

Inteligencia Computacional en Teledetección. Control de contenido de humedad en combustible en superficie terrestre mediante imagen satélite para prevención de incendios

Tesis Doctoral

DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN UNIVERSIDAD DE ALCALA

Junio 2010
Luis Usero Aragonés

A mis padres Paco y Luisa

A Paco, Ramón, Marivi y Araceli (que paciencia tienen)

A mis sobrin@s Javi, Lara, Eva y Nuria

Y sobre todo a mi compañera Esther, a mi hijo César y a los que vendrán detrás

Resumen

La presente Tesis doctoral analiza los problemas relacionados con la estimación del contenido de humedad del combustible para prevenir posibles incendios y saber cómo se comportarán una vez iniciados. La Tesis describe el problema de estimación mediante técnicas de Teledetección. Hemos conseguido un avance en cuanto a investigación en nuevos modelos computacionales no convencionales para la estimación de la cantidad de agua en diferentes entornos de vegetación, uno de los mayores problemas que tiene la Teledetección, ya que se requiere un estudio previo de la zona, recogida selectiva de muestras y tratamiento de éstas para conocer su estructura, composición y contenido de humedad.

A través de estas simulaciones hemos analizado la capacidad de discriminación mediante el análisis de curvas ROC (Receiver Operating Characteristic), para posteriormente utilizar diferentes técnicas de Softcomputing (Inteligencia Computacional), en concreto, modelos neuronales y evolutivos, para validar los experimentos.

La metodología que se presenta en este trabajo para estimación de contenido de humedad en vegetación, se basa en la utilización de modelos de transferencia radiativa mediante la unión de modelos a nivel de hoja Prospect (leaf optical properties spectra) (Jacquemoud et al. 1990)) y de simulación de cubierta SAIL (Scattering by Arbitrarily Inclined Leaves).

Abstract

This doctoral thesis examines the problems related to the estimation of fuel moisture content to prevent possible fires and how to behave once they commence. The thesis describes the estimation problem using Remote Sensing techniques. We have achieved a breakthrough in terms of research into new unconventional computational models to estimate the amount of water in different environments of vegetation, one of the biggest problems of the Remote Sensing, because it requires a prior study of the area, collecting selective treatment of these samples and to understand their structure, composition and moisture content.

Through these simulations we have analyzed the capacity of discrimination by analyzing the ROC curves (Receiver Operating Characteristic), and later use different Soft Computing techniques, namely, neural and evolutionary models, to validate the experiments.

The methodology presented in this paper to estimate moisture content in vegetation is based on the use of radiative transfer models by linking models at Prospect leaf (leaf optical spectra properties) (Jacquemoud et al. 1990) and canopy model SAIL (Scattering by Arbitrarily Inclined Leaves).

Agradecimientos

Este trabajo es gracias principalmente a tres personas que han sido mi trípode de apoyo, Miguel Ángel Patricio, Ángel Arroyo y principalmente a David Riaño, sin ellos no hubiera sido posible. Espero no solo escribir aquí vuestros nombres, sino, poder ayudaros siempre.

Esta tesis es el final de una etapa, en ella han participado numerosas personas moldeando mi personalidad y carácter, no puedo olvidar citar a Richard, el mejor, a Emilio que soporta estoicamente mis numerosas tonterías, a Juantxo que siempre esta ahí, a todo Traíd por existir, a Cote por su belleza, a Eva Rubi por su valentía, a la Tuna de la Universidad de Alcalá, y a la Universidad de Alcalá que es mi única casa, y me ha enseñado lo poco que sé.

Gracias a TODOS.

Contenido

RESUMEN	5
ABSTRACT.....	6
AGRADECIMIENTOS.....	7
ÍNDICE DE FIGURAS	10
ÍNDICE DE TABLAS.....	13
CAPÍTULO 0. INTRODUCCIÓN	14
MOTIVACIÓN.....	14
OBJETIVOS	14
¿QUÉ ES SOFTCOMPUTING?	16
LAS REDES DE NEURONAS ARTIFICIALES COMO CLASIFICADORES	17
COMPUTACIÓN EVOLUTIVA	21
ESTRUCTURA DEL DOCUMENTO.....	21
CAPÍTULO 1. LA TELEDETECCIÓN Y LA ESTIMACIÓN DE HUMEDAD DE COBERTURAS DE VEGETACIÓN	22
INTRODUCCIÓN.....	22
SISTEMAS DE ADQUISICIÓN DE INFORMACIÓN	23
MÉTODOS DE ESTIMACIÓN DE CONTENIDO DE AGUA PARA PREVENCIÓN DE INCENDIOS	26
MÉTODO DE ESTIMACIÓN DIRECTA.....	27
MÉTODOS BASADOS EN TELEDETECCIÓN.....	30
CARACTERÍSTICAS ESPECTRALES DE LA VEGETACIÓN.	30
CAPÍTULO 2. EVALUACIÓN DEL RIESGO DE INCENDIO MEDIANTE CURVAS ROC CAPÍTULO.....	31
INTRODUCCIÓN.....	31
CURVAS ROC	32
MÉTODOS.....	34
ANÁLISIS DE DATOS Y CONCLUSIONES	34
AGRADECIMIENTOS.....	39

CAPÍTULO 3. ESTIMACIÓN DEL CONTENIDO DE HUMEDAD EN VEGETACIÓN MEDIANTE DE REDES NEURONALES 40

INTRODUCCIÓN.....	41
MÉTODOS.....	44
RESULTADOS	49
DISCURSIÓN	54
AGRADECIMIENTOS.....	55

CAPÍTULO 4. ESTIMACIÓN DEL CONTENIDO DE HUMEDAD EN VEGETACIÓN MEDIANTE COMPUTACIÓN EVOLUTIVA

.....	56
INTRODUCCIÓN.....	56
RESULTADOS EXPERIMENTALES.....	60
DISCUSIÓN	65
CONCLUSIONES Y FUTURAS LÍNEAS DE TRABAJO:	67
REFERENCIAS	69

Índice de figuras

Figura 0: Estructura general de una red multicapa.....	18
Figura 1: Simulaciones de la respuesta espectral de la vegetación en el espectro óptico empleando el modelo PROSPECT-SAILH. Se simulan hojas con distinto contenido de humedad. También se muestran donde se localizan las bandas MODIS (adaptado de Zarco-Tejada et al., 2003).	30
Figura 2: Banda 1 Sensor MODIS y tasa de falsos positivos.....	36
Figura 3: Banda 2 sensor MODIS y tasa de falsos positivos	36
Figura 4: Banda 3 Sensor MODIS y tasa de falsos positivos.....	37
Figura 5: Banda 4 Sensor MODIS y tasa de falsos positivos.....	37
Figura 6: Banda 5 Sensor MODIS y tasa de falsos positivos.....	38
Figura 7: Banda 6 Sensor MODIS y tasa de falsos positivos.....	38
Figura 8: Banda 7 Sensor MODIS y tasa de falsos positivos.....	39
Figura 9: Combinación de bandas de MODIS para distinción entre riesgo de incendio alto y bajo con discriminación casi perfecta, valor próximo a 1 en AUC.	39
Figura 10: Ejemplo de un espectrofotómetro. Perkin-Elmer Lambda-19. Capacidad espectral 250-2500 nm con un rango de 1,00 nm (www.grc.nasa.gov/WWW/epbranch/OpticalProps/lambda.htm).....	42
Figura 11: Ejemplo de un espectrorradiómetro de campo. Analytical Spectral Devices FieldSpec Profr, rango espectral 350-2500 nm con un rango 3-10 nm. (www.asdi.com/especificaciones de los productos-FSP.asp)	42
Figura 12: Ejemplo de sensor hiperespectral en avión. Espectrómetro de imagen visible infrarrojo de avanzado rango espectral entre 380 – 2500 nm y una ranura de 10 nm de ancho y una resolución espacial entre 4 y 20 metros dependiendo de la altitud de vuelo. (aviris.jpl.nasa.gov/)	43
Figura 13: Imagen satélite hiperespectral. Espectrómetro de imagen de resolución moderada (MODIS). 36 bandas, rango espectral entre 405 y 14385 nm con una ranura de 20 nm de ancho, dependiendo de las bandas. Resolución espacial desde 250 hasta 1000 metros dependiendo de las bandas (www.latuv.uva.es/modisql/)	43
Figura 14: Arquitectura de la red neuronal IWINAC 2005.....	46
Figura 15: Arquitectura de la red neuronal CEDI 2005.....	47
Figura 16: Estimación de EWT con los datos utilizados para aprendizaje	50
Figura 17: Estimación de EWT con los datos utilizados para validación.....	50
Figura 18: Estimación de DM en muestras frescas con los datos utilizados para el aprendizaje	50

Figura 19: Estimación de DM en muestras frescas con los datos usados para validación	50
Figura 20: Estimación de DM en muestras secas con los datos utilizados para el aprendizaje	51
Figura 21: Estimación de DM en muestras secas con los datos utilizados para validación	51
Figura 22: Estimaciones de EWT en entrenamiento	52
Figura 23: Estimaciones de EWT en validación	52
Figura 24: Estimación de EWT usando para entrenar la red todo el conjunto de muestras LOPEX pero una en cada iteración (n=245).	53
Figura 25: Estimación de EWT usando para entrenar la red el modelo PROSPECT (n=245).	53
Figura 26: Estimación de EWT con datos LOPEX usando una muestra por iteración (n=245).	54
Figura 27: Estimación de EWT y LAI con muestras generadas con PROSPECT y SAILH (n=7400).	54
Figura 28: Ejemplo de espectrofotómetro. Modelo 1100 RS, longitud de onda 335-1000 +/- 2.00 nm de precisión (http://www.lehmanscientific.com/spectro.html).	57
Figura 29: Ejemplo de imagen de sensor hiperespectral AVIRIS (Advanced Visible Imaging Spectrometer), rango espectral entre 380 y 2500 nm. con slot de 10 nm, resolución espacial de 4 a 20 m. depende de altura (http://www.higp.hawaii.edu/.../aviris_color_mosaic.gif).....	57
Figura 30: Ejemplo de imagen de sensor satélite MODIS (Moderate Resolution Imaging Spectrometer) en 36 bandas, rango espectral entre 405 y 14385 nm. , dependiendo de las bandas, resolución espacial de 250 a 1000 nm. (http://www.satellite-images.com/imagery/modis.jpg)	58
Figura 31: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento).....	61
Figura 32: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)	62
Figura 33: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento).....	62
Figura 34: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)	62
Figura 35: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento).....	63
Figura 36: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)	63

Figura 37: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento).....	63
Figura 38: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)	64
Figura 39: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento).....	64
Figura 40: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)	64
Figura 41: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento).....	65
Figura 42: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)	65

Índice de tablas

Tabla 1: Entradas PROSPECT–SAILH en la simulación de matorral y monte bajo. Índice de estructura del mesófilo de hoja (N), cantidad de humedad equivalente a nivel de hoja (Cw), biomasa seca(Cm), índice de área foliar (LAI), zenit de ángulo solar(TS). Adaptado de Yebra et al. (2008).	34
Tabla 2: Rangos de entrenamiento y validación de la red neuronal	46
Tabla 3: Rango de los parámetros de entrada del modelo PROSPECT.(N) Estructura interna de la hoja, (C _{a+b}) contenido de clorofila a + b,.....	48
Tabla 4: rango de los parámetros de entrada de la combinación de los modelos PROSPECT-SAILH, N, EWT, DM, LAI, leaf inclination density function(LIDF), y reflectancia del suelo(ρ_s).....	49
Tabla 5: Número de muestras (N) y correlación (r^2) para la estimación de EWT y DM usando IRTM (Riaño et al., 2006) y redes de neuronas artificiales.	51
Tabla 6: Parámetros utilizados en la programación genética.	61

Capítulo 0. Introducción

Motivación.

La teledetección es una ciencia reciente que permite la adquisición de imágenes recurrentes para la observación de la tierra desde el espacio. Esta tesis parte de la colaboración en el proyecto “GLOBAL ESTIMATION OF CANOPY WATER CONTENT” SUPPORTED BY NASA (NRA-03-OES-02) NASA Wilfire Hazard Research realizado entre febrero del 2004 y julio del 2005 en la Universidad de Davis en California. Este estado sufre todos los veranos violentos incendios que arrasan zonas habitadas y grandes áreas con gran valor medioambiental.

El proyecto consiste en, mediante imágenes del satélite MODIS detectar las zonas donde la humedad del terreno desciende de un umbral y se multiplica la probabilidad de incendio, con estos datos sería posible realizar un aviso para informar a los retenes de incendios.

Las técnicas de ajuste tradicionales estaban basadas en regresiones estadísticas en sus diversas variantes (Ahern et al. 1991, Ji et al. 1992). La inversión del modelo no es singular, ya que determinados valores de reflectividad pueden obtenerse a partir de distintas combinaciones de parámetros de entrada, por lo que la estimación no resulta directa (Combal et al. 2002) a la hora de conocer los valores de humedad (Fuel Moisture Content - FMC) de la zona de estudio. La inversión de los datos retrasaba ampliamente, incluso días, el conocimiento del terreno. Gracias a los conocimientos en técnicas de Inteligencia Artificial, esta tesis propone la modelización e implementación de distintas técnicas en inteligencia computacional, que supongan novedosas aproximaciones en la estimación en tiempo real de los valores de humedad de un terreno mediante Teledetección.

Objetivos

El objetivo de la presente Tesis doctoral es el de analizar los problemas relacionados con la estimación del contenido de humedad del combustible. La Tesis describirá el problema de estimación mediante técnicas de Teledetección. La presente tesis doctoral supone un avance en cuanto la investigación en nuevos modelos computacionales no convencionales para la estimación de la cantidad de agua en diferentes entornos de vegetación, uno de los mayores problemas que tiene la Teledetección, ya que se requiere un estudio previo de la zona, recogida selectiva de muestras y tratamiento de éstas para conocer su estructura y a partir de estos datos, poder conocer su contenido de humedad con imágenes satélite.

Cuando aplicamos nuestra primera red neuronal de base radial en 2005, observamos que la inteligencia computacional podría cubrir las carencias que mostraba la teledetección en el tratamiento de la información.

La metodología que se presenta en este trabajo para estimación de contenido de humedad en vegetación, se basa en la utilización de modelos de transferencia radiativa mediante la unión de modelos a nivel de hoja Prospect (leaf optical properties spectra) (Jacquemoud et al. 1990) y de simulación de cubierta SAIL (Scattering by Arbitrarily Inclined Leaves). Este último modelo nace a partir de un modelo anterior de Suits (Suits, 1972), basado en la teoría de Kubelka-Munk, que describía la transferencia radiativa de la dispersión difusa media, con dos coeficientes, de dispersión y absorción. (Baret et al 1988), y la combinación de las dos, para estudiar la conveniencia o no, de las diferentes técnicas para conseguir conocer la cantidad de humedad de una zona de estudio.

A través de estas dos simulaciones hemos analizado la capacidad de discriminación mediante el análisis de curvas ROC (Receiver Operating Characteristic), para posteriormente utilizar diferentes técnicas de Softcomputing (Inteligencia Computacional), en concreto, modelos neuronales y evolutivos, para validar los experimentos.

El modelo Prospect es un modelo de transferencia radiativa a nivel de hoja, muy utilizado en experimentos científicos. Está basado en el modelo de laminas de Allen et al. 1969 (Jacquemoud et al. 1990) en el que asume una hoja compacta con un coeficiente de refracción, y un coeficiente de absorción.

Prospect asume una estructura interna de la hoja con un número N de capas homogéneas, determinado por el índice de área vacía (VAI), que está dado de $n-1$, esta puede ser una aproximación continua donde N no necesita ser un número entero, un valor de $N=1$ asume una hoja lo más compacta posible representando una disminución de la reflectancia y aumento de la transmitancia en el IR cercano, valores superiores a 2,5 advierten hojas senescentes con una estructura interna desorganizada.

Jacquemoud y Baret calcularon el coeficiente de refracción con una hoja de maíz seca tratada con fluridone para eliminar rastros de pigmentos, por lo que se considerará a esto valores constantes para modelar la reflectancia y transmitancia de distintos tipos de hojas.

El modelo SAIL (Scattering by Arbitrarily Inclined Leaves), utiliza la aproximación de Kubelka-Munk de las ecuaciones de transferencia radiativa para caracterizar una zona de estudio de vegetación. El modelo SAIL se ha utilizado en Teledetección para la investigación de las propiedades de reflectancia espectral y direccional de las cubiertas de vegetación.

El análisis de curva de característica de funcionamiento del receptor (ROC) proviene del análisis estadístico de decisión (Green y Swets, 1966) y originalmente se utilizaba durante la Segunda Guerra Mundial para el

análisis de imágenes de radar. Las primeras aplicaciones de esta teoría en el área médica se produjeron durante finales la década de 1960.

Hoy en día, el análisis ROC es un método generalizado en el campo de medicina, con muchos libros de texto y artículos sobre ello, por ejemplo. (Kraemer, 1992, Metz, 1978, DeLong et al., 1988, Hanley, 1982, McClish, 1989, Armitage y Berry, 1994), contienen descripciones del mismo.

En líneas generales, el análisis de curvas ROC es una técnica estadística de decisión que, a partir de una característica, permite discriminar entre dos grupos o subpoblaciones de una población, proporcionando un criterio de clasificación y una medida de su bondad.

El uso de las curvas ROC está muy extendido en la actualidad y juega un papel fundamental en procesos de diagnóstico y clasificación, por ejemplo en detección de señales, psiquiatría, medicina y psicología, y recientemente también en meteorología, inteligencia artificial y economía. Desde la perspectiva de las ciencias de la computación el análisis ROC se ha utilizado cada vez más como una herramienta para evaluar los efectos discriminantes entre los diferentes métodos.

¿Qué es Softcomputing?

Softcomputing engloba una colección de técnicas computacionales de informática, inteligencia artificial y algunas disciplinas ingenieriles, que intentan estudiar, modelizar y analizar fenómenos muy complejos. Estos fenómenos son aquellos que por los métodos tradicionales no tienen un bajo coste, análisis y soluciones completas.

Softcomputing se diferencia de la computación tradicional (Hard computing) en que al contrario de ésta, tolera la imprecisión, incertidumbre o duda y verdad parcial. El modelo a seguir para el Softcomputing es el cerebro.

Los enfoques computacionales anteriores podían modelar y analizar con precisión sólo sistemas relativamente simples. Sistemas más complejos surgen en la biología, medicina, humanidades, ciencias de administración y campos similares que normalmente permanecían intratables por los métodos matemáticos analíticos convencionales.

Hay que señalar que la simplicidad y complejidad son relativas, y muchos modelos matemáticos convencionales han sido exigentes y muy productivos.

El Softcomputing consiste en: Explotar la tolerancia de la imprecisión, incertidumbre y verdad parcial, para conseguir la maleabilidad, robustez y solución de bajo coste.

La inclusión de la teoría de las redes neuronales en Softcomputing ha llegado algo más tarde. En esta unión, los principales componentes del Softcomputing son la Lógica borrosa o difusa (Fuzzy Logic), la teoría de

redes neuronales y el razonamiento probabilístico, con la unión más tardía de redes de creencia, algoritmos genéticos, la teoría del caos y partes de la teoría del aprendizaje.

Hay que destacar que el Softcomputing no es una mezcla de Lógica difusa, redes neuronales y razonamiento probabilístico. Más bien es una asociación en la que cada uno de los miembros aporta una metodología distinta para tratar los problemas en su campo. Visto así, las aportaciones principales de las tres, son más complementarias que competitivas.

Actualmente, existe un importante grupo de investigación que está enteramente dedicado al estudio, desarrollo y descubrimiento del Softcomputing, y éste es el BISC, o Iniciativa de Berkeley en Softcomputing, dirigido por Lotfi Asker Zadeh

Aunque el concepto de Softcomputing, se introdujo en 1990, fue en 1994 cuando Lotfi Asker Zadeh, el primero en hablar de conjuntos fuzzy, estableció la primera definición de Softcomputing. Fue la siguiente:

«Básicamente, Softcomputing no es un cuerpo homogéneo de conceptos y técnicas. Mas bien es una mezcla de distintos métodos que de una forma u otra cooperan desde sus fundamentos. En este sentido, el principal objetivo del Softcomputing es aprovechar la tolerancia que conllevan la imprecisión y la incertidumbre, para conseguir manejabilidad, robustez y soluciones de bajo costo. Los principales ingredientes del Softcomputing son la Lógica Difusa, la Neurocomputación y el Razonamiento Probabilístico, incluyendo este último a los Algoritmos Genéticos, las Redes de Creencia, los Sistemas Caóticos y algunas partes de la Teoría de Aprendizaje. En esa asociación de Lógica Fuzzy, Neurocomputación y Razonamiento Probabilístico, la Lógica Fuzzy se ocupa principalmente de la imprecisión y el Razonamiento Aproximado; la Neurocomputación del aprendizaje, y el Razonamiento Probabilístico de la incertidumbre y la propagación de las creencias» (Lotfi Asker Zadeh, 1994).

De esta definición se obtiene que el principal objetivo del Softcomputing es, por medio de distintos conceptos y técnicas, afrontar los problemas que se presentan en el mundo, el cual es, por naturaleza, impreciso y difícil de categorizar.

Las Redes de Neuronas Artificiales como clasificadores

Las Redes de Neuronas Artificiales (RNA) son básicamente aproximadores universales, surgieron del movimiento conexionista que nació junto la Inteligencia Artificial (IA) simbólica o tradicional hacia los años 50. La IA simbólica se basa en que todo conocimiento se puede representar mediante combinaciones de símbolos, derivadas de otras combinaciones que representan verdades incuestionables (axiomas). La IA tradicional asume que el conocimiento es independiente de la estructura que maneje los símbolos, siempre y cuando la 'máquina' realice algunas operaciones básicas entre ellos. En contraposición, los 'conexionistas'

intentan representar el conocimiento desde el estrato más básico de la inteligencia: el estrato físico. Creen que el secreto para el aprendizaje y el conocimiento se halla directamente relacionado con la estructura del cerebro: concretamente con las neuronas y la interconexión entre ellas.

Aunque la estructura básica de las RNA es común a todas las redes neuronales, existe una gran variedad de arquitecturas diseñadas para realizar tareas bien distintas. En este apartado vamos a tratar un tipo de redes concebidas para realizar funciones de clasificación. El objetivo de estas redes es clasificar un conjunto de configuraciones dentro de un conjunto de categorías. Para ello, es necesario una etapa inicial de entrenamiento en la cual, se debe enseñar a la red a qué categoría pertenece cada estímulo. Cuanto mayor sea el tiempo dedicado a la enseñanza y mayor el número de muestras a clasificar, mayor será la capacidad de generalización de la red. Esta es una propiedad importante, porque lo que realmente se pretende es que la red sea capaz de clasificar correctamente estímulos nuevos, o simplemente versiones distorsionadas de estímulos presentadas en la sesión de aprendizaje.

Las RNA pueden considerarse como funciones de decisión muy flexibles. En contraste con la filosofía de cálculo secuencial con procesadores muy potentes, las técnicas conexionistas están basadas en un alto número de procesadores muy simples funcionando en paralelo y de forma cooperativa (Proceso Paralelo Distribuido).

La estructura de RNA más general se basa en un conjunto de unidades de proceso $X = \{x_1, x_2, \dots, x_n\}$, $x_i = x_i(e_1, e_2, \dots, e_m)$; cada uno de los cuales realiza una operación sencilla sobre sus entradas $\{e_i\}$. Estos procesadores están interconectados entre sí en forma de red, de modo que las salidas de unos alimentan las entradas de otros. Ciertas entradas y salidas de algunos procesadores se utilizan como entradas y salidas globales de la red. El procesador elemental más común se conoce como neurona formal. Se trata de una función discriminante lineal: realiza una suma ponderada de sus entradas. Sus coeficientes definen un hiperplano en el espacio de entrada que separa los semiespacios de salida positiva y negativa. Mediante una función no lineal, normalmente de saturación, la salida de la neurona indica (por ejemplo con 0 ó 1) en qué semiespacio se encuentra su entrada.

El funcionamiento de una neurona artificial está burdamente copiado al de una neurona cerebral. Aprender significa, entrenar nuestras neuronas mediante el éxito o fracaso de una acción a unos estímulos sensoriales. Cuando cierta acción realizada en respuesta a alguna entrada sensorial es exitosa (por ejemplo, al beber agua calmamos la sed), las conexiones entre un grupo de neuronas se fortalecen, de manera que cuando tengamos una sensación sensorial parecida, la salida será la correcta. De esta forma se establecen fuertes conexiones entre grupos de neuronas, que pueden servir para realizar otras acciones complejas.

El modelo conexionista más conocido y utilizado en la actualidad es el perceptrón multicapa con propagación hacia adelante (*feedforward neural network*), que suele ir asociado al algoritmo de ajuste denominado Retropropagación del Gradiente (*Back-Propagation*).

Consideremos un conjunto de ejemplos de clasificación designados por e_j^i (j-ésimo ejemplo de la clase i-ésima). El objetivo de una RNA como clasificador es la de encontrar una función f que asocie a cada uno de los ejemplos con la clase a que pertenece $f(e_j^i)=i$ ($i=1..n$, $j=1..m_j$). Para ello se utiliza una red de neuronas estructurada en capas consecutivas, en las que las entradas a las neuronas de una capa son las salidas de la capa anterior. Las entradas de la primera capa son las componentes de cada ejemplo e_j^i , y la salida de la última capa codifica la clase a que pertenece la entrada (véase figura 0).

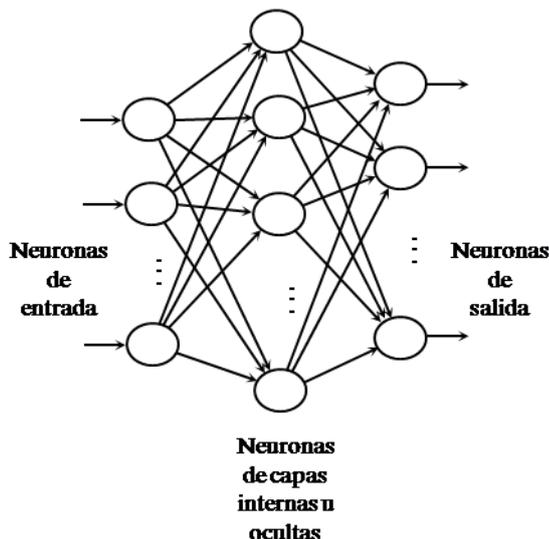


Figura 0: Estructura general de una red multicapa

Las asociaciones que realiza la red dependen de los pesos de cada una de las neuronas, así como de los coeficientes de sus funciones de saturación. Se representa este conjunto de parámetros mediante el “vector de pesos” w de la red. La salida dependerá de la entrada y del vector de pesos.

Definimos el error que comete la red al clasificar un ejemplo como la diferencia al cuadrado entre el vector clase deseado (y_d) y el vector obtenido a la salida de la red (y_s):

$$E(w) = \|y_d - y_s\|^2$$

Cada ejemplo que se presenta a la red define una función de error en el espacio de pesos:

$$E(w) = E(w, e) = \|y_d - f(w, e)\|^2$$

Al ser derivable la función de saturación, es posible calcular analíticamente el gradiente del error, lo que permite minimizarlo utilizando un algoritmo de gradiente descendiente:

$$\mathbf{w}_{k+1} = \mathbf{w}_k - \alpha \nabla E(\mathbf{w}_k, \mathbf{e})$$

donde α es el *coeficiente de aprendizaje* con el que podemos cuantificar el ajuste de los pesos de una iteración a otra. En realidad es necesario minimizar el error total cometido por la red, que es la suma de los errores cometidos en cada uno de los ejemplos:

$$E_T = \sum_i \| \mathbf{y}_{d_i} - f(\mathbf{e}_i) \|^2$$

por lo que la corrección que debe hacerse sobre cada peso es la suma de las correcciones debidas a cada ejemplo:

$$\Delta \mathbf{w}_T = \sum_i \Delta \mathbf{w}_i$$

A la hora de la corrección de los pesos se puede escoger entre dos estrategias. La primera acumula las correcciones correspondientes a cada uno de los ejemplos de entrenamiento y finalmente modifica los pesos con el valor acumulado debido a todos ellos. La segunda alternativa modifica inmediatamente los pesos cada vez que llega un nuevo ejemplo, teniendo en cuenta únicamente la función de error individual de cada ejemplo presentada. Mientras que la primera estrategia nos ofrece una aproximación del gradiente verdadero de la función error global, la segunda converge más rápidamente cuando la función de error global es muy “llana”. De ambas formas, es razonable que tras un número suficiente de iteraciones se alcance un mínimo de la función de error global en el espacio de los pesos.

El algoritmo de retropropagación del gradiente es la técnica más utilizada para el proceso de ajuste de los pesos de las neuronas internas, y tal como su nombre indica tiene la función de ir propagando los errores producidos en la capa de salida hacia atrás. Este algoritmo fue propuesto por Rumelhart, y su idea es la siguiente. En primer lugar se presenta un ejemplo a la red, produciéndose la activación sucesiva de las capas, desde la entrada hasta la salida. A continuación se realiza un proceso análogo de propagación de la “sensibilidad” del error respecto a cada neurona, desde la capa de salida (donde puede calcularse directamente) hasta la entrada, utilizando de forma recursiva las sensibilidades de una capa para calcular las de la capa anterior. El ajuste requerido para cada peso se obtiene directamente teniendo en cuenta cómo afecta el peso a su neurona asociada y cómo afecta esta neurona al error.

Lamentablemente, no existe una regla establecida para determinar el número de capas internas y el número de neuronas de cada una de ellas para un problema de clasificación en concreto. La topología característica de una red de neuronas está fuertemente determinada por el problema a resolver. Las redes de neuronas almacenan en su estructura la información sobre la resolución al problema en concreto para las que se han

diseñado. Por este motivo, el entrenamiento o aprendizaje de las redes de neuronas se convierte en un proceso crítico y vital para cualquier aplicación práctica.

Computación Evolutiva

La Computación Evolutiva intenta imitar la forma de trabajar de la evolución, como Darwin nos la expuso, para solucionar muy variados problemas,

La selección natural determina qué miembros de la población sobrevivirán hasta reproducirse y la reproducción sexual garantizando la mezcla y combinación de sus genes entre la descendencia.

La Computación Evolutiva interpreta la naturaleza como una inmensa máquina de resolver problemas y trata de imitar sus técnicas, con éxito en numerosas ocasiones, para cuestiones de muy variado proceder.

Con el término de Computación Evolutiva se engloba al conjunto de técnicas que basándose en la simulación de los procesos naturales y la genética se utilizan para resolver problemas complejos de búsqueda y aprendizaje.

Estructura del documento

El resto del documento posee la siguiente estructura. En el capítulo 1 de la tesis, se describen los conceptos relevantes para la estimación de cantidad de humedad utilizando técnicas de teledetección. En este capítulo se realiza una aproximación a los métodos clásicos de estimación. A continuación, en el capítulo 2 se valida la capacidad de las bandas del satélite para la predicción del riesgo de incendios mediante el uso del análisis de curvas ROC. En los siguientes capítulos, se describe la exploración realizada para la estimación del riesgo de incendios mediante el uso de modelos neuronales (capítulo 3) y evolutivos (capítulo 4). Por último, se destacan las principales conclusiones de la tesis doctoral y sus futuras líneas, así como la relación de la bibliografía utilizada para la elaboración de la misma.

Capítulo 1. La Teledetección y la estimación de humedad de coberturas de vegetación

Introducción.

(Adaptado y completado de “Remote Sensing of the Environment: State of the Science and New Directions”

Autores: Susan L. Ustin, Pablo J. Zarco-Tejada, Stephane Jacquemoud, Gregory P. Asner)

ADAPTADO Y COMPLEADO TAMBIÉN CON:

Chuvienco, E., D. Riaño, I. Aguado & D. Cocero 2002, Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: applications in fire danger assessment. *International Journal of Remote Sensing*. 23(11): 2145-2162.

Chuvienco, E., I. Aguado, D. Cocero y D. Riaño 2003: Design of an Empirical Index to Estimate Fuel Moisture Content from NOAA-AVHRR Analysis In Forest Fire Danger Studies. *International Journal of Remote Sensing*, 24: 1621-1637

Chuvienco, E., Cocero, D., Riaño, D., Martin, P., Martnez-Vega, J., de la Riva, J., et al. 2004. Combining NDVI and Surface Temperature for the estimation of live fuels moisture content in forest fire danger rating. *Remote Sensing of Environment*, 92(3), 322-331

Chuvienco, E. (2002). *Teledetección ambiental: La observación de la tierra desde el espacio*. Barcelona: Ariel.

La Teledetección es la técnica de adquirir imágenes de la superficie terrestre a través de sensores instalados en plataformas aerotransportadas, en avión o satélite principalmente, y la interpretación de éstas.

Desde hace 30 años, con el lanzamiento del primer satélite de la serie ERTS (*Earth Resources Technology Satellite*) el 23 de julio de 1972, la teledetección se ha convertido en parte integral del estudio y evaluación del medio ambiente a lo largo del planeta. Durante este tiempo, hemos visto el avance en el análisis de los datos desde la simple observación visual a sofisticadas interpretaciones basadas en los principios básicos de la radiación electromagnética. Hemos desarrollado una percepción científica de la Tierra como un sistema vivo, tanto física como biológicamente, y ha creado nuevos áreas científicas del sistema terrestre y ecología global. La Tierra está sujeta a numerosos parámetros ecológicos y ambientales, que combinados, tienen efectos singulares que afectan el bienestar físico y biológico medio ambiente en todas las escalas, desde lo local a lo global.

El contenido de agua de la vegetación condiciona la probabilidad de incendio, ya que aumenta la probabilidad de ignición y propagación del incendio. (Chuvienco et al., 2002; Burgan et al., 1998; Van Wagner, 1967. Jacquemoud 1990.

Entre los principales factores de comportamiento del fuego se encuentra el estado de los combustibles vegetales, concretamente la capacidad de combustión, cantidad, densidad, estratificación y humedad (Viger et al, 2004)

La importancia de este problema explica que se hayan realizado importantes esfuerzos para establecer modelos de predicción y prevención que consideren la variedad de factores que intervienen en el inicio y propagación del incendio (Chuvienco et al., 1996; Vázquez et al., 1999; Chuvienco et al., 2002; San Miguel-Ayán et al., 2003).

Para estimar estas variables a partir de una imagen, es preciso aplicar algún modelo que relacione la variables a estudiar de la vegetación con los valores detectados por el sensor, estos modelos se podrían aplicar a toda la imagen, ser reproducidos por la comunidad investigadora en similares condiciones, y presentar la menor dependencia posible de factores no relacionados con la vegetación (atmósfera, iluminación, ángulos de observación, etc.) También debería ser generalizable a otros lugares o situaciones y tener una precisión conocida (Strahler et al. 1986).

El incremento de nuestro conocimiento sobre las condiciones y la interdependencia de los diferentes ecosistemas, ha estimulado la investigación que intenta solucionar los problemas del medio ambiente, por ejemplo, la gestión de ecosistemas para promover la estabilidad en la producción de bienes y servicios o como restaurar los ecosistemas degradados.

Hay una creciente toma de conciencia entre el público en general, y particularmente entre los científicos sobre las consecuencias que de los sucesivos ascensos de temperaturas globales, la teledetección puede y debe desempeñar un papel importante en el suministro de los datos necesarios para evaluar las condiciones de los ecosistemas y vigilar el cambio climático.

Debido a los recientes desastres naturales, la desaparición de ecosistemas, y el descenso de biodiversidad, se han podido diseñar esquemas en un formato visual e intuitivo de fácil comprensión, que ha concienciado al público que esta preocupación no es exagerada, a pesar de esto, la teledetección nos ofrece datos que a menudo son complicados de convertir en información.

Sistemas de adquisición de información

La adquisición real de datos de satélite, comienza con el lanzamiento por la NASA de los satélites *Landsat Multispectral Scanner* en 1972, el *Thematic Mapper* en 1984 y avanzado de la y el *Advanced Very High Resolution Radiometer* de la serie NOAA lanzado en 1978, estos satélites han demostrado la variabilidad espacial y temporal de muchos ecosistemas y han mejorado profundamente el conocimiento de nuestro planeta y de su condiciones medio ambientales. De este modo, la teledetección ha traído una nueva conciencia del contexto espacial en el que ocurren los procesos medioambientales, poniendo de relieve la

interconexión de los diferentes ecosistemas. El registro de datos global que disponemos tiene una antigüedad de 30 años, en los cuales hemos podido observar los cambios en el entorno de la Tierra.

El radiómetro AVHRR (Advanced Very High Resolution Radiometer) fue diseñado para la observación meteorológica (determinación de cobertura de nubes y temperatura de la superficie), pero posteriormente los datos que suministra han encontrado numerosas aplicaciones en el campo de la Observación de la Tierra, lo que ha llevado a convertir al sensor AVHRR en una de las fuentes de datos de Teledetección más utilizadas.

Un radiómetro, es un sistema óptico-electrónico, que descompone la radiación recibida en varias longitudes de onda (bandas). Cada una de estas bandas se envía a un conjunto de detectores sensibles a esa región del espectro que la amplifican y la convierten en señal eléctrica y por un proceso de conversión analógico-digital, en un valor numérico conocido como Nivel Digital (ND). El AVHRR es ampliamente utilizado para este tipo de estudios (Ceccato *et al.*, 2003).

Su escáner tiene un barrido de banda ancha, capaz de medir radiación en diferentes zonas del espectro electromagnético. Al obtener una misma imagen en diferentes longitudes de onda se puede hacer un análisis multiespectral para definir con gran precisión parámetros hidrológicos, oceanográficos y meteorológicos.

La llegada de la era digital satélite, comienza con la puesta en órbita de los satélites multiespectrales y sensores aerotransportados estimulando el desarrollo de nuevas herramientas informáticas de análisis y métodos de visualización, esta nace de la necesidad de la detección de procesos biológicos y geoquímicos que se pueden medir en diferentes regiones del espectro electromagnético (EM), incluyendo el comportamiento y estudio de los diferentes compuestos atmosféricos, marinos y terrestres de los ciclos biológicos, geológicos y químicos (BGC). No menos importantes son las aportaciones de estos datos a todas las disciplinas dependientes del medio ambiente, lo que ha llevado al reconocimiento que la mayoría de los problemas de interés científico de hoy se encuentran enraizados entre estas disciplinas tradicionales y dependiendo en muchos casos de varias disciplinas con ciclos de transición de unas a las otras.

A través de plataformas aerotransportadas o satélite se obtienen medidas en todo el espectro electromagnético desde el ultravioleta a radiofrecuencias, lo que ha generado una amplia gama de nuevos instrumentos. El desarrollo de estos nuevos y más estables sensores con mayor resolución espectral y espacial, y la mejora en los calibrados los instrumentos de teledetección, ha estimulado el desarrollo de una base teórica para comprender la interacción de la radiación electromagnética con la materia.

El infrarrojo lejano es de gran utilidad para discriminar masas vegetales y concentraciones de humedad, en combinación con el infrarrojo medio, ya que en éste el agua presenta una fuerte absorción de reflectividad

para concentraciones de agua y el infrarrojo cercano nos sirve para discriminar ya que es relativamente insensible al agua. (Ceccato et al. 2002b; Gao 1996).

La reflectancia en el infrarrojo cercano y el infrarrojo de onda corta detecta en gran medida el agua y de materia seca en las hojas (Gao, 1996; Jacquemoud et al., 1996; Peñuelas et al., 1997; Tucker, 1980), mientras que los pigmentos fotosintéticos sólo son absorbidos en el borde visible y en la región espectral del rojo (Carter, 1994; Miller et al., 1990; Ustin et al., 2004a, b; Zarco - Tejada et al., 2004, Zarco - Tejada et al., 2001). Por lo tanto, las mediciones en el infrarrojo cercano y medio proporcionan una estimación cuantitativa contenido de agua de la vegetación, aunque también se ven afectadas por la estructura de la hoja, materia seca, estructura de la cubierta, y el índice de área foliar (Gao, 1996; Jacquemoud et al., 1996; Serrano et al., 2000; Zarco-Tejada et al., 2003).

Estos avances son el resultado de los diferentes proyectos financiados por la NASA principalmente los programas EOS (*Earth Observing System*) con el lanzamiento de los satélites Terra en 1999 y Aqua en 2002. Las nuevas medidas están cambiando rápidamente la manera en que los datos de teledetección son analizados, enfatizando los patrones visuales, utilizando combinaciones simples banda y procedimientos estadísticos, interpretaciones biofísicas basadas en la absorción y propiedades de dispersión del medio, y técnicas de softcomputing. En algunos casos, todavía hay necesidades insatisfechas de nuevos instrumentos para medir propiedades del Espectro Electromagnético para lograr una mayor comprensión del sistema y complementar la base de conocimientos que poseemos. Un ejemplo de lo primero es la necesidad de instrumentos ópticos para la obtención de imágenes de campo, en las regiones del espectro de radar y lidar (este instrumental emite pulsos de luz láser en la banda del espectro electromagnético comprendido entre el infrarrojo y ultravioleta), lo que permitirá a comprobar las hipótesis que predicen las complejas interacciones de luz y energía en los diferentes ecosistemas. Un ejemplo de esto último es la misión Space Shuttle con el reciente radar de interferometría (consiste en combinar la luz proveniente de diferentes receptores, telescopios o antenas de radio para obtener una imagen de mayor resolución) y el proyecto global de mapas topográficos.

El Lidar (Light Detection And Ranging) es un captador activo, análogo al radar pero con tecnología laser. Es un sensor aerotransportado. Leckie (1990) describe el enorme potencial del Lidar y sistemas de espectroscopia de imágenes para inventario de los bosques. Lidar tiene un gran potencial para la obtención de las propiedades estructurales de la vegetación seca que es el principal factor de ignición y propagación de incendios, ya que permite la obtención imágenes tridimensionales de la capa de vegetación (Riaño et al. 2007).

El uso de longitudes de onda mayores, que no se encuentran en los espectros solar o terrestre, implica la necesidad de utilizar sensores activos, es decir aquellos que generan su propio flujo de radiación. Su mayor

ventaja es que debido a su elevada longitud de onda (muy superior al tamaño de las gotas de agua en la atmósfera) no resulta absorbida por esta, además al ser un haz artificial puede manipularse la forma en que se emite para, así, obtener el máximo de información. El radar lateral aerotransportado (SLAR) permitió completar la primera cartografía de la cuenca del Amazonas y de varios países de Centroamérica, detectando recursos hídricos y mineros antes desconocidos (Henderson y Lewis 1998).

La teledetección ya no es dominio exclusivo de los organismos gubernamentales, Muchas plataformas están disponibles y son explotadas a gran escala para la investigación. El objetivo de este capítulo es poner de relieve nuevas y prometedoras áreas de investigación, remarcando nuevas ideas, necesarias para avanzar en la comprensión científica sobre el funcionamiento de los ecosistemas.

Métodos de estimación de contenido de agua para prevención de incendios

Dentro de este epígrafe vamos a explicar los diferentes métodos para estimar la cantidad de humedad en la vegetación, aspecto clave en el nacimiento y la propagación del incendio. Al final veremos los aspectos novedosos que ofrece esta tesis en la estimación de humedad utilizando imagen satélite y diferentes técnicas de inteligencia computacional.

Respecto a los incendios forestales, la vida vegetal representa una categoría de combustibles conocidos como "combustibles vivos". Se reconocen dos clases principales de los combustibles: herbáceos y leñosos. Los combustibles herbáceos incluyen gramíneas, dicotiledóneas y helechos, mientras que los combustibles leñosos incluyen las hojas, las ramas y pequeños arbustos (y, en el caso de los modelos de incendios de copas, también las hojas y ramas de las cubiertas forestales).

El estrés hídrico es todo aquel perjuicio que sufre la vegetación en su crecimiento por problemas de agua. Las plantas herbáceas anuales de raíces superficiales son los primeros en sufrir estrés hídrico.

Las plantas herbáceas perennes son las siguientes en ser afectadas, mientras que los arbustos (y árboles), que son las más profundas raíces, son los últimos en sufrir las consecuencias de la sequía.

Las plantas herbáceas caducas completan su ciclo de crecimiento dentro de un ciclo anual normal. Brotan, crecen, producen semillas, y mueren antes de la primera congelación. Las plantas herbáceas perennes que normalmente se mantienen verdes durante todo el ciclo de crecimiento y no maduran por completo hasta que se producen las temperaturas más frescas. Los combustibles leñosos, si son de coníferas, mantienen verdes durante todo el año, mientras que los árboles de hoja caduca pasan por un ciclo de crecimiento y caída de la hoja.

El cálculo del contenido en humedad de los combustibles vivos implica una gran complejidad, considerando el amplio rango de factores que le afectan y su amplia variación espacial y temporal (Chuvienco et al., 2004).

Para su estimación, se pueden seguir diferentes métodos. En esta sección descubriremos alguno de ellos.

Método de estimación directa

El método más preciso es la estimación directa de la humedad de la vegetación es mediante muestreo destructivo de las plantas. Este ejercicio permite obtener unos valores precisos de humedad pero resulta complicado cubrir un área grande que sea operativa espacial y temporalmente (Chuvieco *et al.*, 2004). En relación con el peligro de incendio se mide el agua en el combustible vivo (Fuel Moisture Content, FMC), la cantidad de agua por unidad de peso seco de la muestra:

$$FMC(\%) = \left(\frac{Pf - PS}{PS} \right) * 100$$

Donde Pf es el peso fresco de la muestra (el medido en cada adquisición) y Ps es el peso seco (el obtenido al final del proceso). Cuanto menor FMC, más seca la vegetación, mayor probabilidad de que se produzca un incendio y se propague a mayor velocidad (Burgan y Rothermel 1984).

En algunos de los estudios reseñados, el contenido de humedad de la planta no se calcula en función del peso seco, sino del peso para un máximo contenido de humedad. Es el denominado contenido relativo de agua (RWC, Relative Water Content), que se define como:

$$RWC(\%) = \left(\frac{Pf - PS}{PSAT - PS} \right) * 100$$

Donde Pf es el peso fresco, Ps es el peso seco y Psat es el peso saturado (con la máxima humedad posible). Se utiliza el FMC, en lugar del RWC, por dos razones. Por un lado, resulta una medida más ajustada, ya que en la literatura relacionada con incendios se pretende estimar el contenido absoluto de agua, más que la relación con un máximo de acogida teórico (Viegas *et al.*, 1990; Desbois *et al.*, 1997).

Por otro lado, la determinación del FMC es más operativa, ya que calcular el peso saturado implica, o bien sumergir las hojas en agua durante varias horas, escurrirlas y secar la superficie (Carter, 1991), para evitar incluir el agua exterior, o mantener las hojas cubiertas durante 24 h con papel absorbente humedecido en un frigorífico (Hunt y Rock, 1989) lo que resulta muy complicado cuando las muestras se toman regularmente y sobre áreas extensas.

El empleo de índices meteorológicos en el cálculo del contenido en humedad de los combustibles vivos facilitan una estimación espacial y diacrónica de la humedad (Camia *et al.*, 1999), y son fácilmente accesibles. Habitualmente se calculan a partir de variables básicas como la temperatura y humedad del aire, precipitación y velocidad del viento, que se combinan para simular el efecto que estos factores tienen sobre la fisiología de la planta.

No obstante, el empleo de índices meteorológicos también presenta dificultades operativas, puesto que las estaciones meteorológicas habitualmente no están próximas a áreas forestales, ni tienen la densidad adecuada (Chuvieco *et al.*, 2003).

Considerando además que el contenido de agua de las plantas no sólo depende de las condiciones atmosféricas, sino también de las características fisiológicas de la vegetación (Cocero et al., 1998), resulta muy probable que sea necesario adaptarlos a cada entorno y especie.

La inmensa mayoría de los incendios son intencionados o debidos a negligencias. No obstante, es indudable el interés de estudiar diversos factores físicos que intervienen en el fuego, ya sea el estado hídrico de los combustibles o las condiciones que imponen distintos factores meteorológicos, debido a que estos incendios tienen una incidencia y severidad mayor cuando estas variables naturales son críticas. (Camia et al., 1999)

Entre los diferentes agentes físicos que intervienen en un incendio forestal el clima ejerce un factor de control predominante en la frecuencia y ocurrencia de incendios, porque determina la sucesión de períodos secos y húmedos y, en consecuencia, la acumulación de biomasa, la carga de combustible y su estado hídrico. No obstante, a pesar del importante papel del clima, los parámetros meteorológicos, considerados de manera individual, presentan generalmente un grado de correlación débil con la ocurrencia de incendios. Por el contrario, cuando estos parámetros se consideran de forma conjunta los resultados son más satisfactorios.

La elaboración de índices meteorológicos que combinan las diferentes variables meteorológicas resulta, por lo tanto, fundamental en las tareas de predicción y prevención de incendios (Carrega, 1990).

Las condiciones climatológicas generales figuran entre los factores permanentes que permiten una caracterización general del área, si bien no influyen de forma decisiva en la variación diaria del Índice. Son los factores variables los que, medidos o estimados periódicamente, proporcionan la oscilación del índice a corto plazo (Vélez, 1985).

En las primeras investigaciones de carácter preventivo se observó que el comportamiento del fuego estaba influido por los fuertes cambios, que a lo largo del día experimentaba el contenido de humedad de los combustibles muertos existentes sobre el suelo (Chandler et al, 1983). Enseguida se planteó el análisis de las

Los índices disponibles en la actualidad varían mucho en función de las variables que consideran, los más sencillos utilizan diferentes variables meteorológicas para estimar la posibilidad de iniciarse un incendio, mientras en los más complejos se combinan un gran número de factores a partir de modelos teóricos y/o empíricos para conseguir índices de ignición y propagación. Aunque todos estos sistemas varían en complejidad, el nexo común a todos ellos es el objetivo de conseguir mediciones simples y comparables de la inflamabilidad del combustible forestal diariamente (Chandler et al., 1983).

El sistema de peligro estadounidense es representativo de los índices más teóricos, ya que tiene una fuerte fundamentación matemática y física en el análisis de la humedad del combustible. A partir de experimentos

realizados en el laboratorio se pretende medir la influencia de determinados tipos de combustible y factores meteorológicos sobre el comportamiento del fuego (Deeming et al., 1977).

Por el contrario, sistemas como el canadiense o el australiano, aunque basados también en fundamentos físicos, se desarrollaron a partir del análisis de numerosos datos de campo. Los índices canadienses se construyeron empíricamente, relacionando de forma conjunta la meteorología, la humedad del combustible y el comportamiento del fuego con datos registrados durante diferentes estaciones de incendios, en una serie de puntos determinados (Van Wagner, 1974; McArthur, 1973).

La mayoría de los índices meteorológicos de peligro poseen un factor acumulativo que considera a partir de las condiciones meteorológicas pasadas las condiciones presentes. El efecto de la meteorología a largo plazo se estima de forma mucho más refinada en los denominados índices de sequía (drought indices). En estos índices generalmente intervienen elementos como la capacidad de la reserva de agua del suelo, modelos de evapotranspiración, precipitación y tipo de combustibles. Un elemento esencial en estos índices es el denominado período de desfase (Van Wagner, 1985).

El índice utilizado en el sureste francés (Risque Numérique Meteorologique) es también un índice mixto (ignición / propagación). Al igual que en los casos anteriores se estima el contenido de humedad en dos tipos de combustible: la vegetación viva a partir de la reserva de agua del suelo y la vegetación muerta en función de distintas variables meteorológicas (Sol, 1990).

En el Índice italiano IREPI (índice di Riduzione Evapotranspirazionale del Percólo di Incendio) se utiliza la reserva de agua del suelo como estimador del comportamiento diario de la vegetación viva, pero en este caso el índice final sólo ofrece información sobre la ignición; las características de propagación del fuego se obtienen a posteriori considerando el factor viento (Bovio et al., 1984).

En España antes de disponer de un sistema de peligro propio se experimentó con diferentes métodos utilizados en distintos países. Desde 1987 se encuentra en vigor un índice de peligro diseñado por el ICONA (Probabilidad de Ignición), que se inspira en el método desarrollado por el laboratorio del fuego de Missoula, Montana (ICONA, 1990).

La mayoría de los modelos para predecir el peligro de incendios incluyen sólo los elementos meteorológicos, y mantienen valores constantes la cantidad de vegetación seca y la topografía (Andrews 1986, en Pyne et al. 1996). Sin embargo, esta simplificación, por muy útil que sea para el cálculo de valores, no nos permite conocer el riesgo de incendio de forma correcta, lo cual nos lleva a introducirnos en la teledetección como única posible alternativa válida para una correcta estimación del contenido de humedad.

Métodos basados en teledetección

La teledetección proporciona observaciones sistemáticas de la tierra con la suficiente resolución espacial, temporal y espectral sensibles a variaciones en el FMC para la prevención de incendios (Chuvienco et al., 2004). Los datos meteorológicos nos dan una idea del estado del combustible muerto pero no tanto del combustible vivo (Chuvienco et al., 2003), ya que dependen más del agua disponible en la raíces. En la siguiente sección haremos un análisis exhaustivo de los conceptos relacionados con la teledetección y su uso a la hora de predecir la cantidad de agua en la vegetación.

Características espectrales de la Vegetación.

La vegetación la vemos de color verde fundamentalmente porque la clorofila absorbe energía en las longitudes de onda del azul (0.4-0.5 nm) y el rojo (0.6-0.7 nm), reflejando más en el verde (Figura 1). Cuando las hojas no están sanas aumentan, disminuye la absorción de la clorofila aumentando la reflectividad en el rojo y azul. En el infrarrojo cercano las hojas no tienen la capacidad de absorber energía presentando una reflectividad muy alta en torno a 900 nm (Figura 1). Debido a la presencia de agua en las hojas la radiación se absorbe en el infrarrojo medio (930 nm a 2400 nm), y cuando el agua se pierde aumenta la reflectividad (Figura 1) Ustin, S.L et. Al 2004) El satélite MODIS, que se empleará en este estudio, contiene bandas en regiones del espectro que son sensibles a variaciones en el contenido de humedad de las hojas (Figura 1).

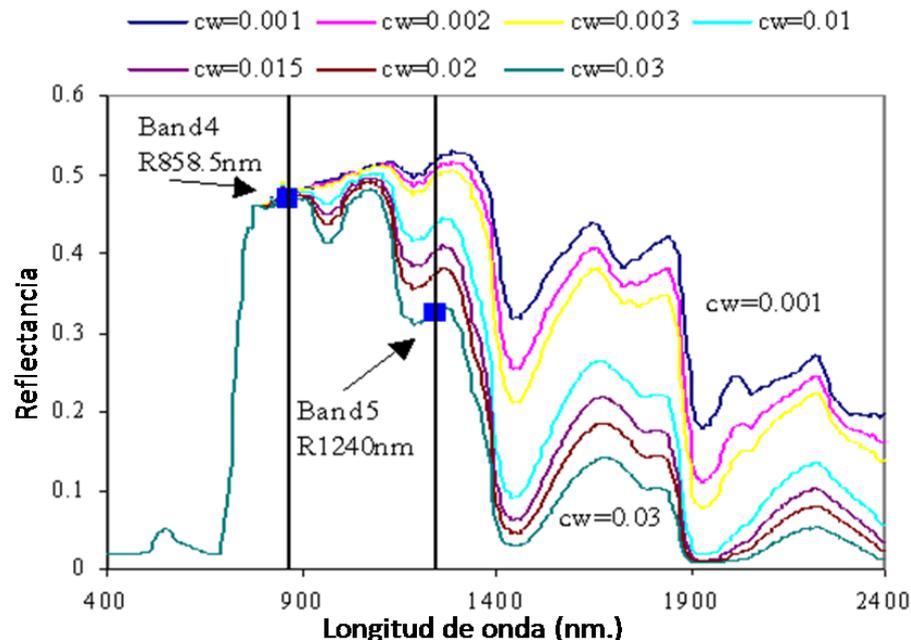


Figura 1: Simulaciones de la respuesta espectral de la vegetación en el espectro óptico empleando el modelo PROSPECT-SAILH. Se simulan hojas con distinto contenido de humedad. También se muestran donde se localizan las bandas MODIS (adaptado de Zarco-Tejada et al., 2003).

Capítulo 2. Evaluación del riesgo de incendio mediante curvas ROC Capítulo

Publicaciones correspondientes:

Evaluation of critical fuel moisture content values from simulated MODIS data for wildfire risk assessment using the Receiver Operating Characteristic curve

Luis Usero^a María Xose Rodríguez-Alvarez^b,

luis.usero@uah.es, mariajose.rodriguez.alvarez@usc.es

^a Departamento de Ciencias de la Computación. Escuela Técnica Superior de Ingeniería Informática. Universidad de Alcalá

^b Universidad de Santiago de Compostela

Resumen

Esta publicación presenta un nuevo método para evaluar los valores críticos de contenido de humedad en combustible para evaluar el riesgo de incendio, utilizando la curva ROC (Receiver Operating Characteristic) y datos simulados del sensor MODIS, ubicado en el satélite Terra (EOS AM).

Las reflectancias simuladas proceden del modelo de transferencia radiativa a nivel de hoja PROSPECT, en unión con el modelo a nivel de dosel SAILH. Los espectros simulados fueron generados de acuerdo a las características biofísicas de matorrales ubicados en Parque Nacional de Cabañeros, adquiridos de Yebra et al. 2008.

Se evaluaron la capacidad de las imágenes MODIS simuladas para discriminar entre alta y baja probabilidad de incendio. Los resultados utilizando la curva ROC mostraron que ninguna de las bandas simuladas del sensor MODIS tiene una buena capacidad de discriminación cuando se usan por separado, pero utilizando los datos de diferentes combinaciones de bandas, ofrecen una pequeña discriminación de errores.

Palabras Clave: Riesgo de incendios, Curva ROC, Contenido de humedad en combustible, Teledetección, MODIS, Modelos de transferencia radiativa

Introducción

La forma más usual de conocer el nivel de humedad de la vegetación relacionado con el peligro de incendios es conocer el contenido de humedad en combustible (FMC), se basa en el cociente entre la humedad de las hojas y la cantidad de biomasa seca, una posible aproximación se basaría en la recolección de muestras frescas y secas y su pesaje (Lawson and Hawkes, 1989). La recolección de muestras en campo requiere mucha mano

de obra y la repetición cada vez que cambiamos de zona de estudio, mediante métodos indirectos podemos conseguir datos espaciales y temporales de forma muy eficiente.

Mediante teledetección conseguimos una alternativa para obtener FMC. El sensor MODIS (Moderate Resolution Imaging Spectroradiometer) ubicado en los satélites Aqua (EOS PM) y Terra (EOS AM) dispone de 36 bandas en el espectro, algunas de las cuales son sensibles a los cambios de agua, por ejemplo la banda número 2 (de 841 a 876 nm.) y la banda número 5 (de 1230 a 1250 nm.) (http://nsidc.org/data/docs/modis_v5/spectral_bands.html).

Los modelos de transferencia radiativa simulan la respuesta espectral de la vegetación basándose en un conjunto de parámetros biofísicos. La inversión de estos modelos nos permite conseguir FMC (Riaño et al. 2005). El modelo de transferencia radiativa PROSPECT-SAILH simula coberturas de vegetación homogéneas basándose en los cambios de humedad en la biomasa, contenido de agua en las hojas, clorofila, estructura del mesófilo de la hoja (N), índice de área foliar (LAI), y en los puntos de vista de las variables geométricas, (zenit solar, azimut relativo) (Jacquemoud 1995).

Este artículo evalúa las bandas del sensor MODIS sensibles a los cambios en FMC usando espectros simulados basados en el modelo de transferencia radiativa PROSPECT-SAILH. Los espectros simulados fueron remuestreados para adaptarse a las bandas MODIS tras considerar su respuesta espectral. El modelo fue parametrizado de acuerdo a las características de matorral y monte bajo en el Parque de Cabañeros usando las simulaciones descritas en Yebra 2008. Para determinar las bandas que nos interesaban para obtener valores de FMC utilizamos la curva ROC (Metz 1978, Bamber 1975, Sweet and Pickett, 1982, Hanley and Mcneil 1982), que sirve para discriminar grupos de valores.

Curvas ROC

El análisis de curva ROC (Receiver Operating Characteristic) proviene del análisis estadístico de decisión (Green y Swets, 1966) y originalmente se utilizaba durante la Segunda Guerra Mundial para el análisis de imágenes de radar. Las primeras aplicaciones de esta teoría en el área médica se produjeron durante finales la década de 1960.

Hoy en día, el análisis ROC es un método generalizado en el campo de medicina, con muchos libros de texto y artículos sobre ello, por ejemplo. (Kraemer, 1992, Metz, 1978, DeLong et al., 1988, Hanley, 1989, McClish, 1989, Armitage y Berry, 1994, y Bamber 1975), contienen descripciones del mismo.

En líneas generales, el análisis de curvas ROC es una técnica estadística de decisión que, a partir de una característica, permite discriminar entre dos grupos o subpoblaciones de una población, proporcionando un criterio de clasificación y una medida de su bondad.

El uso de las curvas ROC está muy extendido en la actualidad y juega un papel fundamental en procesos de diagnóstico y clasificación, por ejemplo en detección de señales, psiquiatría, medicina y psicología, y recientemente también en meteorología, inteligencia artificial y economía. Desde la perspectiva de las ciencias de la computación el análisis ROC se ha utilizado cada vez más como una herramienta para evaluar los efectos discriminantes entre los diferentes métodos.

El problema de clasificación es un tópico recurrente en campos muy diversos: medicina, biología, finanzas, etc. Este problema puede ser descrito de forma general de la siguiente manera: se sabe que un conjunto de individuos u objetos pueden ser clasificados en diferentes categorías o clases (por lo general no observables de forma sencilla), y el objetivo es determinar la categoría a la que pertenece cada individuo a partir de cierta información observada. La importancia o necesidad de realizar esta clasificación dependerá de cada situación de estudio, pero por norma general, proporciona información relevante que podría ser empleada con posterioridad. Por ejemplo, dentro del campo médico, ser capaz de clasificar a un individuo como enfermo (o sano) en base a la información que se posee sería el punto de partida para su posterior tratamiento. En finanzas, la concesión de créditos suele estar supeditada al riesgo de no devolución. Por lo tanto, clasificar a un individuo como idóneo, o probable impago, proporciona una regla en la que basar la concesión de créditos, en nuestro ámbito trataríamos de saber si un área determinada está en peligro de incendio o no, con las consecuentes ventajas para el tratamiento de este problema.

Evidentemente, en muchas circunstancias el número de categorías o clases distintas en los que realizar la clasificación puede ser mayor de dos. Sin embargo, el caso de dos clases es uno de los más recurrentes en la práctica, y, habitualmente, el caso de múltiples clases puede ser descompuesto en una secuencia de casos dos clases (Krzanowski y Hand 2009).

Asociado a cualquier regla de clasificación, y en concreto una prueba de que nos ocupa, aparece la necesidad de evaluar su funcionamiento. Por lo general, la clasificación de alta/baja probabilidad de incendio está en función del resultado del modelo anterior que hayamos usado, que puede tener errores. Es decir, existe la posibilidad de clasificar a una cobertura como alta probabilidad de incendio, cuando no lo es, o viceversa. Por lo tanto, será necesario cuantificar los errores que se realizan en la clasificación para comprobar la validez o invalidez del modelo en la práctica. En otras palabras, tenemos que medir lo que se denota como la exactitud de los valores iniciales. En el caso de un test binario o dicotómico la precisión de alta probabilidad de incendio, es resumida con frecuencia a través de las fracciones de Verdaderos Positivos (TPF) y Falsos Positivos (FPF). La TPF, también denominada en la literatura médica Sensibilidad, es la probabilidad de clasificar correctamente a una cobertura de vegetación como alta probabilidad de incendio, mientras que la FPF o 1-Especificidad es la probabilidad de clasificar a una cobertura en buen estado como de incendio probable. Para los test con resultado continuo u ordinal, la medida de precisión diagnóstica más ampliamente empleada es la curva ROC (Receiver Operating Characteristic). La curva ROC extiende los conceptos de sensibilidad y Especificidad de los test binarios al caso continuo, al representar dichas cantidades para todos

los posibles valores de corte o umbrales de decisión, y que transforman, en la práctica, un test continuo en un test binario o dicotómico.

Relacionado a la curva ROC, se puede utilizar varios índices como resúmenes de la exactitud de una prueba diagnóstica. El más utilizado es el área bajo la curva ROC, se define como AUC (Area Under Curve) y toma valores entre 0,5, en el caso de una prueba de poca información, y 1,0 en un prueba perfecta, otro índice de resumen es el índice de Youden (Youden 1950, Shapiro 1999), el índice varía entre 0 y 1, tomando el valor 0 en el caso de un test no informativo y 1 en un test perfecto.

Métodos

PROSPECT-SAILH fue parametrizado para simular la respuesta espectral del sensor MODIS usando umbrales de los parámetros biofísicos (Tabla 1) generados por Yebra et al. 2008. Usando estos límites simulamos 2100 reflectancias de cobertura, para cada una de ellas determinamos FMC. Se estableció un umbral hipotético del 60% de FMC por debajo del cual consideramos la probabilidad de incendio alta, la curva ROC nos permitió evaluar que bandas eran sensibles a los valores de alto/bajo FMC.

Rango	N	Cw (cm)	Cm(g.cm ⁻²)	LAI	TS (°)
Max	2.5	0.031	0.04	3.2	51
Min	1	0.011	0.018	0.2	19

Tabla 1: Entradas PROSPECT-SAILH en la simulación de matorral y monte bajo. Índice de estructura del mesófilo de hoja (N), cantidad de humeda equivalente a nivel de hoja (Cw), biomasa seca(Cm), índice de área foliar (LAI), zenit de ángulo solar(TS). Adaptado de Yebra et al. (2008).

Análisis de datos y conclusiones

Las técnicas estadísticas apropiadas para evaluar la exactitud de un clasificador continuo dado, distinguiendo entre dos estados (S1 y S2), están basadas en el análisis de la curva ROC (Metz 1978, Bamber 1975, Sweet and Picket, 1982, Hanley and Mcneil 1982).

Siendo Y una variable continua, la clasificación de valores de Y como pertenencia a estados S1 o S2, puede hacerse eligiendo un valor de corte c, si $Y \geq c$ el valor es ubicado en el estado S1 y si $Y < c$ en el estado S2. En este caso la curva ROC es definida como el ratio de verdaderos positivos (TPR probabilidad de clasificación correcta de S1) frente a la tasa de falsos positivos (FPR, probabilidad de errores de clasificación de S2) a través de todos los posibles valores umbral. Tenemos diversos valores asociados a la curva ROC, área bajo la curva (AUC) o el índice de Youden son considerados como fedatarios de la capacidad de discriminación del clasificador. El AUC es uno de los valores asociado de uso más común, toma valores entre 0.5 (sin poder de discriminación) y 1 (perfecta discriminación).

Sin embargo, las reglas de clasificación basadas en los valores de Y que nos proporcionan una baja tasa de errores no tiene que ser el mismo criterio usado en el análisis de curva ROC. Por otra parte, sabemos que la mejor regla de clasificación basada en la probabilidad de que pertenezcan al estado S1 es condicionada por los valores de Y (Neyman y Pearson, 1963; McIntosh y Pepe, 2002). En cuanto a los valores de Y, el mejor clasificador basado en la elección de umbrales de clasificación es:

$$\tilde{Y} = f(\tilde{X}) = P(F=1 | Y). \quad (1)$$

En la práctica, la función de la ecuación 1 no es conocida y es necesaria su estimación.

En este trabajo, hemos evaluado la capacidad de discriminación de varias bandas MODIS para clasificar situaciones de alto o bajo peligro de incendio. Basándonos en valores de nuestro conjunto de datos de FMC, fueron evaluados como alto peligro de incendio las situaciones con valores de FMC por debajo del 60%, y como bajo peligro de incendio los valores por encima de este umbral.

Para este análisis, se consideró el clasificador y no las bandas del sensor MODIS, pero sus valores fueron transformados con la ecuación 1, aplicándolos para estimar la función $f(\tilde{X})$ del modelo de regresión logística flexible (Wood 2006). Explícitamente para cada banda MODIS considerada (de 1 a 7), se ajustó el siguiente modelo de regresión logística.

$$f(\tilde{MB}) = P(F=1 | MB) = g^{-1}(h(\tilde{MB})) \quad (2)$$

Donde MB depende de las bandas variables de MODIS, RF es una variable binaria dada por el valor 1 como alto riesgo de peligro de incendio, y 0 para bajo riesgo de peligro de incendio, g es una función lógica conocida y h es una función que no conocemos en origen.

El conjunto de datos fue dividido en dos grupos, el 60% para ajustar el modelo de regresión (conjunto de calibración) y el 40% restante para estimar la curva ROC (conjunto de validación). Las figuras 2 a 8 muestran la transformación estimada $\hat{f}(\tilde{X})$, junto con los intervalos de confianza del 95%, de la curva ROC y el área por debajo de la curva (AUC), para cada una de las bandas MODIS consideradas. Como se observa en la gráfica, ninguna de las bandas MODIS discrimina correctamente entre alto y bajo riesgo de incendio, con valores cercanos al 0.5 en todas, excepto en la banda 2 que el valor es 0.7.

El siguiente paso fue combinar las diferentes bandas MODIS para obtener un clasificador compuesto, e intentar una mejor discriminación. La metodología en este caso fue la misma que en el análisis anterior. Primero estimamos la probabilidad condicionada de alto riesgo de incendios a través de los diferentes valores de las bandas MODIS, utilizando una regresión logística multivariable (Wood 2006). Las bandas incluidas en el modelo fueron de la 1, 2, 4, 5, 6. Una vez ajustado el modelo, utilizamos las probabilidades estimadas con el conjunto de validación de la curva ROC. El resultado se muestra en la gráfica 8, y como puede verse, la

exactitud de la combinación de varias bandas de MODIS en la distinción entre alto y bajo riesgo de incendio es casi perfecta, con un valor próximo a 1 en AUC.

Todos los análisis estadísticos fueron realizados en R (R development core team 2008). Para estimar la función h en (2) se utilizaron el TPSS (*Thin plate splines smoothers*) (Wood, 2003) optimizando los parámetros de suavizado elegido de forma automática mediante el criterio UBRE (*Un-Biased Risk Estimator criterion*) (Wood, 2004). Todos los modelos han sido equipados por el uso de la función GAM del paquete mgcv. En lo que respecta a la estimación de la curva ROC y AUC, la ROCR, usamos también el paquete R.

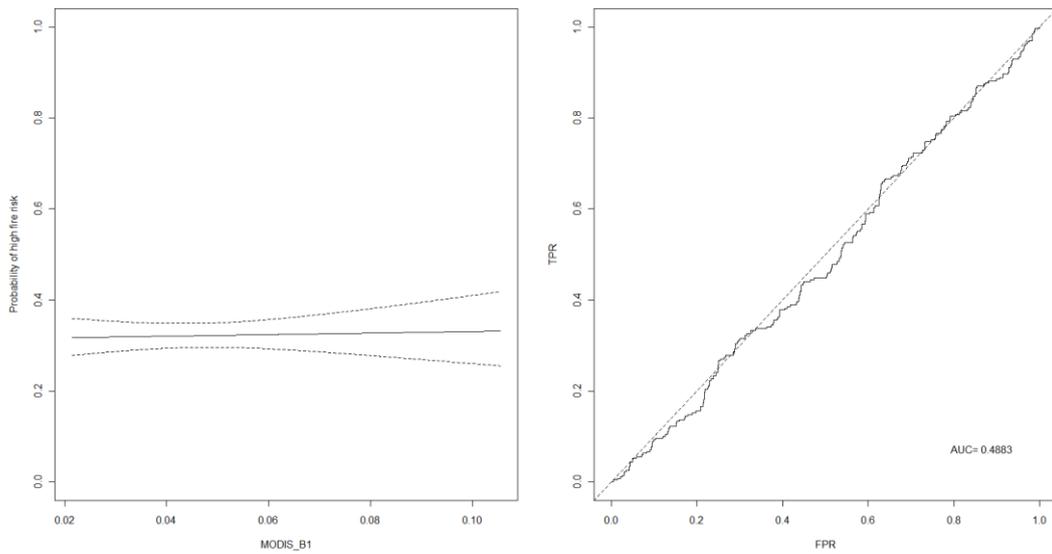


Figura 2: Banda 1 Sensor MODIS y tasa de falsos positivos

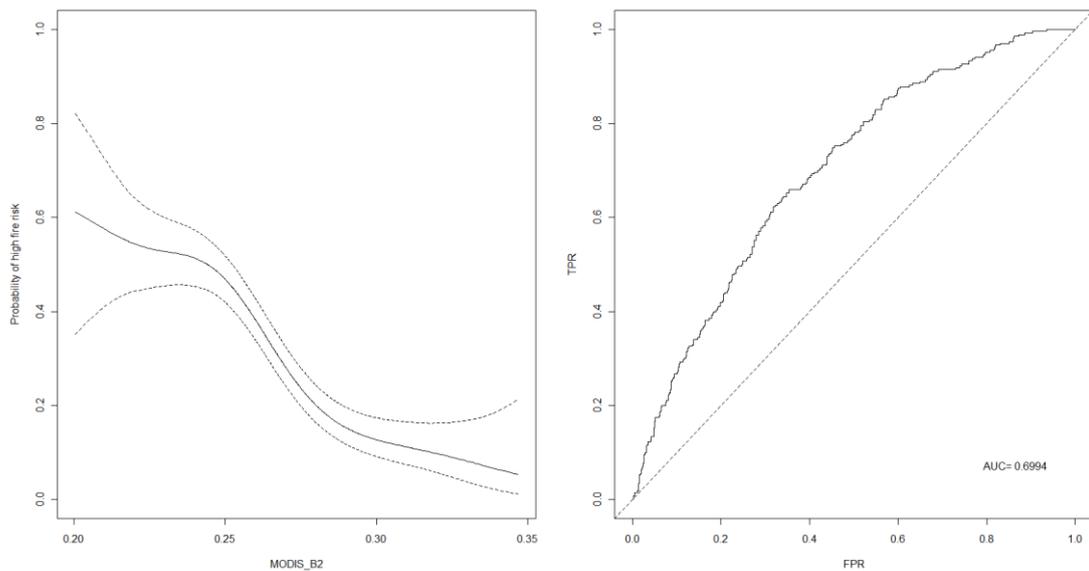


Figura 3: Banda 2 sensor MODIS y tasa de falsos positivos

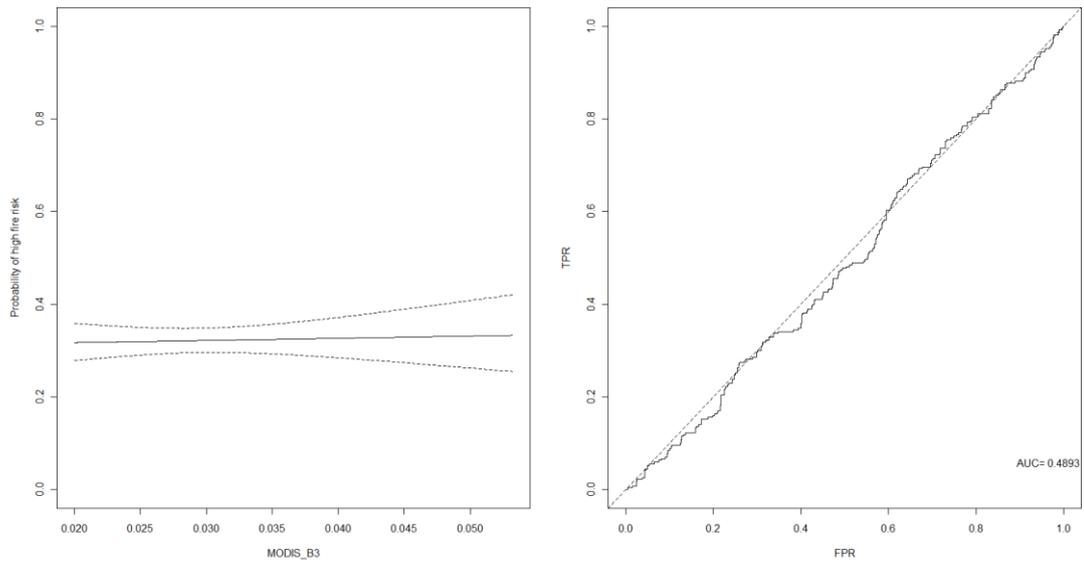


Figura 4: Banda 3 Sensor MODIS y tasa de falsos positivos

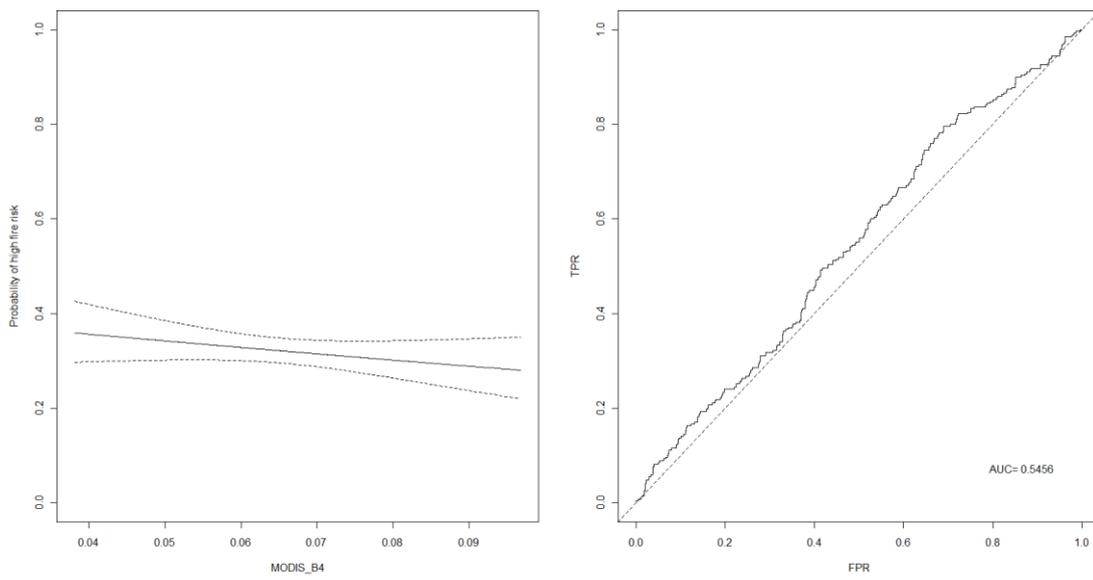


Figura 5: Banda 4 Sensor MODIS y tasa de falsos positivos

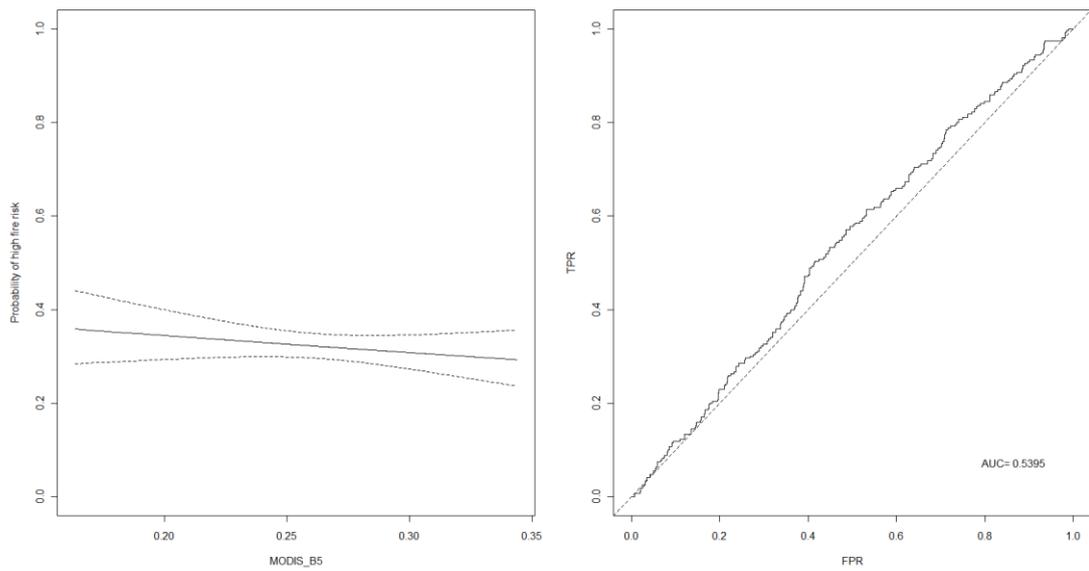


Figura 6: Banda 5 Sensor MODIS y tasa de falsos positivos

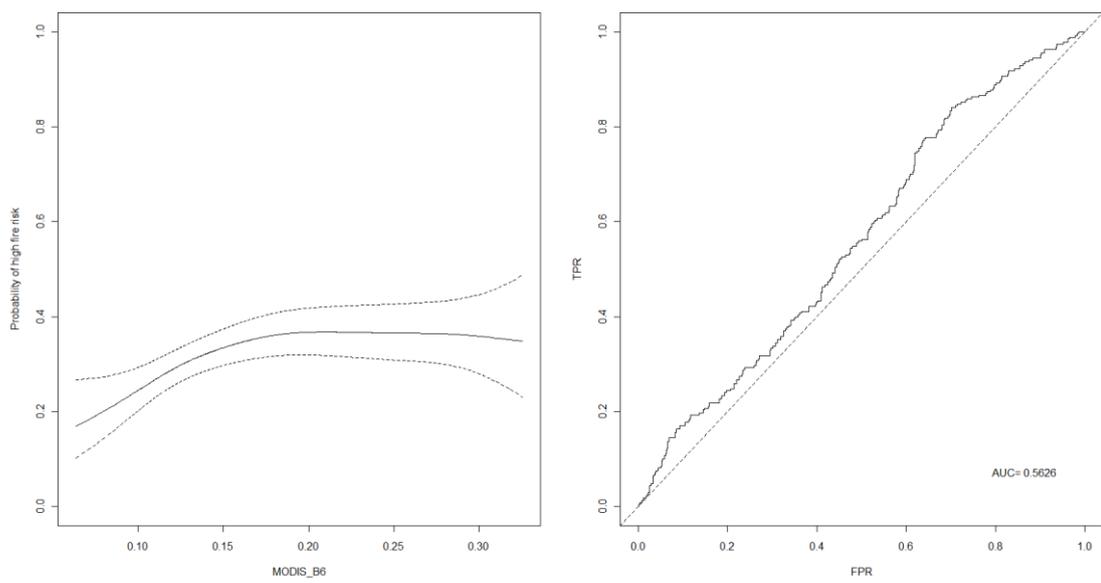


Figura 7: Banda 6 Sensor MODIS y tasa de falsos positivos

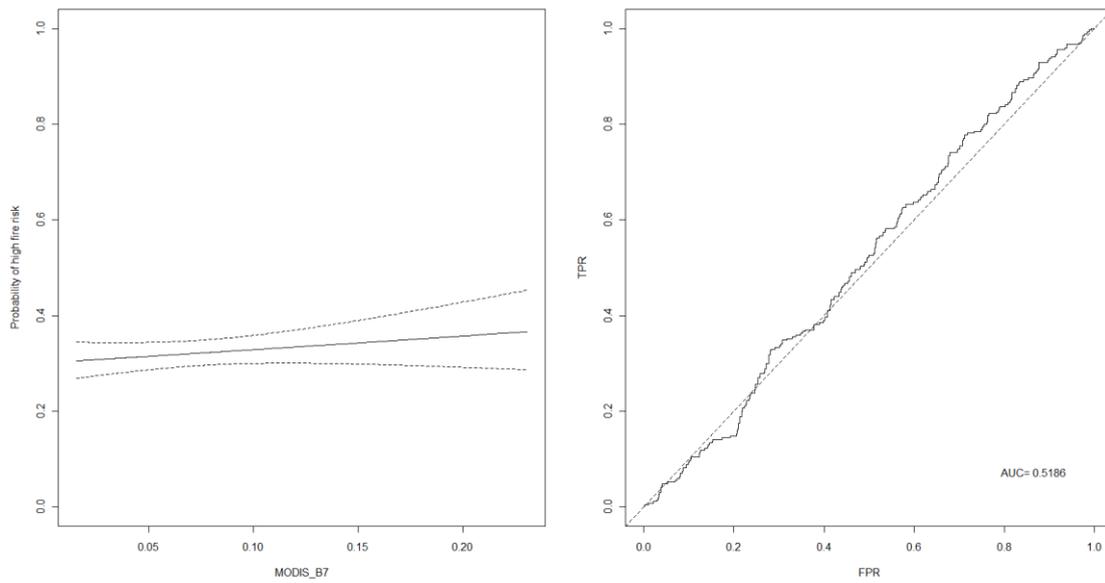


Figura 8: Banda 7 Sensor MODIS y tasa de falsos positivos

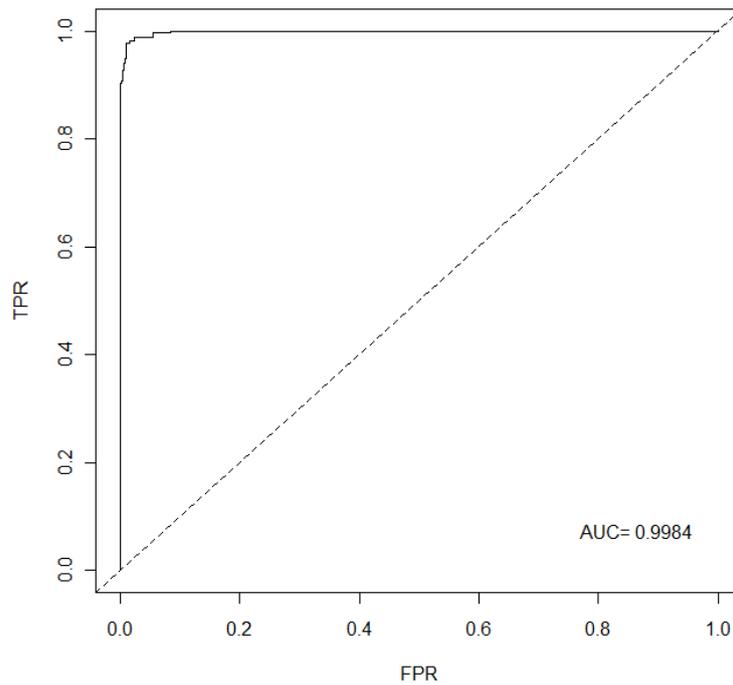


Figura 9: Combinación de bandas de MODIS para distinción entre riesgo de incendio alto y bajo con discriminación casi perfecta, valor próximo a 1 en AUC.

Agradecimientos

Agradecemos a David Riaño por facilitarnos las simulaciones de PROSPECT-SAILH.

Capítulo 3. Estimación del contenido de humedad en vegetación mediante de redes neuronales

Publicaciones correspondientes:

- Estimation of Fuel Moisture Content Using Neural Networks
 - David Riaño, Susan Ustin, Luis Usero y Miguel Ángel Patricio
 - IWINAC 2005 (International Work-conference on the Interplay between Natural and Artificial Computation)
- Estimación del contenido de agua del combustible mediante el uso de redes neuronales artificiales
 - Luis Usero, David Riaño, Pablo Zarco – Tejada, Susan Ustin
 - CEDI 2005 (1º Congreso Español de Informática)
- Estimation of equivalent water thickness using radial basis function neural networks
 - David Riaño, Miguel Ángel Patricio Pablo Zarco-Tejada, Carlos Rueda, Luis Usero y Susan Ustin
 - Proceedings of the 1 St International Conference on Remote Sensing and Geoinformation Processing in the Assessment and Monitoring of Land Degradation and Desertification. Trier, Germany, 7.-9. September 2005

Resumen

En los siguientes trabajos hemos utilizado Redes Neuronales Artificiales (RNA) para estimar FMC, mediante el cálculo de las dos variables implicadas, cantidad equivalente de agua (Equivalente Water Thinkness) y el contenido de materia seca (Dry Matter) a nivel de hoja a partir de la reflectividad y transmisividad obtenida mediante el modelo de transferencia radiativa PROSPECT(Leaf Optical Properties Spectra) y el modelo a nivel de dosel de bosque SAILH (Light Scattering by Leaf Layers with application to canopy reflectance modeling). Los datos fueron remuestreados para que coincidieran con las bandas del satélite AVIRIS. La materia seca (DM) fue estimada para muestras frescas y secas, ya que el agua enmascara el rasgo de absorción en muestras frescas. Se utilizó la base de datos "Leaf Optical Properties Experiment" (LOPEX), LOPEX producido durante un experimento realizado por the Joint Research Center (JRC) of the European Commission (Ispra, Italy) (Hosgood 1994). Para el IWINAC se utilizó una red neuronal artificial del tipo perceptrón multicapa, en la cual el 60% de las muestras se utilizaron para el proceso de aprendizaje de la red y los restantes para la validación. Las estimaciones de EWT y DM en muestras secas fueron tan buenas como otros métodos de prueba sobre el mismo conjunto de datos, tales como la inversión de modelos de transferencia radiativa. Las estimaciones de DM en muestras frescas ($r^2 = 0,86$) mejoraron significativamente los resultados con métodos de inversión de los modelos de transferencia radiativa ($r^2 = 0,38$). En el CEDI utilizamos una red neuronal de base radial, Para el proceso de aprendizaje tomamos el 7% de los datos (1000 espectros) y el resto para validación (14000 espectros). Las estimaciones del Cantidad equivalente de agua (EWT) fueron muy buenos, con un índice de correlación ($r^2=0.974$), lo que nos

confirma que la función de aproximación con la red neuronal nos permite utilizarla en pruebas de campo con muestras reales. En el trabajo presentado en el congreso de teledetección de Alemania utilizamos también una red neuronal artificial de base radial para estimar el contenido equivalente de agua (EWT) en este caso enlazando el modelo a nivel de hoja PROSPECT (Leaf Optical Properties Spectra) con el modelo a nivel de dosel de bosque SAILH (Light Scattering by Leaf Layers with application to canopy reflectance modelling), en el primer caso entrenamos la red datos reales y con datos del modelo PROSPECT y los resultados fueron muy buenos ($r^2=0.89$, 0.91 y 0.92 respectivamente), así como, la combinación de PROSPECT y SAILH los resultados también fueron muy buenos ($r^2=0.89$)

Introducción

El Contenido de humedad del combustible (Fuel Moisture Content) puede ser definido como la cantidad de agua por unidad de materia seca. FMC tradicionalmente es medido directamente a través de muestreos de campo. Este método requiere de mucho trabajo en el área indicada, y muchas veces ésta es de difícil acceso. El uso de la teledetección para medir las propiedades espectrales de las hojas puede proporcionar una estimación indirecta de la FMC, a fin de obtener una distribución espacial y temporal amplia. El estrés hídrico produce cambios en la reflectancia espectral y de transmisión de las hojas (Danson et al. 2003). Los modelos de transferencia radiativa muestran que estas mediciones espectrales están relacionadas con el contenido equivalente de agua (Equivalent Water Thickness), por unidad de área, y la cantidad de materia seca (Dry Matter (Ustin et al. 2004a), (Ustin et al. 1998) y (Zarco et al. 2003). FMC es un cociente de estas dos variables, EWT y DM, que puede estimarse de forma independiente (Jacquemoud et al. 1996). Se han utilizado diferentes instrumentos de teledetección, en diferentes escalas espectrales, para predecir este tipo de variables bioquímicas. Las probabilidades del comienzo y propagación de incendios están estrechamente relacionadas con el contenido equivalente de agua (EWT) en la vegetación (Burgan y Rothemel 1984). El contenido equivalente de agua (EWT) en hoja, cantidad de agua por unidad de área de hoja, es la variable de vegetación más estrechamente ligada a la reflectividad de la vegetación que podemos observar con imágenes satélite (Ustin et al 2004a). En consecuencia, la respuesta espectral a nivel del dosel de vegetación parece estar relacionada con el contenido equivalente de agua (EWT) del Índice de área foliar (Leaf Area Index (LAI)) (Ustin et al 1998). Los Espectrorradiómetros de Campo miden las propiedades del dosel de vegetación en un espectro único, promediando lo que observan en su campo de visión (Figura 10) (Jacquemoud et al. 1996).



Figura 10: Ejemplo de un espectrofotómetro. Perkin-Elmer Lambda-19. Capacidad espectral 250-2500 nm con un rango de 1,00 nm (www.grc.nasa.gov/WWW/epbranch/OpticalProps/lambda.htm)

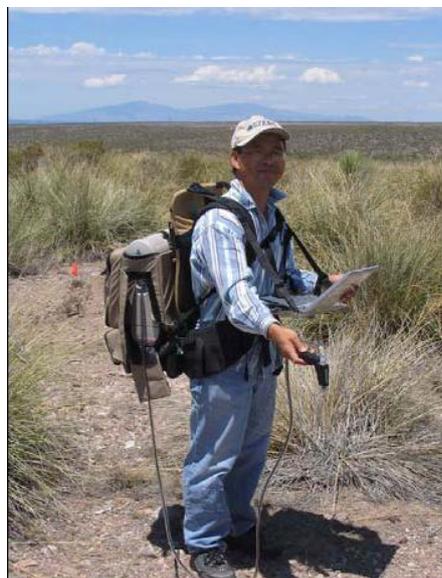


Figura 11: Ejemplo de un espectrorradiómetro de campo. Analytical Spectral Devices FieldSpec Profr, rango espectral 350-2500 nm con un rango 3-10 nm. ([www.asdi.com/especificaciones de los productos-FSP.asp](http://www.asdi.com/especificaciones%20de%20los%20productos-FSP.asp))

Los sensores aerotransportados proporcionan información acerca de las propiedades espectrales de la vegetación (Figura 12) (Serrano et al. 2000, Poggio and F. Girosi 1989).

Los sensores Hiperespectrales y multiespectrales embarcados en satélite nos proporcionan las propiedades espectrales de la vegetación en distintas escalas espaciales, y son capaces de abordar no sólo rangos locales de riesgo de incendio, sino también a escala regional o mundial (Figura 13) (Ceccato et al. 2002b), (Danson and Bowyer 2004), (Riaño et al. 2005) y (Moody and Darken 1989).

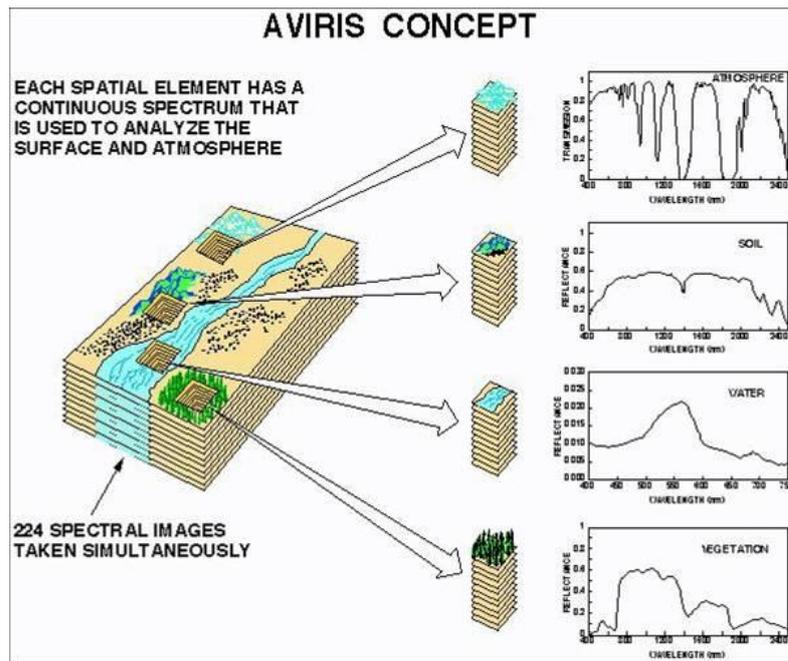


Figura 12: Ejemplo de sensor hiperspectral en avión. Espectrómetro de imagen visible infrarrojo de avanzado rango espectral entre 380 – 2500 nm y una ranura de 10 nm de ancho y una resolución espacial entre 4 y 20 metros dependiendo de la altitud de vuelo. (aviris.jpl.nasa.gov/)

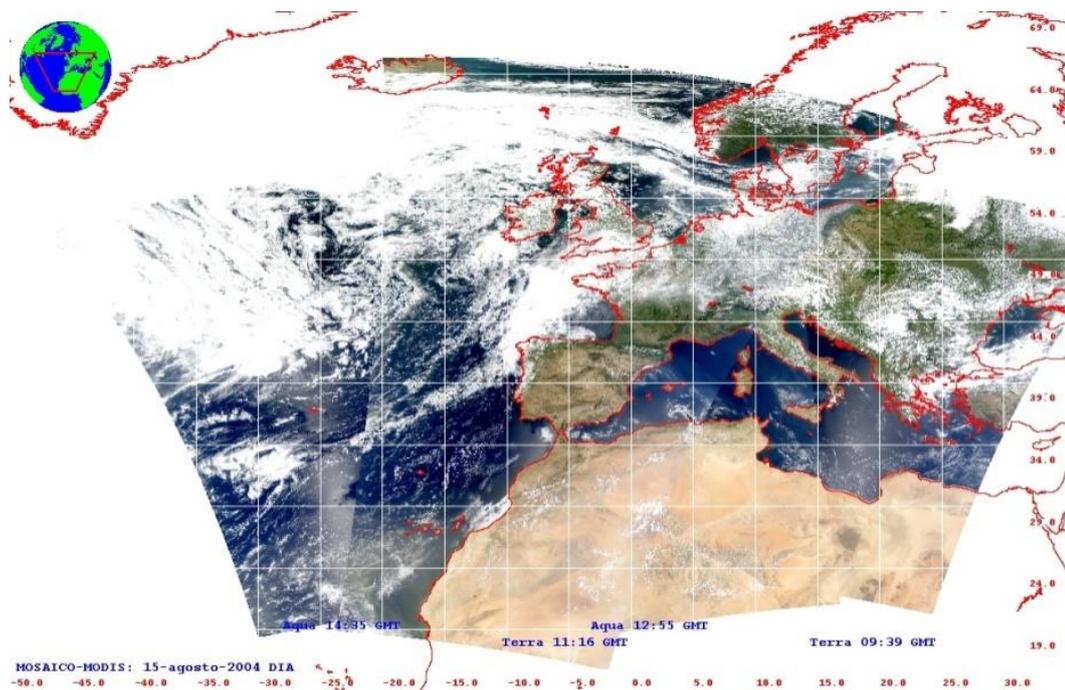


Figura 13: Imagen satélite hiperspectral. Espectrómetro de imagen de resolución moderada (MODIS). 36 bandas, rango espectral entre 405 y 14385 nm con una ranura de 20 nm de ancho, dependiendo de las bandas. Resolución espacial desde 250 hasta 1000 metros dependiendo de las bandas (www.latuv.uva.es/modisq1/)

Hemos utilizado diferentes métodos para extraer la FMC de las mediciones espectrales. Una posible forma de estimar FMC es aplicar un ajuste empírico utilizando regresiones lineales basadas en ratios

de banda (Jacquemoud and Baret 1990), (Danson and Bowyer 2004), (Riaño et al. 2005), (Hardy & Borgan 1999) (Paltridge & Barber 1988).

Este tipo de aproximaciones nos ofrecen un modelo de conocimiento para extraer características de la zona, pero necesita ser recalibrado cuando cambiamos la zona de estudio.

La inversión de modelos de transferencia radiativa nos ofrece una estimación de EWT y DM (Fourty, T. & F. Baret 1997, Jacquemoud et al 1996, Riaño et al 2005, Zarco-Tejada et al 2003). Esta técnica de extracción de valores de agua de la vegetación está ampliada al espectro electromagnético completo, en lugar de ratios de banda como en otros experimentos. Este método se puede extrapolar a otras áreas de estudio, aplicando un modelo diferente incluso con cambios del tipo de vegetación. Por otra parte, su uso requiere de cálculo de tiempo intensivo, siendo difícil aplicar a una imagen completa debido a los complejos cálculos matemáticos, pero con redes neuronales artificiales se puede hacer incluso en tiempo real. Los resultados para DM son más difícil de estimar que EWT, ya que el agua enmascara los rasgos de absorción de DM en la vegetación (Riaño et al 2006).

El trabajo presentado en IWINAC 2005 explora la estimación de EWT y DM utilizando Redes Neuronales Artificiales (RNA) y compara los resultados de la inversión con los del modelo de transferencia radiativa a nivel de hoja PROSPECT (Jacquemoud, & Baret 1990) y (Riaño et al 2006). Para la RNA, la EWT y la estimación de DM no se basan en las características de absorción específica. Este enfoque ha sido aplicado con éxito para estimar el índice de área foliar (Danson et al. 2003) y se tienen la ventaja de que si entrenamos a la red con una amplia gama de muestras de vegetación, el método podría aplicarse a una imagen completa.

Las redes neuronales artificiales se pueden utilizar como un aproximador universal (Poggio & Girosi 1989). Esto significa que las RNA pueden aproximar cualquier función después de una cantidad suficiente de aprendizaje. Esta es una ventaja sustancial sobre los tradicionales modelos estadísticos de predicción, ya que la relación entre los datos de entrada y la salida es altamente no lineal, y con interacciones significativas.

Métodos

En el artículo del IWINAC 2005 se utilizó el conjunto de datos estándar de LOPEX producido durante un experimento realizado por the Joint Research Center (JRC) of the European Commission (Ispra, Italy) (Hosgood 1994).

Se mide la reflectancia y la transmitancia con un espectrofotómetro Perkin-Elmer Lambda (figura 10) y los componentes bioquímicos diferentes, como EWT y DM, de una gran variedad de

especies. Se seleccionaron 49 especies de hoja ancha de la muestra completa, recogidas en un área de 50 km de alrededor del JRC en Ispra.

Cada muestra contiene 5 muestras de hojas frescas y 5 muestras de hojas secas. Aplicamos RNA para predecir EWT y DM a partir de las mediciones de reflectancia y transmitancia de muestras frescas y secas. Se realizaron un total de 2101 medidas entre 400-2500 nm., con una interpolación lineal cada 5 nm. lo cual nos dio un total de 421 valores de reflectancia y transmitancia que se convirtieron en 842 variables de entrada de la RNA.

Entre las diferentes posibilidades de redes neuronales artificiales, optamos por un perceptrón multicapa entrenado con el algoritmo de aprendizaje de backpropagation. Tres diferentes RNA fueron construidas:

- ✓ Para estimar EWT en las muestras de hojas frescas.
- ✓ Para estimar la DM en las muestras de hojas frescas.
- ✓ Para estimar la DM en las muestras de hojas secas.

Utilizamos para el desarrollo de las redes neuronales artificiales el JavaNNS (Java Neural Network Simulator) 1.1 programa desarrollado por la Universidad de Tübingen, en Alemania, así como, programas específicos de Java para la manipulación de los datos. Dividimos las muestras en dos conjuntos de datos: el aprendizaje y los datos de validación. Cada muestra estaba formada de cinco hojas, tres de ellas (60 por ciento de cada muestra) se utilizaron para el conjunto de aprendizaje y el resto de las hojas fueron usadas como validación.

Una función de activación sigmooidal fue utilizada en todas las neuronas con una tasa de aprendizaje de 0.2. Los datos de entrada y salida se normalizaron en un rango de [0, 1] Hay varias maneras de ajustar el tamaño de la red. En nuestro de trabajo, un método de ensayo y error controlado se utilizaba para optimizar el tamaño de la red y conformar su estructura.

La idea básica de este tamaño era, la optimización del método para probar un número diferente de neuronas en la capa oculta con diferentes pesos iniciales. Después de cada entrenamiento, fue testada con los datos de validación. A través de una extensa experimentación, finalmente seleccionamos una RNA con una arquitectura de un centenar de neuronas en la capa oculta y una neurona en la capa de salida (Figura 14).

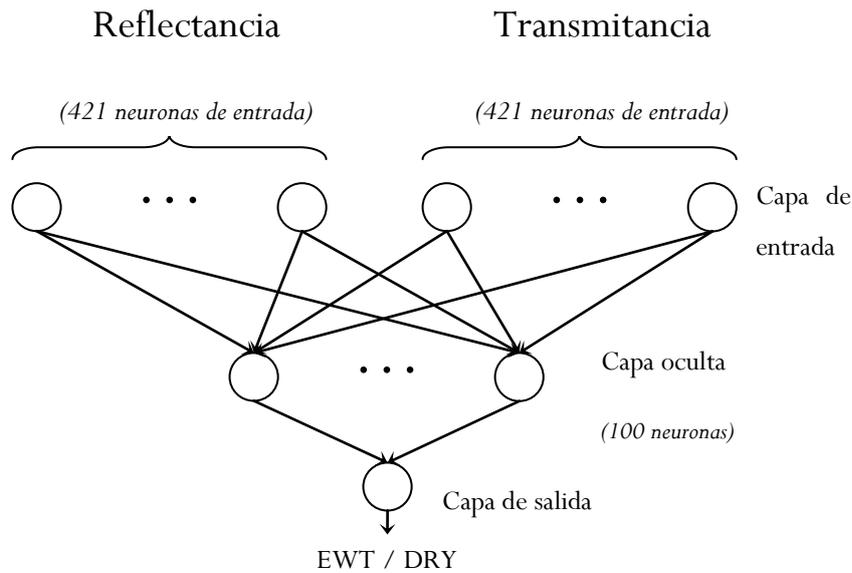


Figura 14: Arquitectura de la red neuronal IWINAC 2005

En el trabajo presentado en el CEDI 2005, utilizamos las bandas del sensor AVIRIS y diseñamos un modelo neuronal tomando datos del modelo físico. Se diseñó un modelo neuronal con objeto de estimar los índices EWT a partir de valores de reflectancia y transmitancia. Se generaron 15000 muestras con el modelo PROSPECT, eligiéndose aleatoriamente 1000 espectros para entrenar la red y 14000 espectros para validar los resultados, utilizando los siguientes rangos:

Numero de capas(N).	Clorofila(cab):	Contenido de agua(cw):	Materia Seca(DM):
• Inicial: 1	• Inicial: 10	• Inicial: 0.001	• Inicial: 0.001
• Salto: 0.6	• Salto: 20	• Salto: 0.001	• Salto: 0.002
• Final: 4	• Final: 90	• Final: 0.05	• Final: 0.02

Tabla 2: Rangos de entrenamiento y validación de la red neuronal

Los datos de reflectividad y transmisividad obtenidos por medio del modelo de transferencia radiativa a nivel de hoja PROSPECT fueron remuestreados a 210 de las 224 bandas del sensor AVIRIS. Las otras 14 bandas no se utilizaron ya que normalmente contienen alto nivel de ruido por absorción total de agua. Se utilizó este sensor para simular este tipo de estimaciones por su alta resolución espectral y amplia utilización en teledetección.

Se utilizaron 210 valores de reflectancia y 210 valores de transmitancia, teniendo un total de 420 variables de entrada, y 290 neuronas en la capa oculta. Entre la multitud de arquitecturas de RNA, hemos elegido una red neuronal con base radial (Radial basis function networks - RBFN). Las RBFN son una de las aportaciones más recientes en la investigación sobre redes neuronales. Fueron creadas por Moody y Darken (Moody and Darken, 1989), cuyo principal objetivo era reducir el tiempo de aprendizaje de las redes multicapa tradicionales. A diferencia de las multicapa tradicionales, sólo tienen tres capas (entrada, oculta y salida). Igual que ellas son perfectas funciones de aproximación. La gran diferencia se encuentra en su capa oculta.

Cada una de las neuronas implementa una función radial (normalmente gaussiana) sobre una porción del espacio de entrada de la red. La capa de salida es una combinación lineal de estas gaussianas y se demuestra que con ellas como base se puede reconstruir cualquier tipo de función.

En nuestra experimentación hemos utilizado la “Neural Network Toolbox” de Matlab para el desarrollo de las redes neuronales, así como para los programas específicos para la manipulación de datos. El modelo neuronal diseñado posee 290 neuronas en la capa oculta y una neurona de salida (figura 15). El error mínimo permitido durante la fase de aprendizaje ha sido 0.001 y el radio de las funciones radiales de la capa oculta de 0.35.

Los datos se dividieron en dos conjuntos: aprendizaje y validación. El 7% de los datos (1000 espectros) fueron usados para el proceso de aprendizaje de la red neuronal y el resto para validación (14000 espectros).

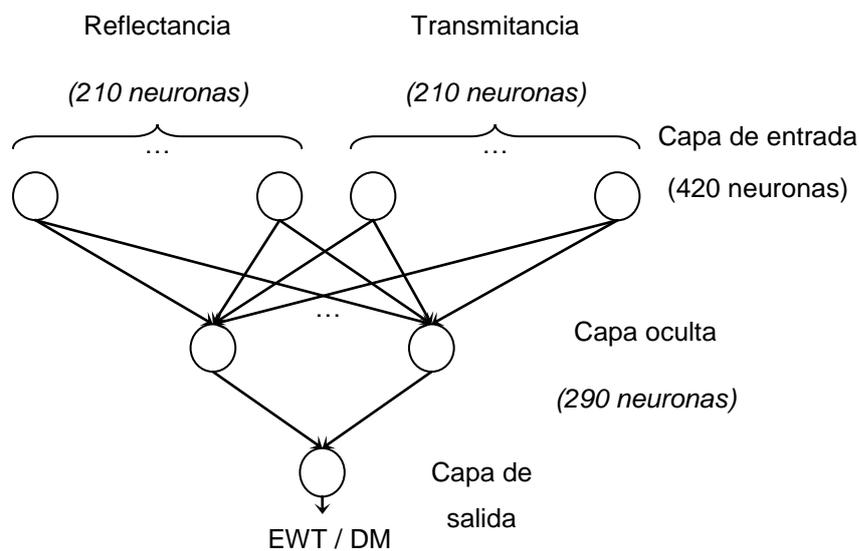


Figura 15: Arquitectura de la red neuronal CEDI 2005

El trabajo presentado en el congreso de Teledetección en Alemania se basa también en el conjunto de datos LOPEX que se produjo durante el experimento en Ispra (Italia) (Hosgood et al. 1994). Mejoramos lo presentado en el CEDI recalibrando la red y mezclando los modelos PROSPECT y SAILH. Las mediciones espectrales se realizaron cada 1 nm. entre 400-2500 nm. Los datos espectrales se interpolaron para proporcionar información comparable a la del sensor AVRIS (<http://aviris.jpl.nasa.gov>). Se obtuvieron un total de 210 reflectancia y transmitancia los valores, con un total de 420 variables de entrada para volver a muestrear los datos espectrales.

Realizamos la red neuronal utilizando tres técnicas diferentes:

- Entrenando la red con el conjunto de datos LOPEX: RNA entrenados con muestras todas las muestras de LOPEX pero una cada iteración. Que fueron 48 muestras, de un total de 240 hojas. La muestra restante se utilizó para la validación (dejando uno fuera para la validación cruzada). El proceso se repitió para cada muestra para desarrollar una validación cruzada.
- Entrenado con los datos del modelo PROSPECT (JACQUEMOUD and BARET 1990): Se realizó la red neuronal con los datos del modelo PROSPECT de modo directo utilizando el rango de los valores y los pasos se muestra en la Tabla 3 con un total de 160.325 casos. Se seleccionaron estos pasos para cubrir mejor la variación completa de todo el espacio del modelo. Se seleccionaron aleatoriamente 600 casos para entrenar la RNA. La validación se realizó con Conjunto de datos LOPEX.
- Entrenando con PROSPECT y el conjunto de datos LOPEX: Entrenamos la red neuronal con una sub-muestra de 100 casos de datos del modelo PROSPECT en modo directo y todas las muestras LOPEX pero una en cada iteración. La muestra LOPEX restante se utilizó para su validación. El proceso se repitió para cada muestra LOPEX para realizar una validación cruzada.

	N	C_{a+b} ($\mu\text{g}/\text{cm}^2$)	EWT (g/cm^2)	DM (g/cm^2)
	1.0–		0.001-	0.001-
Range	4.0	10-90	0.053	0.018
Step	0.3	20	0.001	0.001

Tabla 3: Rango de los parámetros de entrada del modelo PROSPECT.(N) Estructura interna de la hoja, (C_{a+b}) contenido de clorofila a + b,

Aplicamos una red neuronal de base radial (RBF) (MOODY, J. and DARKEN, C. J., 1989) utilizando el toolbox de Matlab 7.0 de redes neuronales (The MathWorks, Inc., www.mathworks.com, EE.UU.). Este tipo de RNA es más rápido de entrenar y puede trabajar con un gran conjunto de variables de entrada. Las Redes Neuronales artificiales de base radial son estructuralmente más simples con sólo tres capas: una entrada, oculta y de salida capa. Las neuronas de la capa oculta tienen una función Gaussiana estática. Ésta función de Gauss responde sólo en una pequeña parte de la región del espacio de entrada donde se centra la función gaussiana. La clave para la implementación exitosa de estas redes en Matlab es encontrar el número adecuado de neuronas para cubrir el espacio de entrada. El número final de las neuronas se ha seleccionado automáticamente cuando se llegó al error máximo. El entrenamiento de la red de base radial fue realizado con una combinación de los modelos PROSPECT Y SAILH (ZARCO-TEJADA et al. 2003) para estimar EWT y LAI. Seleccionamos un amplio rango de valores de entrada para el modelo (Tabla 4). Los valores de los dos tipos de clorofila (C_{a+b}) se fijó a $40 \mu\text{g}/\text{cm}^2$, ya que tiene poca influencia en la estimación de EWT (Zarco-Tejada et al. 2003). El ángulo de visión cenital y los ángulos azimut se establecen en 0^0 y el ángulo cenital sol se fija a 55^0 . Generamos 8000 supuestos seleccionados al azar y utilizamos 600 casos para el

entrenamiento de la red neuronal y los 7400 restantes para la validación. No utilizamos saltos en el rango de las diferentes variables en el modelo de dosel de vegetación.

	N	EWT (g/cm ²)	DM (g/cm ²)	LAI	LIDF	ρ_s
Range	1-2	0.001-0.003	0.001-0.03	1-5	Planophile Erectophile Plagiophile	Bright Médium Dark

Tabla 4: rango de los parámetros de entrada de la combinación de los modelos PROSPECT-SAILH, N, EWT, DM, LAI, leaf inclination density function(LIDF), y reflectancia del suelo(ρ_s).

Resultados

La estimación de EWT y DM para muestras frescas o secas con redes neuronales ha funcionado bien en todos los casos, en el caso del trabajo para el IWINAC pueden verse las gráficas (Figura 16 a Figura 21). Hemos comparado nuestros resultados con la estimación de estas variables mediante la inversión de modelos de transferencia radiativa (IRTM) (Riaño et al., 2006) (Tabla 4). EWT se calculó con gran precisión con ambos métodos, mientras que DM para las muestras secas se predijo un poco mejor por ANN. Por otra parte, DM para muestras frescas sólo es estimado con precisión por RNA. La inversión de modelos de transferencia radiativa ha funcionado mal porque los rasgos de absorción de la materia seca están enmascarados por el agua cuando las muestras son frescas (Riaño et al., 2006). Las complejas interacciones en el espectro afectaron a DM que no fue no consiguió captar los rasgos de absorción en el IRTM ($r^2 = 0,38$), pero fueron estimados notablemente bien en la ANN ($r^2 = 0,96$, el aprendizaje, $r^2 = 0,86$, validación). El proceso de aprendizaje incluyó varias hojas de la misma especie, fue necesario comprobar si RNA funcionará cuando las especies que no fueron utilizados en el proceso de aprendizaje se incluyen en la validación.

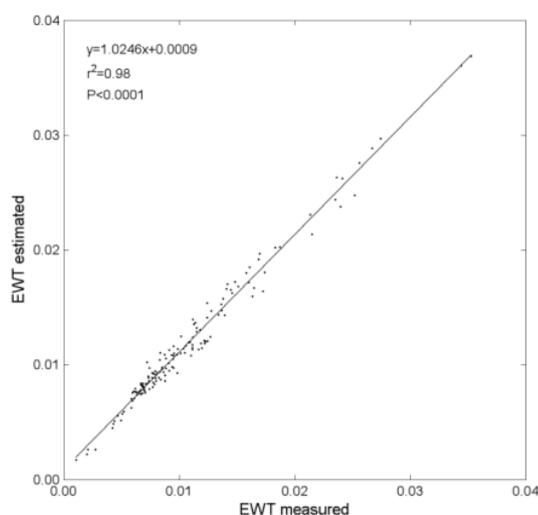


Figura 16: Estimación de EWT con los datos utilizados para aprendizaje

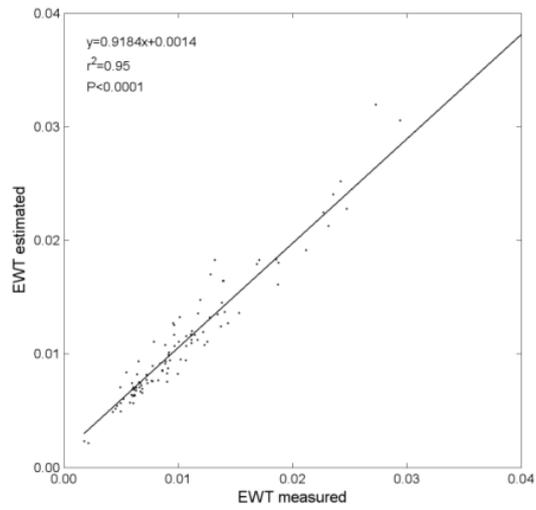


Figura 17: Estimación de EWT con los datos utilizados para validación

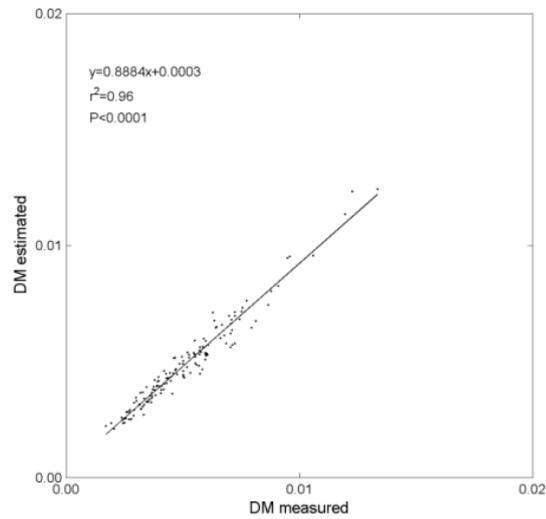


Figura 18: Estimación de DM en muestras frescas con los datos utilizados para el aprendizaje

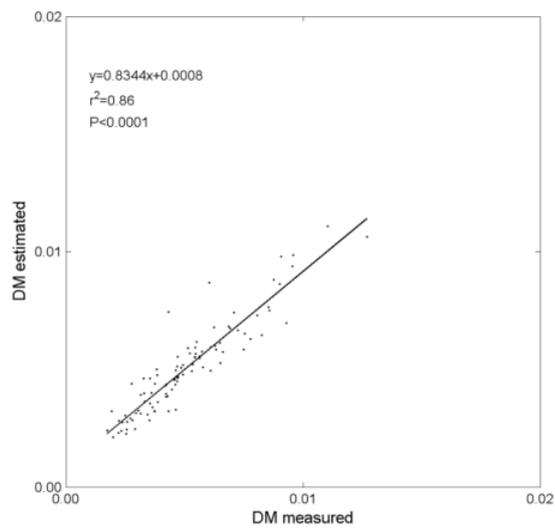


Figura 19: Estimación de DM en muestras frescas con los datos usados para validación

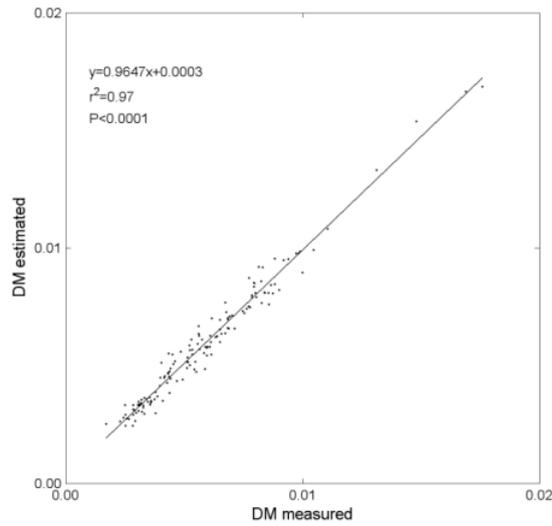


Figura 20: Estimación de DM en muestras secas con los datos utilizados para el aprendizaje

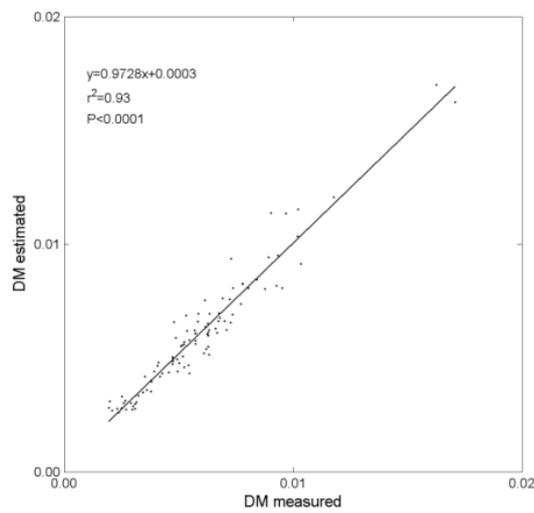


Figura 21: Estimación de DM en muestras secas con los datos utilizados para validación

	N	EWT	DM (frescas)	DM (seca)
IRTM	245	0.94	0.38	0.85
RNA aprendizaje	147	0.98	0.96	0.97
RNA validación	98	0.95	0.86	0.93

Tabla 5: Número de muestras (N) y correlación (r^2) para la estimación de EWT y DM usando IRTM (Riaño et al., 2006) y redes de neuronas artificiales.

En el trabajo del CEDI 2005 las estimaciones de EWT en entrenamiento tienen un error cuadrático ($r^2=0.996$) (figura 22), y las estimaciones de validación habiendo entrenado solo con mil espectros son muy buenos con un ($r^2=0.974$) (figura 23) lo que sugiere que la función de aproximación con la red neuronal nos permitirá utilizarla potencialmente en pruebas de campo con muestras reales.

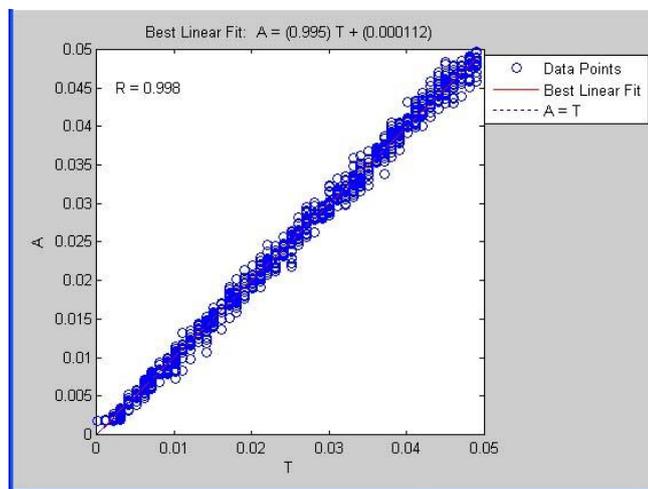


Figura 22: Estimaciones de EWT en entrenamiento

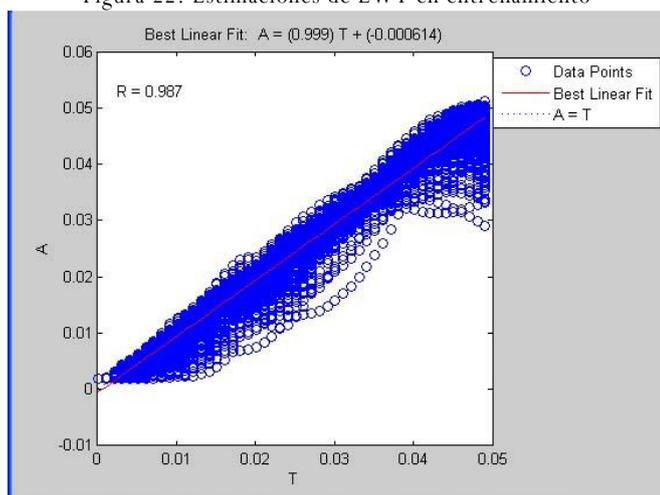


Figura 23: Estimaciones de EWT en validación

En el trabajo presentado en el congreso de Teledetección en Alemania el entrenamiento de la red neuronal y la validación cruzada con el conjunto de datos LOPEX presenta un buen ajuste ($r^2 = 0,89$) (Figura 24). También entrenamos con el modelo PROSPECT y validamos con el conjunto de datos LOPEX y los datos fueron buenos ($r^2 = 0,91$) (Figura 25). Realizamos una normalización de las variables principales con el fin de ser más compatibles las salidas del modelo PROSPECT con la base de datos LOPEX, mejorando la capacidad e predicción de la red neuronal. Las variables que se desviaban más del 0,1% de la varianza fueron retenidas.

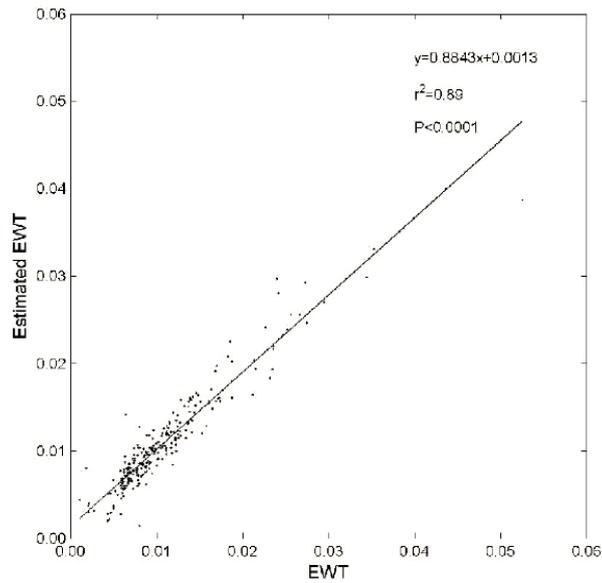


Figura 24: Estimación de EWT usando para entrenar la red todo el conjunto de muestras LOPEX pero una en cada iteración (n=245).

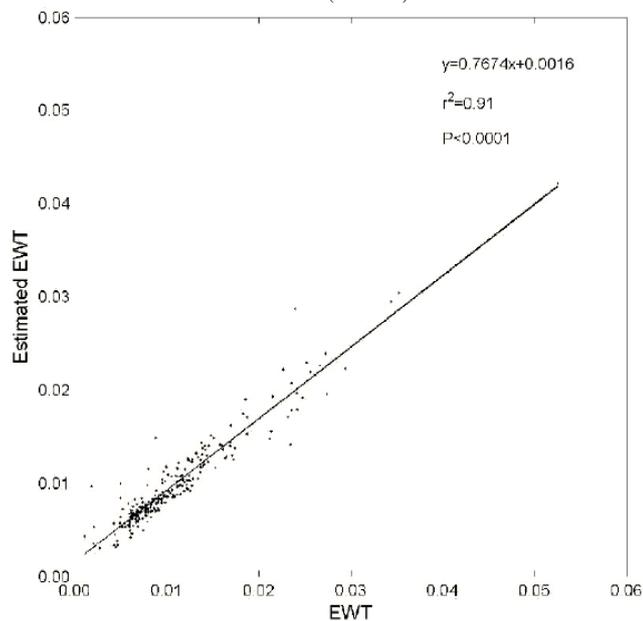


Figura 25: Estimación de EWT usando para entrenar la red el modelo PROSPECT (n=245).

El entrenamiento de la red con el modelo PROSPECT combinado con datos reales LOPEX mejora ligeramente los resultados ($r^2 = 0,92$) (Fig. 26) que únicamente utilizando el modelo PROSPECT o los datos LOPEX en entrenamiento. EWT y LAI se estimaron con precisión entrenando la red neuronal con diferentes muestras generadas por el mismo modelo de dosel de vegetación ($r^2 = 0,98$) (Fig. 27).

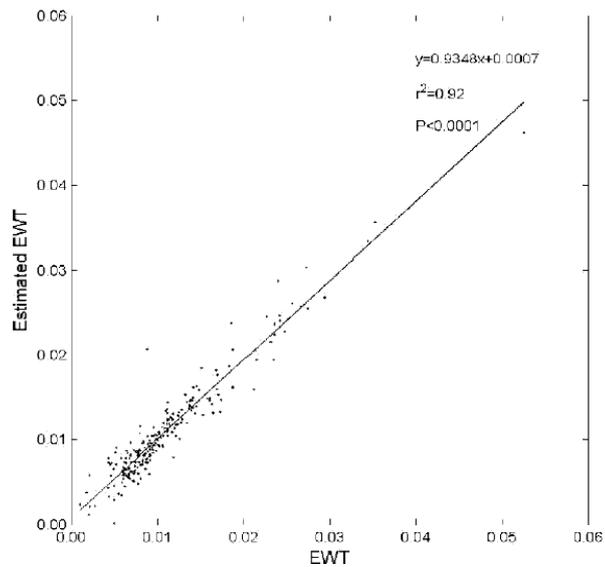


Figura 26: Estimación de EWT con datos LOPEX usando una muestra por iteración (n=245).

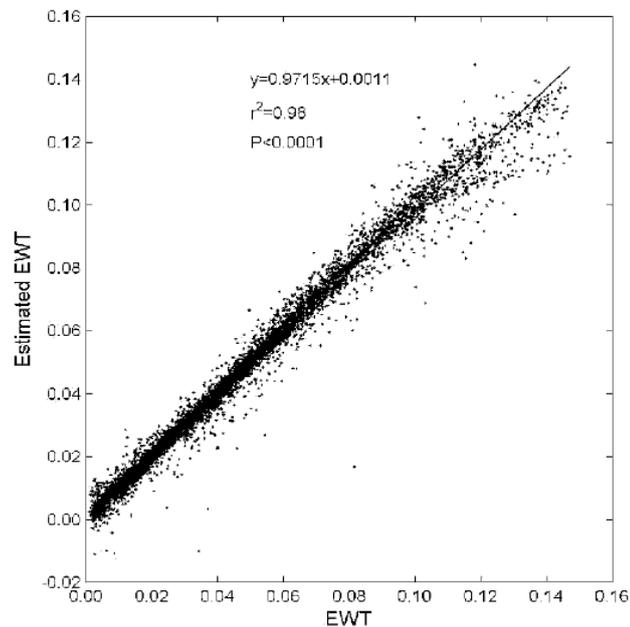


Figura 27: Estimación de EWT y LAI con muestras generadas con PROSPECT y SAILH (n-7400).

Discusión

En el primer artículo que presentamos en el IWINAC, se demostró que una red neuronal del tipo perceptrón multicapa puede estimar EWT a nivel de hoja entrenando la red con muestras de la misma planta ($r^2 = 0,95$). Esto quedó demostrado con la validación cruzada que realizamos en la red neuronal con base radial para predecir EWT a nivel de hoja utilizando un amplio conjunto de datos independientes, de muestras reales procedentes de otras plantas y de diferentes especies ($r^2 = 0,89$). También fue posible que la red de neuronal artificial de base radial fuera entrenada con el modelo de transferencia radiativa

PROSPECT. La ventaja de utilizar un modelo de transferencia radiativa en el entrenamiento de la red, es que, estos modelos abarcan un área global, y se podrían aplicar a cualquier conjunto de datos.

Si en entrenamiento que realizamos de la red es correcto, la bondad de los resultados estará íntimamente relacionada con la inversión modelo de transferencia radiativa ($r^2 = 0,94$) (Riaño et al. 2005). La inversión de modelos de transferencia radiativa es mucho menos costosa computacionalmente. Para mejorar la precisión en conjuntos de datos específicos, la red neuronal de base radial se puede entrenar combinando datos de modelos de transferencia radiativa con datos reales. En nuestro caso, la combinación de datos PROSPECT Y LOPEX mejora ligeramente la estimación de EWT ($r^2 = 0,92$). Estamos en proceso de investigación para saber cómo combinar los datos de inversión de modelos de transferencia radiativa con datos reales en el proceso de entrenamiento de la red, ya que la normalización y los componentes principales no funcionaron bien en alguno de los casos. Las redes de base radial funcionaron con los modelos PROSPECT-SAILH muy bien ($r^2 = 0,98$), aunque no se hizo uso de rangos para cubrir todo el espacio de variación. El número de muestras de entrenamiento fue suficientemente alto. Una vez automatizada la red neuronal, los resultados dependerán de la bondad de la inversión del modelo utilizado, pero computacionalmente mucho más rápido. Por lo tanto este tipo de modelo se podría aplicar imágenes completas de teledetección, en tiempo real

Agradecimientos

A Brian Hosgood por la realización del conjunto de datos LOPEX, y su ayuda. Al proyecto de la NASA, Global Estimation of Water Content (NNG04GQ42G)

Capítulo 4. Estimación del contenido de humedad en vegetación mediante Computación Evolutiva

Publicación correspondiente:

- Context Information for understanding Forest Fire using Evolutionary Computation.
 - Luis Usero, Ángel Arroyo, Julio Calvo.
 - IWINAC 2007 (International Work-conference on the Interplay between Natural and Artificial Computation)

Resumen.

Una de los principales componentes para comprender el riesgo de incendios forestales y el comportamiento del fuego es el combustible (se define así, la cantidad de vegetación en la zona que se quema). El contenido de humedad en vegetación es un factor muy importante para el riesgo de incendios y el comportamiento del fuego. La información sobre la el contenido de agua en vegetación es vital para comprender los procesos involucrados en la iniciación y la propagación de los incendios forestales. En ese sentido, utilizamos un nuevo método para estimar el contenido de agua en un área específica de vegetación (Canopy water content, CWC) mediante simulación de imágenes del sensor MODIS embarcado en satélite. Desarrollamos la inversión de un modelo de transferencia radiativa denominado Forest Light Interaction-model (FLIM) utilizando computación evolutiva. El contenido de agua en el dosel de vegetación (CWC) es fundamental en la evaluación del riesgo de incendios forestales, desde el momento del inicio ya que aumenta las posibilidades de ignición hasta su avance por el terreno en caso de que la humedad sea muy baja. Las simulaciones fueron realizadas con el modelo de transferencia radiativa FLIM para una amplia gama de características de dosel del bosque y valores de CWC. Se utilizaron 50 muestras para el proceso de entrenamiento y otras 50 muestras para la el proceso de validación, proporcionando un error medio cuadrático $RMSE = 0,74$ y $r^2 = 0,62$. Actualmente estamos investigación la posibilidad para aplicar este método en imágenes reales MODIS.

Introducción

Detectar el contenido de agua en el dosel de vegetación (CWC) es un indicador para conocer el estrés hídrico en la vegetación. Los datos adquiridos mediante teledetección son vitales para entender el comportamiento del fuego y el riesgo de incendio forestal. Medir las propiedades espectrales de las hojas puede proporcionar una información estructural indirecta para la estimación del CWC a fin de obtener una perspectiva completa espacial y su distribución temporal. El contenido de agua del dosel vegetación (CWC) es el peso del agua por unidad de área de hoja y por unidad de superficie del suelo. Los datos de CWC son especialmente importantes para diversas aplicaciones medioambientales, incluyendo el riesgo de incendios forestales (Chuvieco et. Al 2004). La velocidad de avance del frente de incendio esta íntimamente relacionada con el nivel de CWC.

El contenido de agua de la vegetación (CWC) es una relación entre el contenido de la hoja de agua (Cw) y el índice de área foliar (LAI). Las estimaciones del contenido de agua en vegetación son muy importantes para evaluar el estado de la vegetación en la agricultura y la silvicultura (Gao et al. 1995, Gao 1996, Peñuelas et al. 1997, Ustin et al. 2004b, Zarco-Tejada et al. 2003), y se han utilizado para la evaluación de la sequía (Ustin et al. 2004b) y la predicción de la susceptibilidad a los incendios forestales. (Chuvieco et al. 2004, Riaño et al. 2005, Ustin et al. 1998)

Diferentes instrumentos de teledetección se han utilizado para predecir las variables bioquímicas incluyendo Convención en diferentes escalas de medición de sus propiedades espectrales. Espectrofotómetros de campo que pueden predecir las propiedades de las hojas individuales (Ceccato et al. 2002, Riaño et al. in press)



Figura 28: Ejemplo de espectrofotómetro. Modelo 1100 RS, longitud de onda 335-1000 +/- 2.00 nm de precisión (<http://www.lehmanscientific.com/spectro.html>).

Los Espectroradiómetros de campo miden las propiedades de las cubiertas de vegetación en un espectro único, un promedio de lo que se ve en su campo de visión (Riaño et al., En prensa). Los sensores hiperespectrales aerotransportados proporcionan información espacial de las propiedades espectrales de la vegetación como la CWC (Fig. 29)

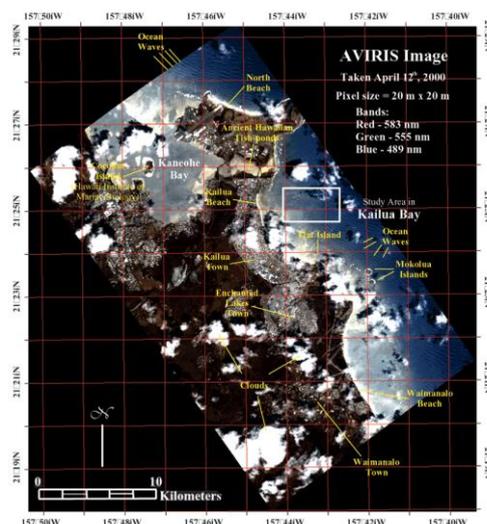


Figura 29: Ejemplo de imagen de sensor hiperespectral AVIRIS (Advanced Visible Imaging Spectrometer), rango espectral entre 380 y 2500 nm. con slot de 10 nm, resolución espacial de 4 a 20 m. depende de altura (http://www.higp.hawaii.edu/.../aviris_color_mosaic.gif)

Los sensores satélite Hiperespectrales y multiespectrales proporcionan propiedades espectrales de la vegetación en diversas escalas espaciales y son capaces de abordar no sólo el riesgo de incendios a escala

local, sino, el riesgo a escala regional o mundial (Chuvieco et al., 2003; Chuvieco et al., 2004; Chuvieco 2002;Zarco-Tejada et al., 2003).



Figura 30: Ejemplo de imagen de sensor satélite MODIS (Moderate Resolution Imaging Spectrometer) en 36 bandas, rango espectral entre 405 y 14385 nm. , dependiendo de las bandas, resolución espacial de 250 a 1000 nm. (<http://www.satellite-images.com/imagery/modis.jpg>)

Se han desarrollado múltiples técnicas para conocer los rasgos de absorción de agua por la vegetación con estimaciones de contenido de agua a través de imágenes hiperespectrales (Gao et al. 1995, Ustin et al. 1998, Green et al. 1993, Roberts et al. 1997, Serrano et al. 2000)

Se ha incrementado el uso de reflectancias hiperespectrales como entradas a los modelos de inversión para la estimación de atributos biofísicos y bioquímicos de la vegetación (Atzberger 2000, Fernandes et al. 2004, Hu et al. 2000, Jacquemoud et al. 2000, Kotz et al. 2004, Meroni et al. 2004, Zarco-Tejada et al. 2001).

Pueden distinguirse cuatro categorías de modelos de transferencia radiativa (Atzberger 2003, Goel 1988, Roberts, G. 2001) modelos geométricos (Li & Strahler 1985), modelos de análisis (Suits, G. H. 1972, Verhoef, W. 1984), los modelos híbridos son una combinación de modelos geométricos y analíticos, los modelos geométricos son mejores para describir el dosel de vegetación mediante formas geométricas (conos o cilindros), que arrojan sombras sobre el terreno. Los modelos analíticos describen la cubierta como una capa horizontal y semi-infinita con elementos de vegetación representados a través de geometría y densidad. Estos modelos se utilizan con frecuencia para describir cubiertas homogéneas y densas, como los cultivos. Los modelos híbridos son una combinación de ambos.

Hemos aplicado diversos métodos empíricos para determinar CWC mediante datos de satélite basados en los cambios de reflectancia del infrarrojo cercano, e infrarrojo de onda corta(Riño et al. in press,). Funcionan bien para una localización específica, pero necesitan ser recalibrados cuando cambiamos de lugar.

Los modelos de transferencia radiativa y los datos de reflectancia de la cubierta vegetal, representan una amplia gama de condiciones biofísicas (Goel 1988). Por lo tanto, estos modelos, pueden ser aplicados para determinar CWC en diferentes lugares con condiciones de diversos ecosistemas.

Los modelos de transferencia radiativa más simples suponen que la vegetación forma una cubierta continua de capas de hojas. Modelos más sofisticados tienen cuenta el efecto de las sombras del árbol en la respuesta de reflexión en el espectro electromagnético, suponiendo que los árboles están distribuidos de forma homogénea y de igual tamaño. Otros modelos completos tienen en cuenta cubiertas heterogéneas con árboles de diferentes tamaños y una capa uniforme de sotobosque (Ceccato et al. 2002).

La simplificación de modelos nos obliga a hacer suposiciones, por lo que podríamos estar lejos de la realidad, pero son más fáciles de parametrizar, con menor número de variables de entrada.

Una de las limitaciones de los modelos de transferencia radiativa es que la inversión para obtener CWC es computacionalmente muy costosa. Estas limitaciones computacionales han quedado resueltas (Usero et al. 2005, Riaño et al. 2005, Xiao 2005) con el uso de redes neuronales artificiales y algoritmos genéticos. El modelo FLIM asume un dosel del bosque homogéneo, con representación de sombras de árboles (Rosema et al. 1995).

Este trabajo utiliza el modelo de transferencia radiativa FLIM para obtener CWC mediante datos obtenidos del satélite MODIS. El modelo fue seleccionado a pesar de su complejidad, pero simplifica el parametrizar las variables de la heterogeneidad en la distribución de los árboles. La computación evolutiva se aplicó para testear la sensibilidad de varios índices de vegetación, obteniendo CWC y nos proporcionó un modelo sólido para predecir esta la variable de humedad mediante datos de reflectividad.

En los últimos años, se han realizado nuevas técnicas para obtener la distribución de contenidos de agua e información espacial de las variables bioquímicas de vegetación en escalas locales, regionales y, finalmente, a nivel mundial a través de datos de Teledetección. Estos nuevos enfoques están relacionados con técnicas de Softcomputing e Inteligencia Artificial. En el artículo “Estimation of fuel moisture content using neural networks”, analizado en el capítulo anterior, se realizaron estimaciones de EWT y DM en muestras secas con redes de neuronas artificiales, y quedó demostrado que eran tan buenas o mejores, como otros métodos de prueba sobre el mismo conjunto de datos, tales como la inversión de los modelos de transferencia de radiación, pero con bastante menos carga computacional. Las estimaciones de DM en muestras frescas utilizando RNA ($r^2 = 0,86$) mejoraron significativamente los resultados de la inversión de modelos de transferencia radiativa ($r^2 = 0,38$). El uso de algoritmos genéticos (AG) mejora los problemas de optimización en Teledetección (Fang et al. 2003).

Fang en su artículo del 2003 estima el índice de área foliar (leaf area index, LAI) mediante la integración de un modelo reflectancia a nivel de dosel y técnicas de optimización con algoritmos genéticos. Este método fue usado para obtener valores de LAI a través de medidas de reflectancia in situ y datos atmosféricos

normalizados del satélite Landsat ETM+. De las 8 bandas que observa el satélite Landsat ETM+, cuatro combinaciones diferentes fueron probados para evaluar su eficacia. El impacto del uso del número de algoritmos genéticos también fue testado.

El principal objetivo de este trabajo es evaluar la exactitud de la estimación de LAI por medio de la computación evolutiva. En la siguiente sección se describe nuestro método de cálculo evolutivo. Por último, los autores presentan varios experimentos con éxito y sus conclusiones

Resultados experimentales

El objetivo de este experimento fue encontrar un índice que sea capaz de correlacionar tan próximo como sea posible el valor CwLAI, contenido de agua del índice de área foliar. Este índice estará formado por una combinación de datos de obtenidos más allá de las primeras siete bandas del sensor MODIS embarcado en satélite.

Los índices utilizados normalmente en la teledetección (por ejemplo, NDVI) no son útiles para este propósito. Estos índices carecen de correlación con el valor buscado de CwLAI. (Todos ellos tienen el coeficiente de determinación inferior a 0,1). Con el fin de encontrar este índice, dentro del conjunto de todos los posibles valores, que se forman por medio de la combinación de datos obtenidos por el sensor MODIS y sus numerosas combinaciones, hemos decidido utilizar Técnicas de Programación genética y, en concreto, operadores de reproducción.

Hemos utilizado un sistema de investigación en Computación evolutiva, basado en el lenguaje de programación Java (EJC), que desarrolló el Laboratorio de Computación Evolutiva (ECLab) en la universidad de George Mason en Virginia (EEUU). Para el desarrollo del sistema de prueba, hemos utilizado un conjunto de 1000 muestras obtenidas a través modelo de transferencia radiativa FLIM.

De este conjunto, 500 muestras se han utilizado en la fase de entrenamiento, los otros 500 casos se han utilizado para evaluar las soluciones obtenidas en esta fase de entrenamiento (ver Tabla 6).

Población	1000 individuos
operadores	Cruce y reproducción
Probabilidad de Cruce	0.9
Probabilidad de Reproducción	0.1
Tipo de Selección	Selección por torneo

Tamaño del torneo	10
Elitismo	Entre 1 and 5 individuos
Nodos finales	Siente bandas MODIS (M1 .. M7)

Tabla 6: Parámetros utilizados en la programación genética.

Las pruebas se han obtenido teniendo en cuenta, tanto los índices típicos de Teledetección como las bandas del sensor MODIS. Las modificaciones en la función de aptitud se han llevado a con el fin de incorporar los valores RMSE en el proceso de optimización.

En la figura 31 a la figura 34, podemos observar los índices que presentan la mayor correlación, entre todas las posibles diferentes ejecuciones, con el esquema evolutivo propuesto, cada una de ellas con su gráfico correspondiente. Los test de validez, se realizaron con el conjunto de datos de entrenamiento validado con los datos de prueba. Podemos observar los valores de RMSE, el grado de correlación de Pearson (r) y el coeficiente determinación (r^2). Cada figura está dividida en cuatro gráficos. Cada fila representa la misma expresión pero uno está usando un conjunto de datos de entrenamiento y el otro es de trabajo con una base de datos de prueba.

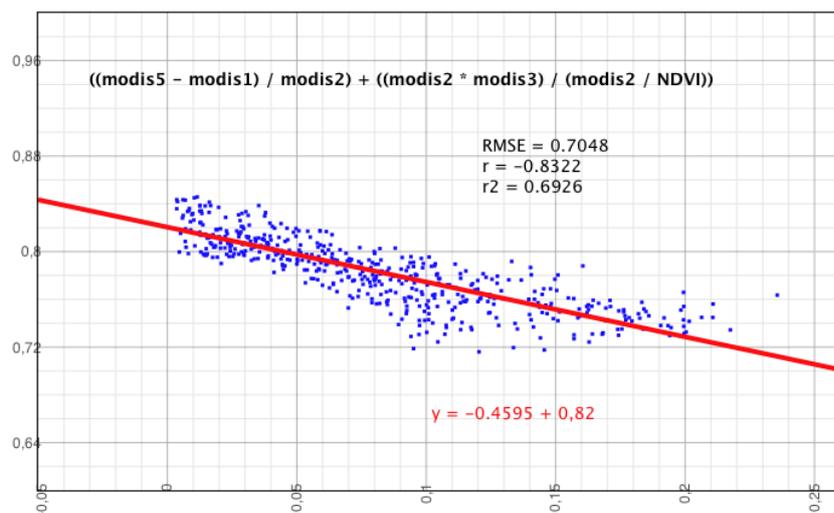


Figura 31: Correlación entre los valores $Cw \cdot LAI$ obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento)

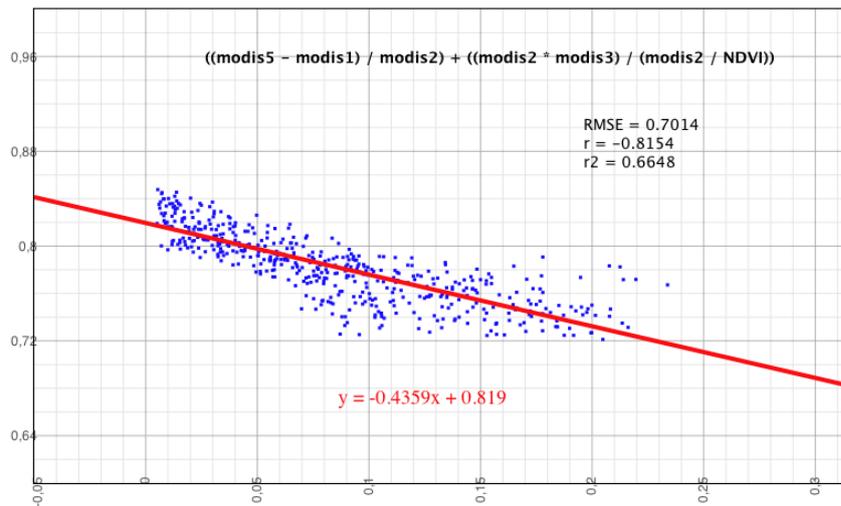


Figura 32: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)

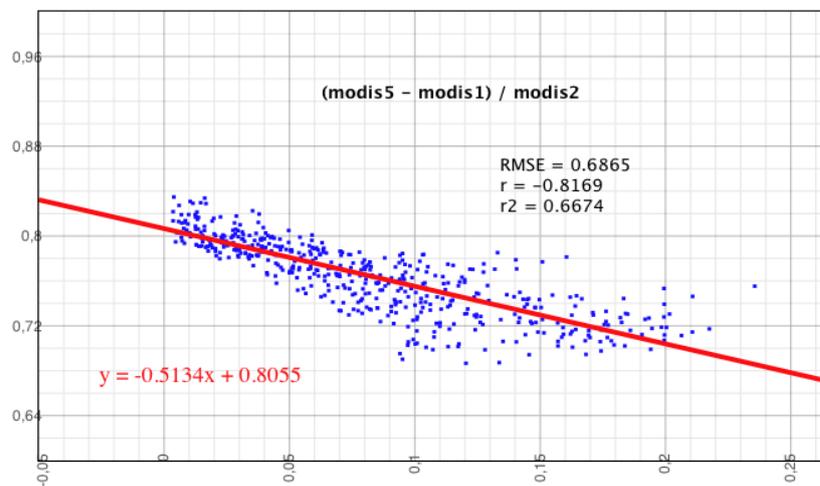


Figura 33: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento)

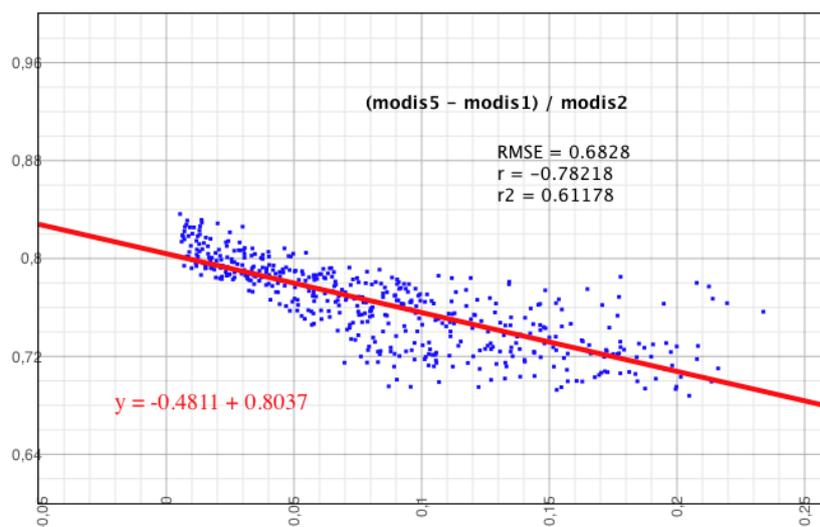


Figura 34: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)

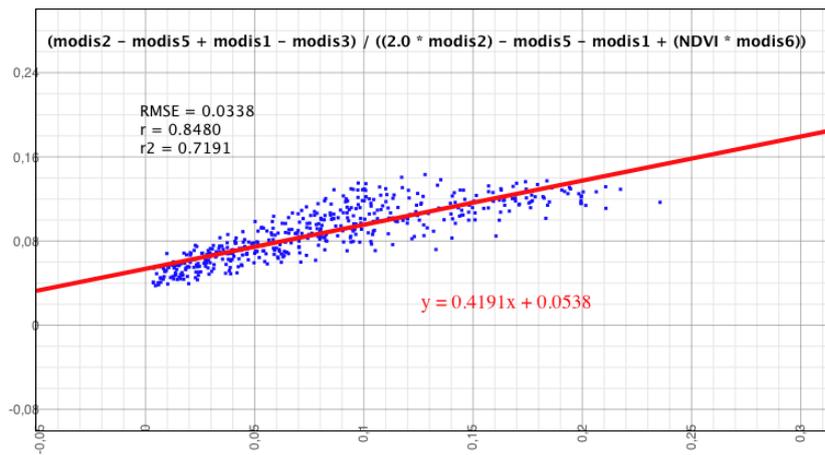


Figura 35: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento)

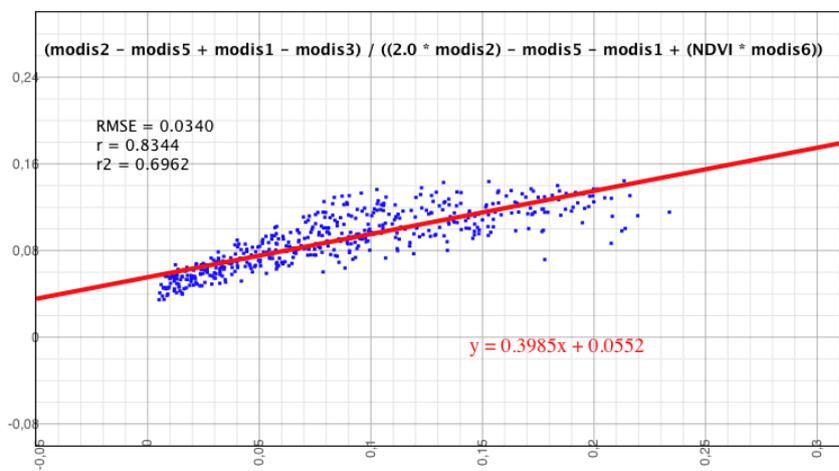


Figura 36: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)

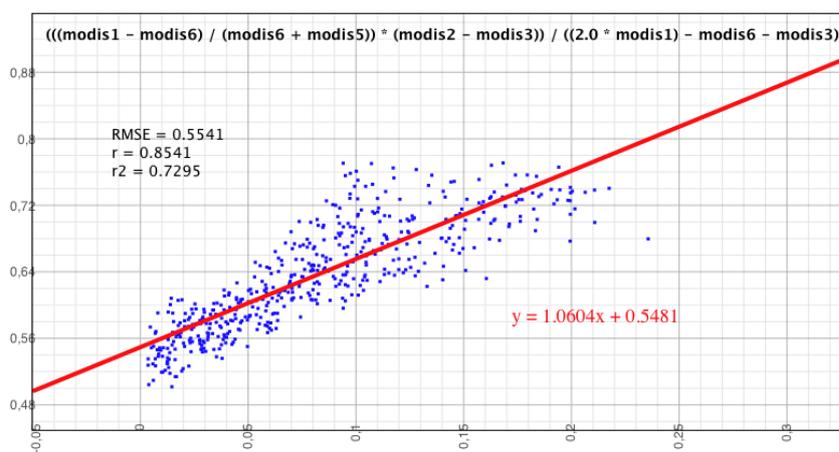


Figura 37: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento)

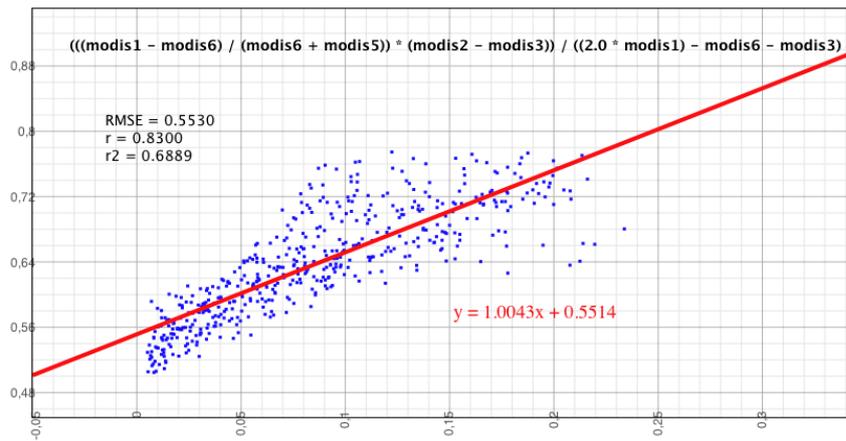


Figura 38: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)

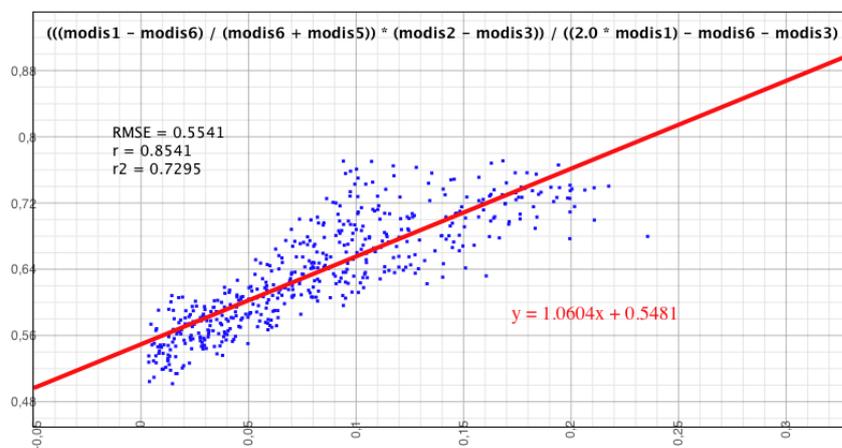


Figura 39: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento)

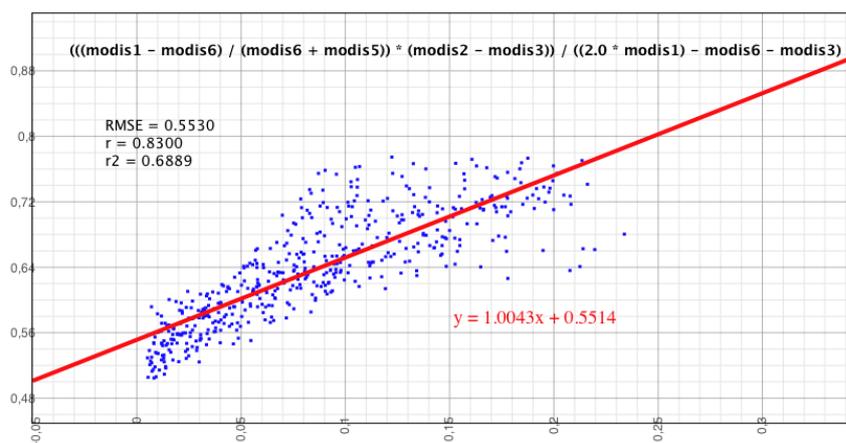


Figura 40: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)

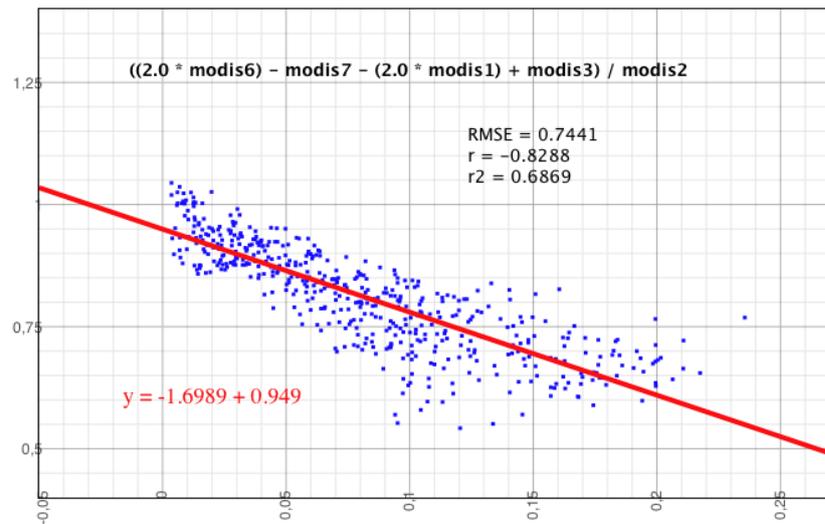


Figura 41: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS que señala la gráfica (datos de entrenamiento)

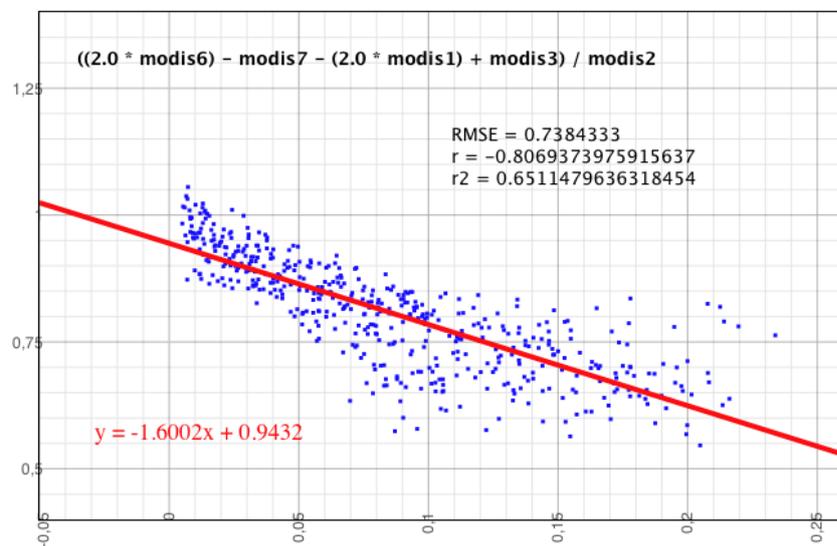


Figura 42: Correlación entre los valores Cw • LAI obtenidos con las bandas MODIS señala la gráfica (datos de validación)

Podemos apreciar doce resultados diferentes para el índice propuesto. Seis con datos de entrenamiento conjunto (lado izquierdo en las figuras) y seis con el conjunto de datos de prueba (lado derecho en las figuras), como hemos explicado antes. Idealmente, el mejor resultado sería el más cercano a 1, a pesar de los índices obtenidos son cerca de 0,7 (aceptable grado de correlación), es muy importante señalar que hemos tratado de no ejecutar expresiones complejas en el algoritmo genético, para obtener índices asequibles y útiles. Así, los resultados son fáciles de manejar por un sistema de teledetección real.

Discusión

Hemos demostrado cómo la Programación genética mejora la estimación del contenido de agua en vegetación. Esta información es vital para evaluar el riesgo de incendios forestales y su comportamiento. El siguiente paso para mejorar esta técnica evolutiva, es la incorporación de algunas características nuevas en el esquema evolutivo, por lo que estamos inmersos en nuevas experimentaciones empleando técnicas multipoblación que posibilitan la migración de individuos entre poblaciones con lo que aumentaremos la diversidad de los individuos de las diferentes poblaciones empleadas. También estamos trabajando en la inclusión de operadores simbióticos entre individuos de distintas poblaciones y en la inclusión de técnicas híbridas entre AG y PG ya que se ha demostrado su validez en la mejora de la variabilidad de los datos numéricos empleados como símbolos terminales.

Conclusiones y futuras líneas de trabajo:

En la presente Tesis Doctoral se realiza una exploración de los diferentes métodos basados en teledetección para estimar el contenido de humedad en diferentes tipos de coberturas de vegetación, siendo este un aspecto clave en el inicio y la propagación del incendio.

Hasta la fecha, las técnicas empleadas se basan en los modelos de transferencia radiativa, que simula la respuesta espectral de la vegetación basándose en un conjunto de parámetros biofísicos.

El objetivo es obtener el contenido de humedad a través de la inversión de estos modelos. La principal aportación de la tesis doctoral ha sido demostrar que, las técnicas basadas en inteligencia computacional están especialmente indicadas para trabajar con imágenes satélite, especialmente en prevención de incendios, ya que reducen los tiempos de respuesta casi a cero, una vez calibradas, y que pueden discriminar diferentes coberturas de vegetación analizando sus probabilidades de incendio.

Inicialmente, hemos llevado a cabo un análisis de la capacidad predictiva de las bandas de los sensores satélite (en este caso MODIS) mediante el uso de curvas ROC. Con dicho análisis validamos la hipótesis de que dichas bandas son capaces de estimar situaciones de alto o bajo peligro de incendio, y que, por lo tanto, podremos descubrir el modelo computacional capaz de estimar el peligro de incendio de un área. Basándonos en valores de nuestro conjunto de datos de FMC, fueron evaluados como alto peligro de incendio las situaciones con valores de FMC por debajo del 60%, y como bajo peligro de incendio los valores por encima de este umbral. En la tesis, se demuestra la necesidad de usar todas las bandas del satélite para poder estimar el peligro de incendio y, que por lo tanto, valida nuestra hipótesis de poder encontrar un modelo computacional que relaciona la información de las bandas del satélite con el riesgo de incendio.

El primer modelo computacional analizado se basa en redes neuronales artificiales. En la tesis se demuestra que una red neuronal del tipo perceptrón multicapa puede estimar EWT a nivel de hoja entrenando la red con muestras de la misma planta ($r^2 = 0,95$). Esto quedó demostrado con la validación cruzada que realizamos en la red neuronal con base radial para predecir EWT a nivel de hoja utilizando un amplio conjunto de datos independientes, de muestras reales procedentes de otras plantas y de diferentes especies ($r^2 = 0,89$). También fue posible que la red neuronal artificial de base radial fuera entrenada con el modelo de transferencia radiativa PROSPECT. La ventaja de utilizar un modelo de transferencia radiativa en el entrenamiento de la red, es que, estos modelos abarcan un área global, y se podrían aplicar a cualquier conjunto de datos.

Por otra parte, se ha llevado a cabo un estudio sobre la capacidad de la computación evolutiva en la mejora de la estimación del contenido de agua en vegetación. En concreto se ha demostrado la validez de estos modelos para obtener un índice que sea capaz de correlacionar tan próximo como sea posible el valor

CwLAI, contenido de agua del índice de área foliar. Este índice está formado por una combinación de datos de obtenidos más allá de las primeras siete bandas del sensor MODIS embarcado en satélite. El siguiente paso para mejorar esta técnica evolutiva, es la incorporación de algunas características nuevas en el esquema evolutivo, por lo que estamos inmersos en nuevas experimentaciones empleando técnicas multipoblación que posibilitan la migración de individuos entre poblaciones con lo que aumentaremos la diversidad de los individuos de las diferentes poblaciones empleadas. También estamos trabajando en la inclusión de operadores simbióticos entre individuos de distintas poblaciones y en la inclusión de técnicas híbridas entre AG y PG ya que se ha demostrado su validez en la mejora de la variabilidad de los datos numéricos empleados como símbolos terminales.

Como futuras líneas de trabajo podemos continuar con las técnicas de Softcomputing que no hemos evaluado todavía, como son la lógica difusa, razonamiento probabilístico y las Meta heurísticas combinando estas para conseguir modelos híbridos, como son, Meta heurísticas Evolutivas, Meta heurísticas de relajación, Meta heurísticas de búsqueda y Meta heurísticas individuales

Todas estas Meta heurísticas individuales avanzan en el espacio de búsqueda, basándose en la evaluación de la función objetivo del problema que tratamos de solucionar, lo que implica hacer valoraciones numéricas en un espacio claramente limitado.

Referencias

- Ahern F. J., Erdle T., MacLean D. A., Kneppeck I. D. 1991, A quantitative relationship between forest growth rates and Thematic Mapper reflectance measurements: *International Journal of Remote Sensing*, Volume 12, Issue 3 March 1991 , pages 387 – 400.
- Andrews P. L. 1986, BEHAVE: Fire Behavior prediction and modeling systems - Burn subsystem, part 1. General Technical Report INT- 194. Ogden, UT, US Department of Agriculture, Forest Service, Intermountain Research Station.
- Armitage, P. and Berry, G. 1994. *Statistical Methods in Medical Research*. Blackwell Scientific Publications.
- Atzberger, C. 2000. Development of an invertible forest reflectance model: The INFOR-Model. In Buchroithner (Ed.), *A decade of trans-European remote sensing cooperation. Proceedings of the 20th EARSeL Symposium Dresden, Germany, 14– 16. June 2000* (pp. 39– 44).
- Atzberger, C. 2003. *Möglichkeiten und Grenzen der fernerkundlichen Bestimmung biophysikalischer Vegetationsparameter mittels physikalisch basierter Reflexionsmodelle*. *Photogrammetrie, Fernerkundung, Geoinformation*, 1, 51–61.
- Bamber, D. 1975. The Area above the Ordinal Dominance Graph and the Area below the Receiver Operating Characteristic Graph. *J. Math Psychol*, 12:387_415
- Baret, F., Andrieu, B., and Guyot, G. 1988, A simple model for leaf optical properties in visible and near infrared: application to the analysis of spectral shifts determinism, in *Applications of Chlorophyll Fluorescence* (H. K. Lichtenthaler, Ed.), Kluwer Academic, pp. 345-351.
- Bovio, G., Quaglino, A. and Nosenzo, A. 1984 – Individuazione di un indice di previsione per il pericolo di incendi boschivi. *Monti e Boschi*. 35, 4: 39-44.
- Burgan, R.E. & R.C. Rothermel 1984. BEHAVE: Fire Behavior Prediction and Fuel Modeling System. Fuel Subsystem. GTR INT-167, USDA Forest Service, Ogden, Utah
- Burgan, R.E., Klaver, R.W. and Klaver, J.M., 1998. Fuel models and fire potential from satellite and surface observations. *International Journal of Wildland Fire*, 8(3): 159-170.
- Camia, A., Bovio, G., Aguado, I. and Stach, N., 1999. Meteorological fire danger indices and remote sensing. *Remote Sensing of Large Wildfires in the European Mediterranean Basin*. Springer-Verlag, Berlin, pp. 39-59
- Carrega P. 1990. *Climatology and Index of Forest Fire Hazard in Mediterranean France*. *International Conference Forest Fire Research. Coimbra. Proceedings*

- Carter, G.A., 1991. Primary and secondary effects of water content on the spectral reflectance of leaves. *American Journal of Botany* 78, pp. 916–924
- Carter, G. A. 1994. Ratios of leaf reflectances in narrow wavebands as indicators of plant stress. *International Journal of Remote Sensing*, 15, 697–704.
- Ceccato, P., N. Gobron, S. Flasse, B. Pinty & S. Tarantola 2002b, Designing a spectral index to estimate vegetation water content from remote sensing data: Part 1 - Theoretical approach. *Remote Sensing of Environment*. 82(2-3): 188-197
- Ceccato, P., B. Leblon, E. Chuvieco, S. Flasse y J. D. Carlson 2003: Estimation of Live Fuel Moisture Content, en *Wildland Fire Danger Estimation and Mapping. The Role of Remote Sensing Data* .Singapore, World Scientific Publishing: 63-90.
- Chandler, C., Cheney, P., Thomas, P., Trabaud, L., Williams, D. 1983. *Fire in Forestry. Vol 1: Forest Fire Behaviour and Effects*. John Wiley & Sons, New York.
- Chuvieco, E. y F. J. Salas 1996: Mapping the spatial distribution of forest fire danger using GIS. *International Journal of Geographical Information Systems*, 10: 333-345
- Chuvieco, E., D. Riaño, I. Aguado & D. Cocero 2002, Estimation of fuel moisture content from multitemporal analysis of Landsat Thematic Mapper reflectance data: applications in fire danger assessment. *International Journal of Remote Sensing*. 23(11): 2145-2162.
- Chuvieco, E., I. Aguado, D. Cocero y D. Riaño 2003: Design of an Empirical Index to Estimate Fuel Moisture Content from NOAA-AVHRR Analysis In *Forest Fire Danger Studies*. *International Journal of Remote Sensing*, 24: 1621-1637
- Chuvieco, E., Cocero, D., Riaño, D., Martin, P., Martnez-Vega, J., de la Riva, J., et al. 2004. Combining NDVI and Surface Temperature for the estimation of live fuels moisture content in forest fire danger rating. *Remote Sensing of Environment*, 92(3), 322-331
- Cocero, D., D. Riaño y E. Chuvieco 1998: Estimación de la humedad de la vegetación usando imágenes de satélite e índices meteorológicos del peligro de incendios. *Serie Geográfica*, 7: 59-72.
- Combal B., Baret F. & Weiss M. 2002, Improving canopy variables estimation from remote sensing data by exploiting ancillary information. Case study on sugar beet canopies. *Agronomie* Volume 22, Number 2, March 2002, Parameter estimation for crop models, 205 – 215
- Danson, F.M., C.S. Rowland & F. Baret 2003, Training a neural network with a canopy reflectance model to estimate crop leaf area index. *International Journal Of Remote Sensing*. 24(23): 4891-4905
- Danson, F.M. & P. Bowyer 2004, Estimating live fuel moisture content from remotely sensed reflectance. *Remote Sensing of Enviroment*. 92(3): 309-321

- Deeming, J. E., Burgan, R. E. y Cohén, J. D. 1977. The National Fire-Danger Rating System - 1978. Ogden, UT, USDA Forest Service
- DeLong E.R., DeLong D.M., & Clarke-Pearson D.L. 1988, Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach, *Biometrics*, 44, 837-845.
- Desbois, N., Deshayes, M. y Beudoin, A. 1997. Protocol for fuel moisture content measurements. A review of remote sensing methods for the study of large wildland fires (pp. 61-72). Alcalá de Henares: Departamento de Geografía, Universidad de Alcalá.
- Fang, H., Liang, S., and Kuusk, A. 2003. Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. *Remote Sensing of Environment*, 85, 257-270.
- Fernandes, R., Miller, J. R., Chen, J. M., & Rubinstein, I. G. 2004. Evaluating image-based estimates of leaf area index in boreal conifer stands over a range of scales using high-resolution CASI imagery. *Remote Sensing of Environment*, 89, 200– 216.
- Fourty, T. & F. Baret 1997, Vegetation water and dry matter contents estimated from top-of-the atmosphere reflectance data: a simulation study. *Remote Sensing of Environment*. 61: 34-45.
- Gao, B. -C., and Goetz, A. F. H. 1995. Retrieval of equivalent water thickness and information related to biochemical components of vegetation canopies from AVIRIS data. *Remote Sensing of Environment*, 52(3), 155-162.
- Gao, B.C. 1996. NDWI. A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58, 257-266
- Goel, N. S. 1988. Models of vegetation canopy reflectance and their use in estimation of biophysical parameters from reflectance data. *Remote Sensing Reviews*, 4, 1 – 212
- Green, D. M., y Swets, J. A. 1966. Signal detection theory and psychophysics. New York: John Wiley and Sons, Inc.
- Green, R. O., Conel, J. E., & Roberts, D. A. 1993. Estimation of aerosol optical depth and calculation of apparent surface reflectance from radiance measured by the airborne visible-infrared imaging spectrometer (AVIRIS) using MODTRAN2. SPIE Conference 1937: Imaging Spectrometry of the Terrestrial Environment, the International Society for Optical Engineering (SPIE), Orlando, FL., 14-15 April
- Hanley J.A., McNeil B.J. 1982 The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29-36. (PMID 7063747)
- Hanley, J. A. 1989. Receiver Operating Characteristic (ROC) Methodology: The State of the Art. *Critical Reviews in Diagnostic Imaging*, 29(3):307-335.

- Hardy, C.C. & R.E. Burgan 1999, Evaluation of NDVI for monitoring live moisture in three vegetation types of the Western U.S. *Photogrammetric Engineering and Remote Sensing*. 65: 603-610
- Henderson, F. M. and Lewis, A. J. 1998 (Eds.): *Principles and Application of Imaging Radar, Manual of Remote Sensing*, 3rd Edition, Vol. 2. John Wiley and Sons, New York.
- Hosgood, B., S. Jacquemoud, G. Andreoli, J. Verdebout, A. Pedrini & G. Schmuck 1994. The JRC Leaf Optical Properties Experiment (LOPEX'93). CL-NA-16095-EN-C, EUROPEAN COMMISSION, Directorate - General XIII, Telecommunications, Information Market and Exploitation of Research, L-2920, Luxembourg
- Hunt, E. R. y B. N. Rock 1989: Detection of changes in leaf water content using near and middle-infrared reflectances. *Remote Sensing of Environment*, 30: 43-54.
- Hu, B., Inannen, K., & Miller, J. R. 2000. Retrieval of leaf area index and canopy closure from CASI data over the BOREAS flux tower sites. *Remote Sensing of Environment*, 74, 255–274.
- ICONA 1990: V curso sobre defensa contra incendios forestales, Madrid ICONA.
- Jacquemoud, S. & F. Baret 1990, Prospect - a Model of Leaf Optical-Properties Spectra. *Remote Sensing of Environment*. 34(2): 75-91.
- Jacquemoud, S., Baret, F., Andrieu, B., Danson, F.M., & Jaggard, K. 1995. Extraction of Vegetation Biophysical Parameters by Inversion of the Prospect Plus Sail Models on Sugar-Beet Canopy Reflectance Data - Application to Tm and Aviris Sensors. *Remote Sensing of Environment*, 52, 163-172
- Jacquemoud, S., Ustin, S. L., Verdebout, J., Schmuck, G., Andreoli, G., & Hosgood, B. 1996. Estimating leaf biochemistry using the PROSPECT leaf optical properties model. *Remote Sensing of Environment*, 56(3), 194–202.
- Jacquemoud, S., Bacour, C., Poilve, H., & Frangi, J. -P. 2000. Comparison of four radiative transfer models to simulate plant canopies reflectance: Direct and inverse mode. *Remote Sensing of Environment*, 74, 471–481.
- JI W. CIVCO D. & KENNARD W. 1992, Satellite Remote Bathymetry: A New Mechanism for Modelling, *Photogrammetric Engineering & Remote Sensing*, 58(5), pp. 545-549, Bethesda (USA), ASPRS.
- Kotz, B., Schaepman, M., Morsdorf, F., Bowyer, P., Itten, K., & Allgower, B. 2004. Radiative transfer modeling within a heterogenous canopy for estimation of forest fire fuel properties. *Remote Sensing of Environment*, 92,
- Kraemer, H. C. 1992. *Evaluating Medical Tests*. Sage Publications.

- Krzanowski, W. J. and Hand, D. J. 2009. ROC curves for continuous data. BocaRaton, Fl.: Chapman and Hall/CRC.
- Lawson, B.D., Hawkes, B.C., 1989. Field evaluation of moisture content model for medium-sized logging slash. In: Proceedings of the 10th Conference on Fire and Forest Meteorology. Ottawa, Canada pp. 247-257.
- Leckie D.G. 1990, Advances in remote sensing technologies for forest surveys and management. Can. J. For. Res. **20** (1990), pp. 464–483.
- Li, X., & Strahler, A. H. 1985. Geometric-optical modeling of a conifer forest. IEEE Transactions on Geoscience and Remote Sensing, GE-23(5), 705-721
- McArthur A.G. 1973. Forest fire danger meter. Mark V. Forest Research Institute, Forest and Timber Bureau of Australia
- McClish, D.K. 1989. Analyzing a portion of the ROC curve. Medical Decision Making, 9, pp 190-195.
- McIntosh MW, Pepe MS.2002 Combining Several Screening Tests: Optimality of the Risk Score. Biometrics 2002; 58: 657–664
- Meroni, M., Colombo, R., & Panigada, C. 2004. Inversion of a radiative transfer model with hyperspectral observations for LAI mapping in poplar plantations. Remote Sensing of Environment, 92, 195– 20
- Metz, C.E. 1978. Basic principles of ROC analysis. Seminars in Nuclear Medicine 8, 183 - 298.
- Miller, J. R., Hare, E. W., & Wu, J. 1990. Quantitative characterization of the vegetation red edge reflectance: An inverted-Gaussian model. International Journal of Remote Sensing, 11, 1755–1773.
- Moody, J., Darken, C.J., 1989. Fast learning in networks of locally tuned processing units. Neural Computation 1 (2), 281-294.
- Neyman J, Pearson ES 1963. On the problem of the most efficient tests of statistical hypothesis. Philosophical Transactions of the Royal Society of London, Series A 1933; 231:289-337
- Paltridge, G.W. & J. Barber 1988, Monitoring grassland dryness and fire potential in Australia with NOAA/AVHRR data. Remote Sensing of Environment. 25: 381-394
- Peñuelas, J., Piñol, J., Ogaya, R., & Filella, I. (1997). Estimation of plant water concentration by the reflectance Water Index WI (R900/R970). International Journal of Remote Sensing, 18, 2869–28.
- Poggio, T. & F. Girosi 1989. A theory of networks for approximation and learning. Technical Report AI 1140, MIT, Cambridge, MA
- Pyne, S.J.; Andrews, P.L.; Laven,. 1996 “Introduction to wildland fire” New York, NY John Wiley & Sons.

- Riaño, D., Vaughan, P., Chuvieco, E., Zarco-Tejada, P.J., & Ustin, S.L. 2005. Estimation of fuel moisture content by inversion of radiative transfer models to simulate equivalent water thickness and dry matter content. Analysis at leaf and canopy level. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 819-826.
- Riaño, D., Chuvieco, E., Ustin, S.L., Salas, J., Rodríguez-Pérez, J.R., Ribeiro, L.M., Viegas, D.X., Moreno, J.M. y Fernández, H. 2007. Estimation of shrub height for fuel type mapping combining airborne LiDAR and simultaneous color infrared ortho image. *International journal of wildland fire*, 16 (3) 341–348.
- Riaño, D., P. Vaughan, E. Chuvieco, P.J. Zarco-Tejada & S.L. Ustin in press, Estimation of fuel moisture content by inversion of radiative transfer models to simulate equivalent water thickness and dry matter content. Analysis at leaf and canopy level. *IEEE Transactions on Geoscience and Remote Sensing*
- Rivero D., Vidal R., Dorado J., Rabuñal J. R. and Pazos A.. “Restoration of Old Documents with Genetic Algorithms”. *EvoWorkshops 2003*, LNCS 2611, pp. 432-443, 2003.
- Rivero D., Rabuñal J. R. , Dorado J, and Pazos A. “Using Genetic Programming for Character Discrimination in Damage Documents”. *EvoWorkshops 2004*, LNCS 3005, pp. 349-358, 2004.
- Roberts, D. A., Green, R. O., & Adams, J. B. 1997. Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS. *Remote Sensing of Environment*, 62(3), 223–240.
- Roberts, G. 2001. A review of the application of BRDF models to infer land cover parameters at regional and global scales. *Progress in Physical Geography*, 25(4), 483–511.
- Rosema, A., Verhoef, W., Noorbergen, H., and Borgesius, J. J. 1995. A new forest light interaction model in support of forest monitoring. *Remote Sensing of Environment*, 42, 23- 41
- San Miguel-Ayanz, J., J. D. Carlson, M. Alexander, K. Tolhurst, G. Morgan, R. Sneeuwjagt y M. Dudley 2003: Current Methods to Assess Fire Danger Potential, en *Wildland Fire Danger Estimation and Mapping. The Role of Remote Sensing Data* (editado por E. Chuvieco). Singapore, World Scientific Publishing: 21-61
- Shapiro, D. E. 1999. The interpretation of diagnostic tests. *Statistical Methods in Medical Research*, 8, 113-134
- Serrano L., Ustin S. L., Roberts D. A., Gamon J. A., & Peñuelas, J. 2000. Deriving water content of chaparral vegetation from AVIRIS data. *Remote Sensing of Environment*, 74(3), 570–581.
- Sol, B., 1990 - Estimation du risque meteorologique d’incendies de forêts dans le Sud-est de la France. - *Revue Forestière Française*, Nancy. n° spécial, 263-271
- Strahler A. H., Woodcock C. E. & Smith J. A. 1986, On the nature of models in remote sensing, *Remote Sensing Environ.* 20: 121-139

- Suits, G. H. 1972, The calculation of the directional reflectance of a vegetative canopy, *Remote Sens. Environ.*, 2, 117-125, 1972.
- Swets, J.A. and Pickett, R.M. 1982. Evaluation of diagnostic systems: Methods from signal detection theory. Academic Press, New York
- Tucker, C. J. 1980. Remote sensing of leaf water content in the near infrared *Remote Sensing of Environment*, 10, 23–32
- Ustin, S.L. et al. 1998, Estimating canopy water content of chaparral shrubs using optical methods. *Remote Sensing of Environment*. 65: 280-291.
- Ustin, S. L., Darling, D., Kefauver, S., Greenberg, J., Cheng, Y. -B., &Whiting, M. L. 2004a. Remotely sensed estimates of crop water demand. S.P.I.E. the international symposium on optical science and technology. 49th annual meeting, Denver, CO, 2–6 August
- Ustin, S. L., Jacquemoud, S., Zarco-Tejada, P. J., and Asner, G. 2004b. Remote Sensing of Environmental Processes, State of the Science and New Directions. in *Manual of Remote Sensing Vol. 4. Remote Sensing for Natural Resource Management and Environmental Monitoring*. ASPRS. John Wiley and Sons, New York 768p. +cd. (S.L. Ustin, vol. Ed.). ASPRS. John Wiley and Sons, New York, pages 679-730
- Van Wagner, C. E. 1967. Seasonal variation in moisture content of Eastern Canadian tree foliage and the possible effect on crown fires, *Forestry Branch, Canada*: 15 pp.
- Van Wagner, C. E. 1974. Comparison of American and Canadian Forest Fire Danger Rating Systems, Chalk River, Ontario, Canadian Forestry Service, Petawawa Forest Experiment Station, Information Report PS-X-2
- Van Wagner, C. E. 1985. Drought, Timelag and Fire Danger Rating. 8th National Conference on Fire and Forest Meteorology. Detroit; Michigan, Society of American Foresters: 178-185
- Vázquez, M. F. y B. J. Denore 1999: FOMFIS, un sistema para la cartografía de combustibles forestales a partir de imágenes de satélite de alta resolución, *Teledetección. Avances y Aplicaciones*, Albacete, España, Asociación Española de Teledetección: 279-282.
- Verhoef, W. 1984. Light scattering by leaf layers with application to canopy reflectance modeling: The SAIL model. *Remote Sensing of Environment*, 16, 125– 141
- Viegas, D. X., T. P. Viegas y A. D. Ferreira 1990. Characteristics of some forest fuels and their relation to the occurrence of fires. *International Conference of Forest Fire Research*. Coimbra: B.03: 1 - 13.
- Wood, S. N. 2003. Thin-plate regression splines. *Journal of the Royal Statistical Society B* 65 (1):95–114

- Wood, S. N. 2004. Stable and efficient multiple smoothing parameter estimation for generalized additive models. *Journal of the American Statistical Association* 99 (467):673–686
- Wood, S.N. 2006. *Generalized Additive Models. An Introduction with R*. Chapman and Hall/CRP Press, New York
- Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., et al. 2005. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. *Remote Sensing of Environment*, 95(4), 480-492.
- Yebra, M., Chuvieco, E. and Riaño, D., 2008. Estimation of live Fuel Moisture Content from MODIS images for fire risk assessment. *Agricultural and Forest Meteorology*. 148: 523-536.
- Youden, W.J. 1950. Index for rating diagnostic tests. *Cancer*, 3, 32-35
- Zarco-Tejada, P. J., Miller, J. R., Noland, T. L., Mohammed, G. H., & Sampson, P. H. 2001. Scaling-up and model inversion methods with narrow-band optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 39(7), 1491–1507.
- Zarco-Tejada, P. J., Rueda, C. A., & Ustin, S. L. 2003. Water content estimation in vegetation with MODIS reflectance data and model inversion methods. *Remote Sensing of Environment*, 85(1), 109–124
- Zarco-Tejada, P. J., Miller, J. R., Morales, A., Berjón, A., & Agûera, J. 2004. Hyperspectral indices and model simulation for chlorophyll estimation in open-canopy tree crops. *Remote Sensing of Environment*, 90(4), 463–476.

