (programa en lenguaje del sistema operativo) en Uninx ya que los decodificadores de los datos también son compilados en C/C++ y otras utilidades en FORTRAN considerado en la segunda fase del proceso del sistema en mención.

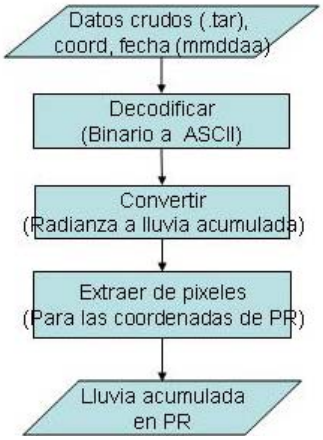

Fig. 5.4-1 Procesamiento de datos de radar

La imagen siguiente (Fig. 5.4-2) es el resultado de aplicar el sistema anterior a las observaciones de radar, la barra de colores muestra la lluvia acumulada (en pulgadas) durante 24*h*.

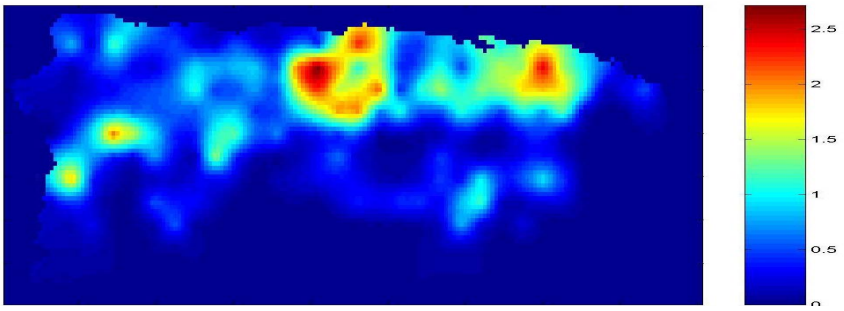

Fig. 5.4-2 Observaciones de Radar, Nov 1, 2003

La gran ventaja de obtener información de lluvia de radar es su cobertura, y su desventaja esta en la pobre resolución espacial, ademas de los problemas mencionados antes (sección 3.9)

5.5 Interpolación de Observaciones de Raingauge

Actualmente la USGS provee datos de lluvia para US y Puesto Rico, con más de 120 estaciones provistas de *Raingauge* (pluviómetro, medidor de lluvia acumulada por unidad de tiempo), las cuales pueden ser interpoladas para obtener imágenes que reflejan estimaciones de lluvia, a intervalos de 15 *min*, una muestra como resultado de interpolación de esos datos se muestra en la siguiente figura, lluvia acumulada de 5 horas, 5 nov del 2003.

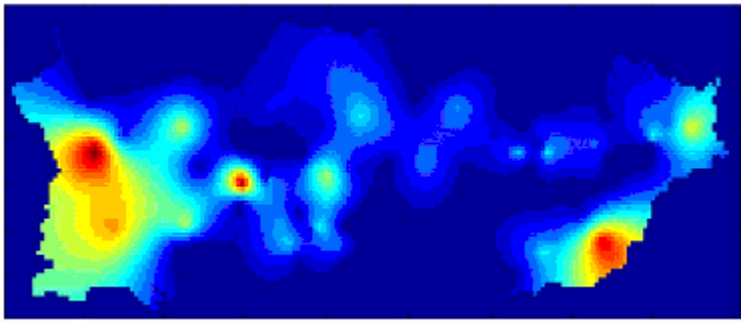

Fig. 5.5-1 Observaciones de *Raingauge*, Nov 5 2003

Una logica pregunata surge ¿Qué similitud hay con las imágenes obtenidas de lluvia de ambas fuentes?, debido a la escasa información provista por las estaciones de raingauge no fue posible obtener las imágenes correspondientes a las mismas fechas y el mismo número de horas acumuladas. Comparar los datos de las dos fuentes sera uno de los siguientes trabajos por seguir. Por tanto, la informació en cuanto a observaciones de raingauge se tomara por esta vez solo como posibilidad. De hecho sería mejor tener información de raingauge sobre toda la región debido a su mejor presición comparado a la del radar.

5.6 Interpolación Espacial Humedad de Suelo

Para interpolar la humead de suelo espacialmente se requiere construir una función de transferencia por cada región, se considerará la función de transferencia (FT) del Modelo "Modelo de Función de Transferencia Autoregresivo" dado por:

$$H_t = \frac{\alpha_{11} + \alpha_{12} B}{1 - \delta_1 B} H_{t-1} + \frac{1}{1 - \delta_2 B} R_t^{\lambda} \quad \lambda > 0$$

En esta sección se analizan con datos de dos estaciones experimentales (Guilarte y Bosque Seco), con la finalidad de construir las funciones de transferencia (determinación de los coeficientes de la función), por simplicidad se llamará coeficientes de la FT al conjunto ordenado $[\alpha_{11}, \alpha_{12}, \delta_1, \delta_2, \lambda]$. El modelo es inicializado con una serie una serie de tiempo de humedad de suelo. La figura siguiente (

Fig. 5.6-1) muestra una aproximación (linea suave) de la humedad de suelo mediante la función de transferencia, las series de tiempo de humedad de suelo consideradas con longitud de 700 (horas de datos) provistas por la estación Maricao (ver características de las estaciones en la Tabla 4.3-1).

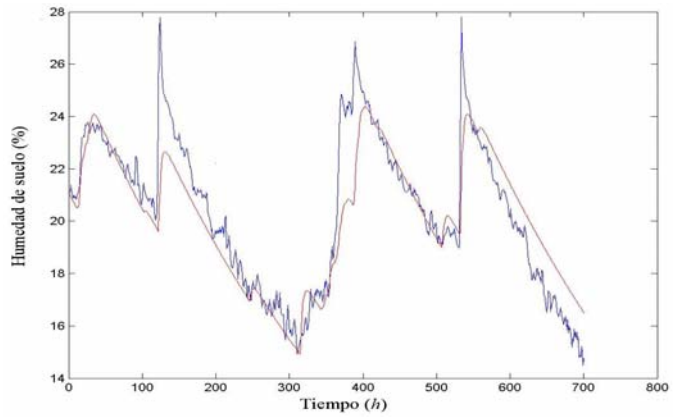

Fig. 5.6-1 Aproximación mediante FT, Maricao, Dic 2003

Los coeficientes de la FT obtenidos para este caso es la correspondiente a la de longitud 700*h* que se muestran en la tabla (Tabla 5.6-1).

La figura siguiente (Fig. 5.6-2) muestra la aproximación (línea suave) a la serie de tiempo de humedad de longitud 1100 (horas de datos), tomadas de la estación Bosque Seco (ver características del en la Tabla 4.3-1).

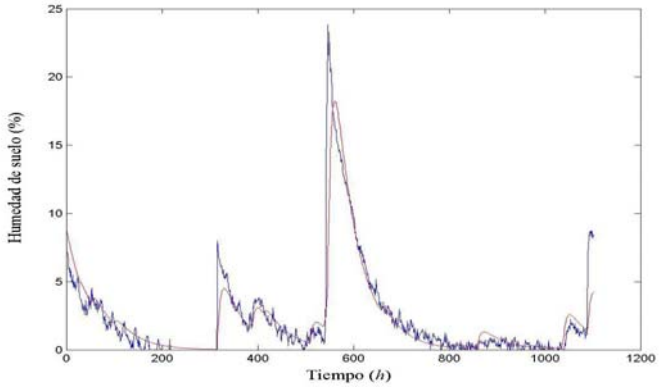

Fig. 5.6-2 Aproximación mediante FT, Bosque Seco, Nov-Dic. 2003

Los coeficientes de la FT para este último caso se muestra en la tabla siguiente (Tabla 5.6-1), que son las correspondientes al de longitud 1100*h*.

Tabla 5.6-1 Coeficientes óptimos de la FT

Estación	***horas***	α_{11}	α_{12}	δ_1	δ_2	λ
Maricao	200	-7.5997	0.3373	0.1000	-0.7013	0.5
(ID: 15)	500	-7.3944	0.3416	0.1026	-0.7008	0.5
	700	-7.4694	0.3397	0.1017	-0.7014	0.5
	600	-9.6754	-2.9280	0.4526	-0.4368	0.4457
Bosque	300	-5.0624	-1.1770	0.5312	-0.7390	0.4410
Seco	800	-5.0587	-1.1785	0.5310	-0.7588	0.4838
(ID: 2067)	900	-5.7245	-1.5721	0.5396	-0.8711	0.6748
	1100	-6.3998	-1.6196	0.4817	-0.6160	0.5

	1200	-5.9845	-1.4179	0.4817	-0.5518	0.4920

Esta tabla muestra los estimados de la FT mediante un procedimiento de optimización para diferentes longitudes de datos (dados en horas) para cada una de las dos estaciones mencionadas antes.

Por otro lado, las estimaciones en promedio mensual resultan buenas, por ejemplo el promedio mensual de humedad (en %) para la estación Bosque Seco es 7.4181, y con las estimaciones resulta 7.4076, aproximación similar ocurre con la estación de Maricao lo cual puede ser observada en las gráficas anteriores, sin embargo para periodos de menos de 24 horas puede esperarse una aproximación poco confiable debido al uso del filtro mencionado en el capítulo anterior (3.3).

Para obtener una estimación de humedad de suelo en todo un área o región (interpolación espacial), esta es dividida en subregiones, en las que la superficie tienen similar cubierta de vegetación de modo que la humedad en puntos con similares características es estimada con una sola función de transferencia y bajo la suposición de que las clases de vegetación están relacionadas con las clases de suelos [11] se consideran las regiones determinadas por las clases de vegetación. Optimizada la FT para un punto de la superficie (con datos de dada estación), la misma FT es usada para estimar la humedad en otros puntos dentro de la región al que pertenece dicho punto, los datos experimentales muestran que en puntos cercanos en condiciones similares de superficie (con la misma vegetación y el mismo tipo de suelo) las funciones (series de tiempo) de humedad de suelo tienen similares características tales como las pendientes de las caídas con los periodos de sequías y los saltos de niveles de humedad con los eventos de lluvia, en la figura siguiente (Fig.

5.6-3) se muestran tres series de tiempo correspondientes a la humedad (moist1, moist2 y moist3) obtenidas mediante tres sensores colocados aproximadamente a una distancia de 10*m* entre sí, con intervalos de tiempo de 15 *min*. durante ~9 días. Se puede notar las diferencias de niveles de humedad entre las tres series, y la relación con los eventos de lluvia.

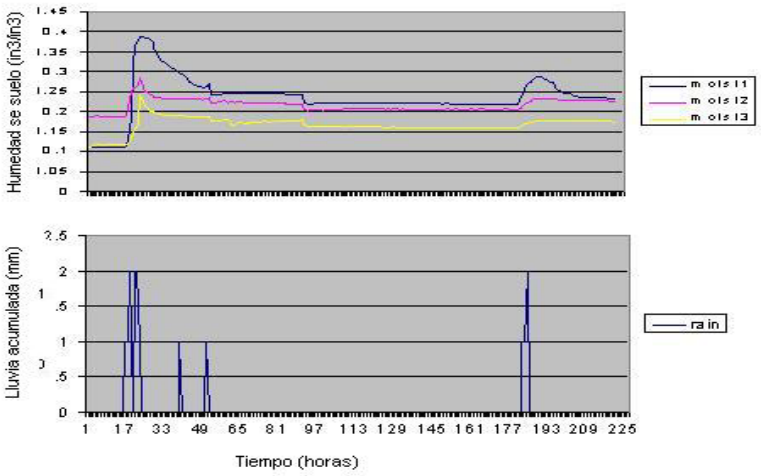

Fig. 5.6-3 Tres niveles de humedad de suelo (moist1, moist2, moist3), y Lluvia

La serie de tiempo que se muestra en la figura de arriba (Fig. 5.6-3) es uno de los resultados del trabajo de campo llevados a cabo en el área oeste de de Puerto Rico desde Octubre hasta Diciembre del 2003.

Las diferencias en los niveles de humedad son atribuidas a la forma ondulada del terreno. Esto hace pensar que en modelos de mayor resolución (<1*Km*) sobre la necesidad de considerar las elevaciones y pendientes (características geométricas del terreno), de estas observaciones (gráficas) puede suponerse que la relación entre las distintas series de tiempo de una misma región es lineal

$$moist_i = a_i * amoist_1 + b_i, \; i = 1,2$$

donde las constantes a_i y b_i estarían en función de las elevaciones y pendientes.

En seguida se muestran los coeficientes de la FT obtenidos optimizándolo con observaciones del raingauge y los datos de radar para dos puntos correspondientes a las estaciones de *Bosque Seco* y *Maricao*, así como los resultados de su interpolación espacial. En cada gráfica que se muestran en las figuras siguientes las aproximaciones son las líneas de puntos de color verde, y las líneas continuas de color azul son las observaciones de Raingauge.

- ***Estimación de la humedad en la región asociado a la estación Bosque Seco***
    - Usando los de datos de Raingauge, la aproximación con la FT se muestra en la figura (Fig. 5.6-4), y los coeficientes obtenidos fueron: [-5.0783 -1.1667 0.5303 -0.5674 0.4364]

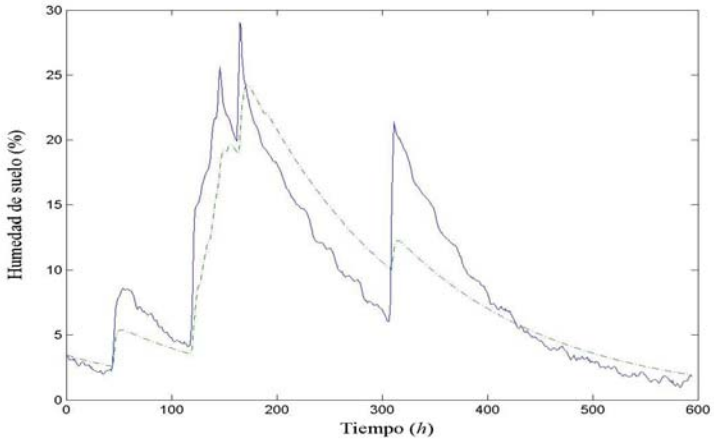

Fig. 5.6-4 Aproximación con datos de Raingauge, Dic. 2003

- Usando los de datos de Radar, la aproximación con la FT se muestra en la figura (Fig. 5.6-5), y los coeficientes obtenidos fueron: [-7.2944 -1.9586 0.4687 0.3366 0.3181]

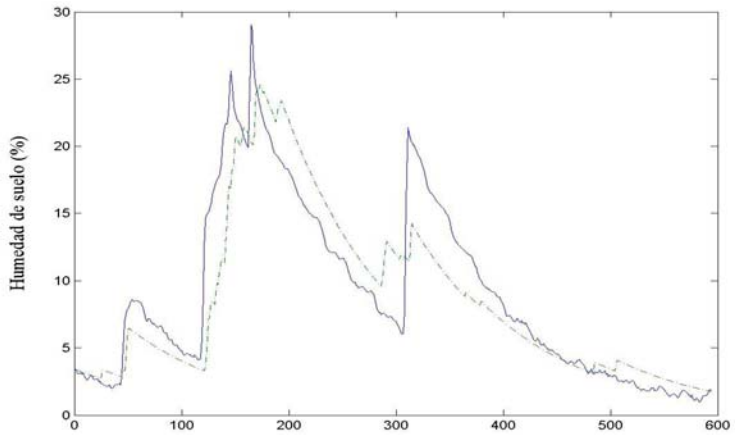

Fig. 5.6-5 Aproximación con datos de Radar, Dic. 2003

El siguiente (Fig. 5.6-6) es el resultado de aplicar la función de transferencia a los datos de lluvia y visualizada como promedio de todo el periodo modelado (~1 mes) correspondiente a la región a la que pertenece la región *Bosque Seco*.

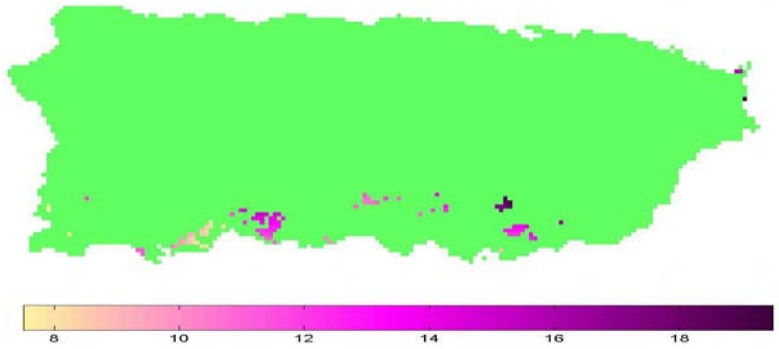

Fig. 5.6-6 Humead en la región asociado a *Bosque Seco*

- ***Estimación de la humedad en la región asociado con la estación Maricao***
    - Para caso de datos de Raingauge la aproximación con la FT se muestra en la figura (Fig. 5.6-7), y los coeficientes obtenidos fueron: [-7.2358 0.3528 0.1033 -0.5302 0.3544]

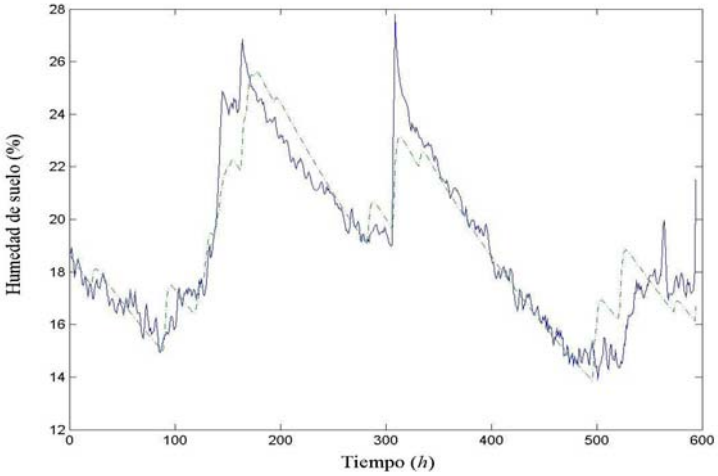

Fig. 5.6-7 Aproximación con datos de Raingauge, Dic 2003

- Para caso de datos de Radar (lluvia acumulada), la aproximación con la FT se muestra en la figura (Fig. 5.6-8), y los coeficientes obtenidos fueron: [-3.6362 -0.0080 -0.1279 -0.9953 0.5604]

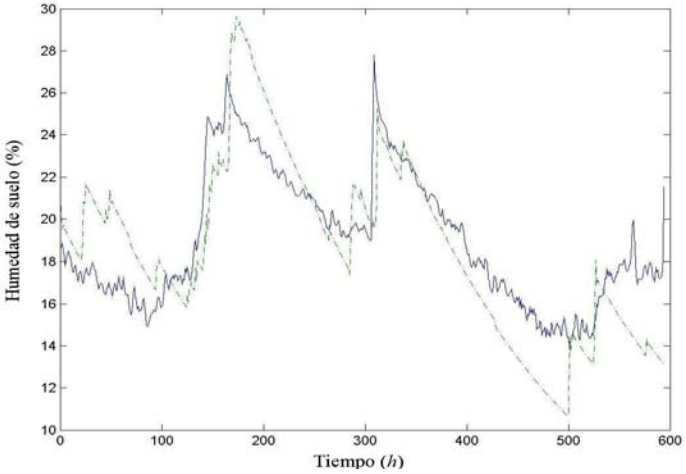

Fig. 5.6-8 Aproximación con datos de Radar, Dic. 2003

El siguiente (Fig. 5.6-9) es el resultado de aplicar la función de transferencia a los datos de lluvia y visualizada como promedio de todo el periodo modelado (~1 mes) correspondiente a la región a la que pertenece la región Maricao.

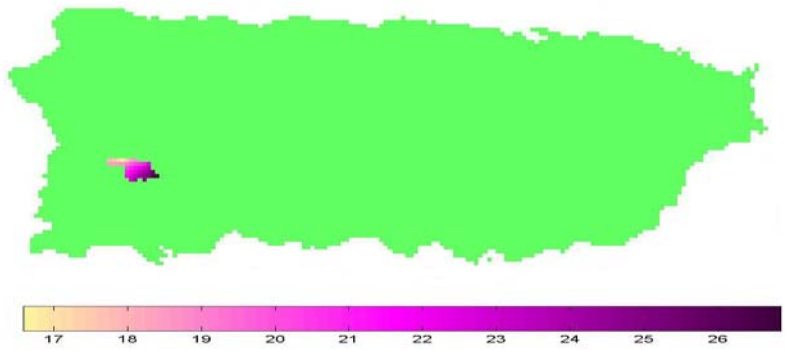

Fig. 5.6-9 Humead en la región asociado a *Maricao*

La figura siguiente (Fig. 5.6-10) muestra el algoritmo para obtener las estimaciones de humedad de suelo por región, el cual consiste en los siguientes procesos: Primero, determinar la región a la que g (geolocalización de la región) pertenece haciendo uso el mapa de las clases de vegetación, esto es determinar $i / g \in V_i$ (ver notaciones hechas en la sección 3.4), y se denotará los datos de lluvia de radar en esta región por R_i.

El segundo proceso es la optimización de los coeficientes de la FT con datos de lluvia de radar y los datos de humedad de suelo de la estación experimental de localización g.

Por último la FT es evaluado con los datos de lluvia obtenidas de radar para la región R_i.

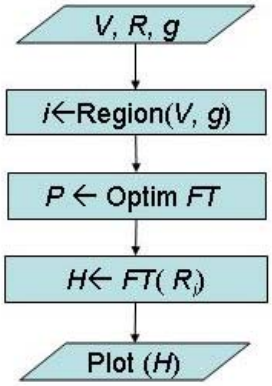

Fig. 5.6-10 Algoritmo de estimación e interpolación humedad de suelo

6 Resutados

En la sección anterior se obtuvo los mapas de temperaturas y de índices de vegetación (NDVI) y humedad de suelo, para los cuales se dieron los algoritmos de estimación e interpolación.

Con respecto al algoritmo de estimación de humedad, el modelo "Modelo de Funcion de Tranferencia Auroregresivo" propuesto muesta hasta aquí ser consistente, se obtuvieron buenas estimaciones en promedios cada 24 horas como se muestran en las figuras siguientes, las gráficas de líneas discontinuas representan la aproximación mediante el Modelo usando datos de *raingauge* y datos de *radar*.

- Resultados con datos de la estación experimental Maricao
    - Aproximación mediante TF usando datos de lluvia de la estación

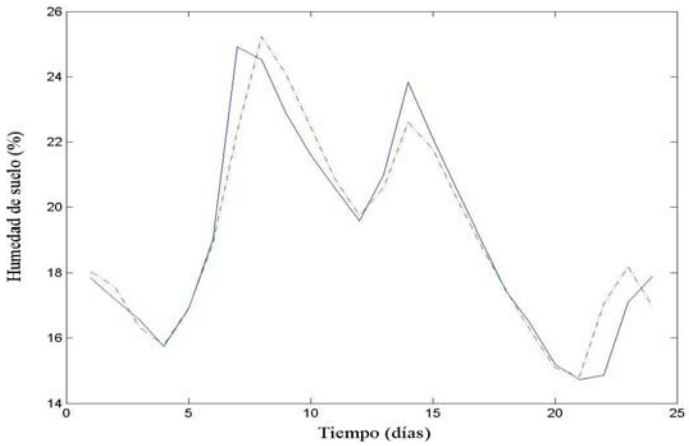

Fig. 5.6-1 Promedio aproximado cada 24h, Raingauge

- Aproximación mediante TF usando datos de lluvia de Radar

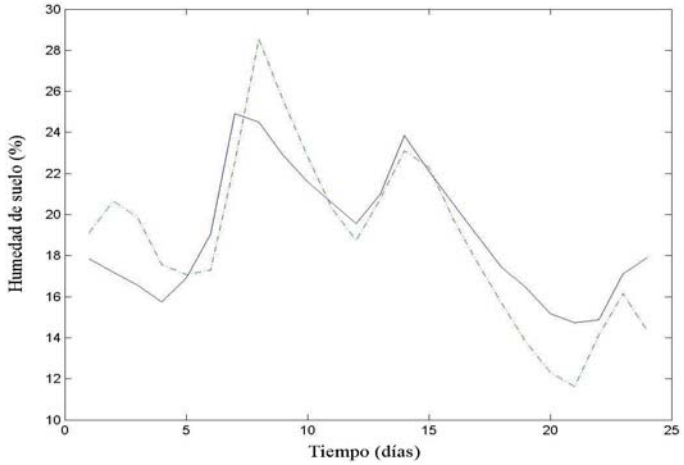

Fig. 5.6-2 Promedio aproximado cada 24h, Radar

- Resultados con datos de la estación experimental Bosque Seco
  - Aproximación mediante TF usando datos de lluvia de la estación

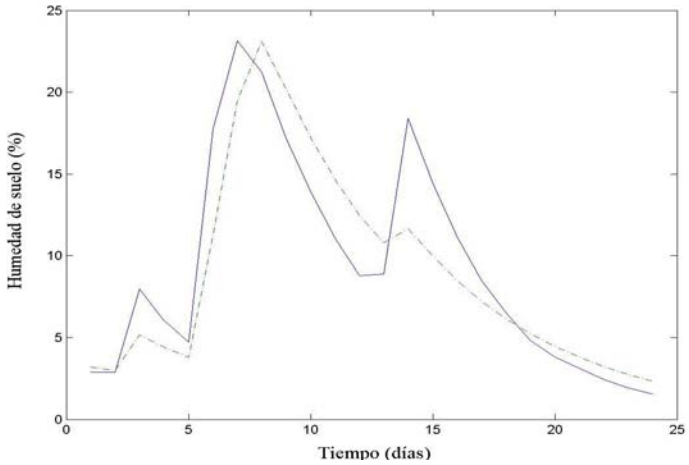

Fig. 5.6-3 Promedio aproximado cada 24h, Raingauge

- Aproximación mediante TF usando datos de lluvia de Radar

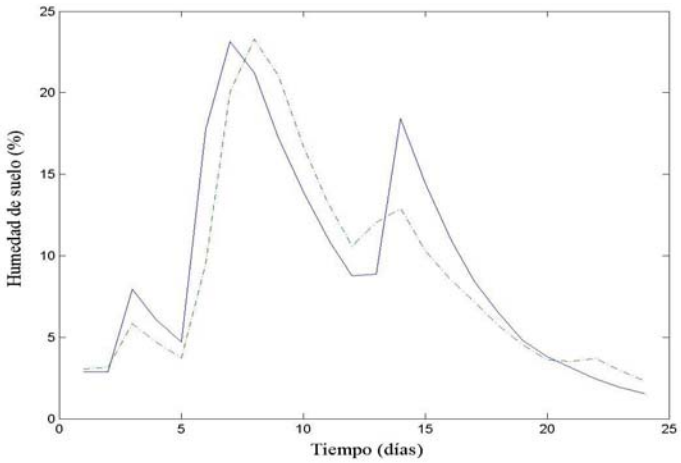

Fig. 5.6-4 Promedio aproximado cada 24h, Radar

La tabla siguiente (Tabla 5.6-1) resume los promedios obtenidos directamente de los datos de *Raingauge* y las estimaciones mediante TF usando datos de Raingauge y dataos de Radar.

Tabla 5.6-1 Promedios mensuales, Dic. 2003

Estación experimental	***Raingauge***	***TF Raingauge***	***TF Radar***
Bosque Seco	9.0560	8.7117	8.6389
Maricao	19.7000	19.6799	19.3581

7 Conclusiones

El modelo propuesto de estimación de humedad de suelo da la posibilidad de monitorear regiones donde se disponen datos de lluvia, ya sea de *Radar* o *Raingauge*.

La dependencia de una sola variable de entrada hace que el modelo sea económico.

Es necesario tener tantas estaciones experimentales como regiones se determinen dentro de un área.

La funcion de transferencia esta dada especialmete para parámetros de entrada en intervalos cortos, si como intervalos de 30 min, o 1hora.

De los datos obtenidos con trabajo de campo puede verse que la humedad de suelo cambia incluso en pequeñas áreas (con dos sensores ubicados a 10m entre sí se pueden obtener 5 diferentes resultados), esto se debe a diversos factores: altitud, elevación (pendiente), tipo de vegetación, tipo de suelo, entre otros.

Existe la posibilidad de estimar la temperatura de la superficie incluso en presencia de nubes, los datos de las fronteras de las regiones nubladas pueden dar información para la interpolación, y el método de Kriging demuestra ser buen interpolador de temperatura.

La clsificación de la vegetación puede ser usada como estrategia para la interpolacion de NDVI.

Para la implementación del los algoritmos se usó el software MATLAB (lenguaje funcional) del cual puede decirse que aún es una herramienta en la actualidad fácil de programar, solo en pocos existen la necesidad de implementación

en otros lenguajes como el C/C++ y FORTRAN debido a que las rutinas para la preparación de datos especialmente la de radar (una imagen por cada hora) que requerían demasiado tiempo y fue necesario crear rutinas en estos lenguajes, puesto que no se disponían de las componentes adicionales de MATLAB para obtener esto ni el motor para ejecutar sus rutinas desde otros programas, los ejecutables creados fueron usados inclusive.

8 Trabajo Futuro

Dado el modelo como alternativa de solución del problema de humedad de suelo, para su aplicación requieren diversos trabajos de largo plazo realizados:

1) En la determinación de las regiones de muestreo, probar que las clases de vegetación pueden aproximar las clases de suelo
2) La instalación de estaciones de recolección de datos por regiones.
3) Analizar los errores de estimación.
4) Ver la manera de estimar humedades de suelo a profundidades mayores que $10cm$.
5) Explorar otros métodos para conseguir los coeficientes óptimos de la función de transferencia.
6) Analizar cómo influirían en el modelo las condiciones atmosféricas extremas: demasiada lluvia, demasiado viento, etc.
7) Probar el modelo con datos de satélite.
8) Cuantificar los errores del modelo dado los errores de radar
9) Ver si las variables de entrada pueden estar en intervalos de tiempo más largas como observaciones en intervalos de 1 día.

9 Bibliografía

[1]. George E P. Box, Gaily M Jenkins; “TIME SERIES ANALISYS Forecasting and Control”; *HOLDEN-DAY*, 1976.

[2]. William W., S. Weir; “TIME SERIES ANALISYS, University and Multivariate Methods”; *Addison-Wesley Publishing Company*, 1990.

[3]. Peter J. Brock well, Richard A. Davis; “Introduction to Time Series and Forecasting”; *Springer-Vela New York*, Second Edition, 2002.

[4]. James A. Cardio; “DISCRETE TIME SYSTEMS, Introduction with Interdisciplinary Applications”; *Prentice- Hall*, 1073.

[5]. Hans Wackernagel; “MULTIVARIATE GEOSTATISTICS, An introduction with Applications”; *Springer-Vela Berling Heidelberg*, Third Edition, 2003.

[6]. John A. Richards, Xiuping Jia; “Remote Sensing Digital Image Analysis”; *Springer*, Third Ed, 1986.

[7]. Zhengming Wang; “Modis Land-Surface Temperature: Algorithm Theoretical Basis Document”; *www.modis.gsfc.nasa.gov*, Version 3.3, 1999.

[8]. Alfredo Huete, Chris Justice; “Modis Vegetation Index: Algorithm Theoretical Basis Document”; *www.modis.gsfc.nasa.gov*, Version 3, 1999.

[9]. Richard G. Allen, Luis S. Pereira Dirk Raes, Martin Smith; “CROP EVAPOTRANSPIRATION Guidelines for Computing crop Water Requirements”. *FAO*, 1998.

[10]. E. H. Helmer; “Mapping the Forest Type and Land Cover of Puerto Rico, a Component of the Caribbean Biodiversity Hotspot”; *Caribbean Journal of Science*, Vol. 38, No. 3-4, 165-138, 2002.

[11]. John J. Ewel, Jacob L. Whitmore; “The Ecological Life Zones of Puerto Rico and the U. S. Virgin Islands”; Institute of Tropical Forestry,1973.

[12]. Martin T. Hagan, Howard B. Demuth, Mark Beale; “Neural Network Design”; *PWS Pushing Company*, 19996.

[13]. Hongli Jiang* and W. R. Cotton; “Estimation of Soil Moisture Using An Artificial Neural Network”; *Department of Atmospheric Science of Colorado State University Fort Collins*, 2002.

[14]. Venkata Atluri, Chih-Cheng Hung, Tomy L. Colema; “An Artifitial Neural Network for Classifying and Predicting Soil Moisture and Temperature Using Levenberg-Marquardt Algorithm”; *Southeastcon* '99, *Proceedings*, *IEEE*, 10-13, 1999.

[15]. Wen Jin_Wei, Zhao Jia_Li, Luo Si_Wei and Zhen; “The Improvements of BP Neural Network Learning Algorithm”; *Proccedings of ICPS 2000*. 1647-1649, 2000

[16]. Jin Huang, Huug van den Dol and Konstantine P. georgakakos; “Analysis of Model-Calculated Soil Moisture over the United States (1931-93) and Application to Long-Range Temperature Forecasts”; *Journal of Climate,* Vol 9, No 6,1350-1362, 1996.

[17]. Huug Van den Dool, Jin Huang, Yun Fan; "Performance and Analisys of the Constructec Analogue Method Applied to U. S. Soil Moisture Over 1981-2001"; *Journal of Geophysical Research*, Vol 108, No 16, 2003.

[18]. T. J. Jackson; "Measuring Surface Soil Moisture Using Passive Microwave Remote Sensing"; *Hydrol. Proc*., 7, 139-152, 1993.

[19]. T. J. Jackson, David E. Le Vine; "Mapping Soil Moisture Using an Aircraft-based Passive Microwave Instrument: Algorithm and example"; *Journal of Hydrology*, Vol 184, 85-99, 1999.

[20]. Wang, J. R. and Schmugge, T. J.; "An Empirical Model for the Complex Dielectric Permittivity of Soils as a Function of Water Content"; *IEEE Trans. Geosci. Remote Sens.*, GE-18, 288-295, 1980.

[21]. Myron C. Dbson, Fawwaz T. Ulaby, Martti T. Hallikainen, Mohamed A. El-Rayes; "Microwave Dielectric Behavior of Wet Soil--PratII: Dielectric Mixing Models"; *IEEE Trans. Geosci. Remote Sens.*, GE-18, 35-46, 1985.

[22]. Pascale C. Dubois, Jakob van Zyl; "An empirical Soil Moisture Estimation Algorithm Using Imaging Radar"; *IEEE Trans. Geosc. and Remote Sensing*, Vol. 34, No. 2, 1573-1575, 1994.

[23]. T. J. Jackson, T. J.Schmugge; "Vegetation effects on the microwave emission of soils"; *Remote Sens. Enviro*n., 36, 203-212, 1991

[24]. B. J. Choudhury, T. J. Schmugge, Newton, R. W., and Chang, A. T. C; " Effect of surface roughness on microwave emission of soils"; *J. Geophys. Res.*, 84, 5699-5706, 1979.

[25]. Pascale C. Dubois, Jakob van Zyl, Ted Egngman; "Measuring Soil Moisture with Imaging Radars", *Trans. Geosc. and Remote Sensing*, Vol. 33, No. 4, 915-1340, 1995.

[26]. Richard F. Fulton; "WSR-88D Polar to HRP Mapping"; Hydrologic Research Laboratory- National Water Service-Silver Spring Maryland, 1998.

[27]. Gerard E. Klazura, Jessica M. Thomale, D. Scott Kelly, Paul Jendrowski; "A comparison of WsR-88D Rain Estimates UIT Gauge Measurements for high and Low Reflectivity Gradient Precipitation Events"; *National Weather Service, Honolulu, Hawaii*, 1999.

[28]. Steven M. Hunter; "WSR-88D Radar Rainfall Estimation: Capabilities, Limitations and Potential Improvements"; *National Weather Digest*, v. 20, p. 26-38, 1996.

[29]. Richard J. Doviak, Dušan S. Zrnicć; "Doppler radar and Weather Observations"; *Academic Press*, 1993.

[30]. G. V. Rekaitis, A. Ravindran, K. M. Ragsdell; "Engineering Optimization: Methods and Aplications"; *John Wesley and Sons*, 1983.

[31]. Mitsuo Gen, Runwei Cheng; "Genetic Algorithms and Engineering Design"; *John Wesley and Sons*, 1997.

[32]. Mitsuo Gen, Runwei Cheng; "Genetic Algorithms and Engineering Optimization"; *John Wesley and Sons*, 2000.

[33]. Fred Brauer, John Nohel; "INTRODUCTION TO THE DIFFERENTIAL EQUATIONS, with Applications"; *Harper & Row*, 1986.

[34]. N. Finizio, G. Lades; "ORDINARY DIFFERENTIAL EQUATIONS, with Modern Applications"; *Wadsworth*, 1989.

[35]. James A. Cadzow; "DISCRETE TIME SYSTEMS, Introduction with Interdisciplinary Applications"; Prentice-Hall, 1973.

[36]. Robert G. Bartle; "The Elements of Real Analysis"; 1972.

[37]. Nazario D. Ramirez Beltran; Thesis: "Multilinear Intervention Modeling with an Application to Time Series Forecasting", 1988.

10 Apéndice

A Métodos e Instrumentación de Medición de Humedad de Suelo

Se tienen en la actualidad métodos prácticos y automatizados para el propósito de medición de humedad de suelo.

A.1 Métodos

A.1.1 Método Gravimétrico

Es un método estándar y muy conocido de calcular el contenido volumétrico de una muestra de suelo, este método consiste en secar en horno a 105° C y relacionar los cambios de masa del volumen de la muestra:

$$\theta = \frac{W_w}{W_d}\frac{\rho_b}{\rho_w}$$

donde θ es la fracción del contenido volumétrico de humedad de suelo, W_w es el volumen de agua contenido en la muestra, W_d es el volumen de la muestra seca, ρ_b es la densidad del suelo bruto, ρ_w es la densidad el agua.

A.1.2 Método de Dispersión de Neutrones

Es una manera directa de determinar el contenido de humedad de suelo, consiste en la emisión de neutrones por una fuente radiactiva sumergida en el suelo, y detectar

la pérdida de energía, por ejemplo la perdida de energía en colisión con el hidrogeno es mayor que con otros componentes del suelo (por tener un peso atómico mas bajo).

A.1.3 Método de Atenuación de Rayos Gamma

Es una técnica de radiación par capas delgadas de suelo (1-2 cm), este método asume que el esparcimiento y la absorción de los rayos gamma están relacionados a la densidad de la materia, ya que la densidad cambia con la humedad, la humedad es determinado de este cambio de densidad.

La ventaja de este método es que la medición se puede realizar desde pequeñas distancias verticales. La desventaja es que es costosa y difícil de usar y solo es usado en laboratorios para evitar danos en la salud.

A.1.4 Método de Conductividad Eléctrica

La conductividad eléctrica del suelo es una función de la humedad volumétrica de suelo, conductividad eléctrica de la mezcla suelo-agua.

A.1.5 Probetas de Conductividad Eléctrica

Las probetas de conductividad eléctrica para medir la humedad de suelo consisten en varios (2 y 4) electrodos. Se adoptan ciertos materiales para evitar incertidumbres debido a la porosidad de los suelos.

Su principal ventaja es que es barato y fácil de usar, y que puede medir volúmenes de suelo relativamente grandes.

A.2 Bloques de Resistencia Eléctrica

Este es un método indirecto estriba en el equilibrio del potencial del agua del suelo en el bloque

A.3 Inducción Electromagnética

En este método el suelo es sometido a los campos magnéticos primarios y secundarios generados por la bobina de un transmisor colocado en la superficie del suelo, la razón de esos dos campos en el receptor provee la estimación de conductividad eléctrica del suelo.

La mayor ventaja de este es que no necesita insertar al suelo, es rápido y de fácil manejo. Una desventaja es que no es fácil determinar la humedad a profundidades particulares o específicas.

A.4 Método de Tensiómetro

Este método de medida indirecta, mide la capacidad de retención de una columna de agua del suelo, las cuales deben ser convertidas a medidas de humedad de suelo.

La ventaja de este método es su bajo costo. La desventaja es que ellos pueden fácilmente averiarse.

A.5 Método Hygrometric

La relación entre el contenido de humedad en materiales porosos y la humedad relativa de la atmósfera circundante es bien conocida. Muchos sensores relativamente simples para medir humedad relativa han sido diseñados, básicamente estos pueden ser clasificados en 6 grupos: resistencia eléctrica, capacitancia, absorción piezoeléctrica, absorción infrarroja y transmisión, punto de condensación, y psicometría.

La ventaja es simplemente la simplicidad del aparato y su bajo costo. La desventaja es que los sensores se deterioran en contacto con el suelo, y requiere calibración para cada material en prueba.

A.5.1 Método Dieléctrico

La constante dieléctrica (conocida también con permisividad o capacidad inductiva) ε, es una medida de cuan polarizable es un material cuando es sometido a un campo magnético.

El suelo es compuesto por aire partículas de suelo y agua; debido a la gran variabilidad de la constante dieléctrica de ~80 (agua) a ~3 (suelo seco), esta es usada para determinar la humead del suelo.

A.5.2 Probetas de Capacitancia

Cuando un potencial de voltaje es colocado a través de platos.

A.6 Probetas TDR

En TDR (Time Domain Reflectometry) una onda electromagnética de longitud L es propaganda a en el material en prueba, el sensor es capaz de detectar el tiempo t de viaje de la honda, este tiempo t dependerá de la constante dieléctrica del material, el tiempo crece con el crecimiento de la humedad.

Las ventajas de este instrumento incluyen: monitoreo de humedad de suelo por diferentes profundidades y perfiles. Las principales desventajas son que las mediciones son afectadas por regiones inmediatamente adyacentes, y la atenuación causada por la salinidad.

B Trabajo de Campo

B.1 Instrumentación

La instrumentación usada en este proyecto fueron diversos: medidores de lluvia, temperatura, humedad de suelo.

A continuación se dan algunas características de los instrumentos usados:

B.1.1 Medidor de Temperatura

Este instrumento (Fig. 5.6-1) es diseñado para medir la temperatura del aire o suelo, que conectados al *datalogger* podemos monitorear la temperatura.

Fig. 5.6-1 Sensor d e Temperatura

Para el caso de estimación de temperatura del aire es necesario el dispositivo que se muestra en la figura de la derecha (radiationshilt). Para medir la temperatura del suelo, el sensor es colocado en contacto directo con el suelo. En nuestro caso estamos interesados en profundidades cerca de la superficie, esto es de 0 a 10 cm. de profundidad.

B.1.2 Medidor de Humedad de Suelo (ECHO)

Este instrumento (Fig. 5.6-2) permite monitorear el contenido de agua por unidad de volumen del suelo, su bajo costo nos permitió tener mediciones de un número suficiente para validar el modelo que tratamos aquí.

Este sensor detecta las variaciones de voltaje para estimar la constante dieléctrica del suelo, para luego estimar el contenido volumétrico de agua del suelo, además, este instrumento es calibrado para estimar el parámetro de humedad de suelo en mas de cien tipos de suelo.

Fig. 5.6-2 Sensor de humedad de suelo

La figura muestra dos diferentes tamaños de probetas (20 cm. y 10 cm.), los cuales conectadas al *datalogger* permiten monitorear en intervalos cortos de tiempo (~ 10ms), en condiciones de hasta 50º C de temperatura.

B.1.3 Medidor de Lluvia (ECHORain)

Este instrumento nos permite medir la lluvia acumulada por unidad de tiempo

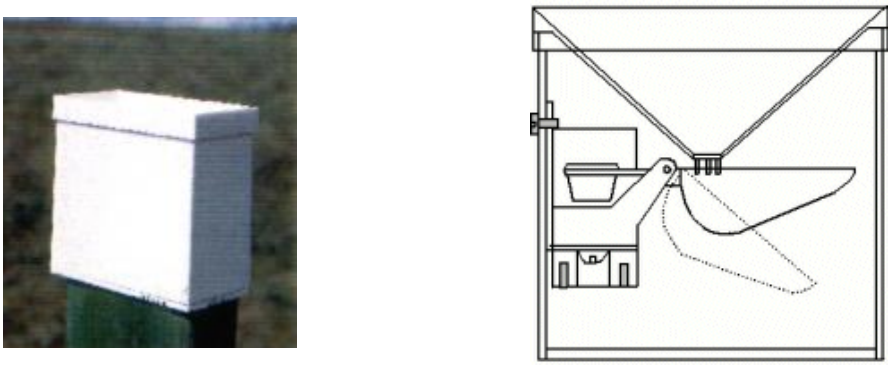

Fig. 5.6-3 Raingauge, mecanismo interno

a la derecha la figura muestra como este trabaja, la cuchara que muestra se vacía cada vez que alcanza un nivel de volumen, generalmente a cada un milímetro cúbico.

B.1.4 Data Logger (EM5)

Este instrumento posee 5 canales de entradas para colectar data provenientes de diferentes (no más de 5) instrumentos, la figura siguiente (Fig. 5.6-4) muestra el dispositivo de de almacenamiento de datos (Data Logger EM5)

Fig. 5.6-4 Datalogger

Puede verse más información a cerca de estos instrumentos en la página Web http://www.decagon.com

C Percepción Remota

En la actualidad existen cientos de satélites en el espacio observando la tierra y otros objetos en el espacio, las observaciones que se pueden obtener desde los satélites dependerá de la instrumentación a bordo, las resoluciones espectrales y espaciales han ido mejorando de modo que ahora se pueden distinguir objetos cada vez más pequeños (resolución espacial del orden de 1m). Para este proyecto se eligieron satélites los cuales proporcionan datos de observaciones de la tierra gratuitamente y algunas de bajo costo.

Existen dos modos de percepción remota: *activo* y *pasivo*, la diferencia entre ellos es la fuente de energía, Toda materia por encima del cero absoluto (0 K) emite radiación electromagnética debido a al movimiento de las partículas cargadas de sus átomos y moléculas. La percepción remota pasiva mide esta radiación natural emitida por al Tierra en la región de microondas del espectro electromagnético. En contraste la percepción remota activa, conocida como Radar, envía pulsos de radiación electromagnética y mide la cantidad la cantidad esparcida (conocida como *scattering*) que regresa en dirección del sensor definida como promedio de *scattering cross-sección* por unidad de área.

Designación de Bandas	Longitud de onda	Frecuencia
Ka	0.75-1.1	40.0-26.5
K	1.1-1.67	26.5-18.0
Ku	2.67-2.40	18.0-12.5

X	2.40-3.75	12.5-8.0
C	3.75-7.50	8.0-4.0
S	7.50-15.00	4.0-2.0
L	15.0-30.0	2.0-1.0
P	30.0-100	1.0-0.3

Tabla 5.6-1 Designación de banda de microonda

Debido a las altas diferencia de temperatura a la que se encuentran sometidas los satélites, estas tienen un tiempo limitado de vida que es aproximadamente 4 años.

C.1 TERRA/MODIS

El radiómetro MODIS (Moderate Resolution Imaging Spectroradiometer) es un instrumento del tipo *pasivo* a bordo de los satélites Terra lanzado el 8 de Dic. De 1999 y el satélite Aqua lanzado el 4 de Mayo de 2002, durante un tramo de su (~10 AM) trayectoria permite observar parte de la región del caribe.

Las especificaciones de este satélite se dan en la tabla siguiente:

Tabla 5.6-2 Bandas del sensor MODIS

Bandas: 1-2 (µm) Resolución 250m	Bandas: 3-7 (µm) Resolución: 500m	Bandas: 8-12, 13L, 13H, 26 (µm) Resolution 1 km	Bandas: 20-25, 27-36 (µm) Resolución: 1 km
Banda 1 (.620-.670)	Band 3 (.459-.479)	Band 8 (.405-.420)	Band 20 (3.660-3.840)
Banda 2 (.841-.876)	Band 4 (.545-.565)	Band 9 (.438-.448)	Band 21 (3.929-3.989)
	Band 5 (1.230-1.250)	Band 10 (.483-.493)	Band 22 (3.939-3.989)
	Band 6 (1.628-1.652)	Band 11 (.526-.536)	Band 23 (4.020-4.080)
	Band 7 (2.105-2.155)	Band 12 (.546-.556)	Band 24 (4.433-4.498)
		Band 13L (.662-.672)	Band 25 (4.482-4.549)
		Band 13H (.662-.672)	Band 27 (6.535-6.895)
		Band 14L (.673-.683)	Band 28 (7.175-7.475)

		Band 14H (.673-.683)	Band 29 (8.400-8.700)
		Band 15 (.743-.753)	Band 30 (9.580-9.880)
		Band 16 (.862-.877)	Band 31 (10.780-11.280)
		Band 17 (.890-.920)	Band 32 (11.770-12.270)
		Band 18 (.931-.941)	Band 33 (13.185-13.485)
		Band 19 (.915-.965)	Band 34 (13.485-13.785)
		Band 26 (1.360-1.390)	Band 35 (13.785-14.085)
			Band 36 (14.085-14.385)

Este satélite orbita la tierra en forma sincronizada (sun synchronous) a una altura de 708Km observando casi toda la tierra en un día.

Fig. 5.6-5 MODIS

C.2 NOAA/AVHRR

El primer satélite NOAA (National Oceanic and Atmospheric Administration) de los Estados Unidos fue lanzado el año 1979 (NOAA-6). Hasta la fecha han sido lanzados 7 satélites NOAA. El instrumento AVHRR (Advanced Very High Resolution Radiometer) a bordo del satélite NOAA ha ido cambiando hasta la

actualidad cada 4 años aproximadamente. Las características del último satélite (NOAA-17) lanzado en junio del 2002 se muestran en la siguiente tabla:

Tabla 5.6-3 Bandas del sensor AVHRR

Bandas:1-5 *Resolución 1.1Km*	*especificaciones*
Band 1: 0.52-0.68	Rojo
Band 2: 0.725-1.10	Infra Rojo
Band 3: 3.35-3.93	Infra Rojo Medio
Band 4: 10.30-11.30	Termal
Band 5: 11.50-12.50	Termal

La orbita de este satélite (Fig. 5.6-6) es sincronizado (sun synchronous) orbitando 14 veces al día a una distancia de 833Km de la tierra.

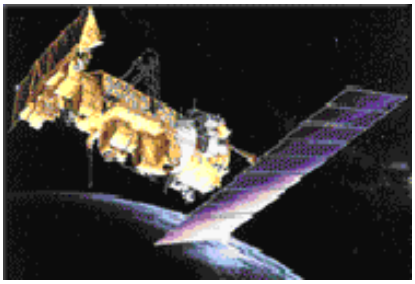

Fig. 5.6-6 NOAA

La siguiente tabla muestra algunas características del instrumento

Tabla 5.6-4 Características del satélite NOAA

Ancho del Swath	2399km
Resolución espacial	1.1km approx.
Altitud	833km
Cuantificación	10 bit
Typo de órbita	Sun synchronous
Número de órbitas por día	14.1 approx.

C.3 Radares

Un sistema de radar esta basada en el envió de señales hacia un objeto y la recepción del rebote de la misma (forma activa), que pueden estar a bordo de un satélite o avión o pueden estar colocados en un punto de la superficie de la tierra. Este sistema es usado en diferentes campos, tales como en la milicia, aeropuertos, investigación científica, etc. Su forma de trabajo hace costoso su operación, y por lo general son comerciales. La ventaja con estos tipos de sensores es que pueden operar a cualquier hora del día.

C.4 RADARSAT/SAR

Uno de los satélites con instrumento a bordo SAR (Synthetic Aperture Radar) es el RADARSAT el cual es administrado por la agencia espacial canadiense (CSA), fuel lanzada el 4 de Nov. de 1995 y clasificado como instrumento activo.

Fig. 5.6-7 RADARSAT

Este instrumento tiene diversas formas de operación, más información puede ser vista en http://www.rsi.ca/.