

APLICACIÓN DE REDES NEURONALES ARTIFICIALES PARA DETERMINAR LA DISTRIBUCIÓN ESPACIAL DE LA HUMEDAD DEL SUELO EN UNA PEQUEÑA CUENCA DE DRENAJE. ESTUDIOS PRELIMINARES

M. Maneta y S. Schnabel

Grupo de Investigación Geoambiental. Dpto Geografía y OT. Área Geografía Física. Universidad de Extremadura. Avda Universidad s/n 10071Cáceres. maneta@unex.es

RESUMEN. La distribución espacial de la humedad superficial del suelo se ha determinado tradicionalmente mediante técnicas de interpolación o mediante métodos geoestadísticos a partir de valores de humedad medidos puntualmente. Los valores interpolados son casi siempre función exclusiva de los valores vecinos conocidos. En el método propuesto, se considera que el contenido de humedad del suelo en cuencas semiáridas es en gran medida explicado por factores topográficos locales tales como curvaturas horizontales y verticales, factores topográficos de entorno más amplio tales como el índice topográfico $\ln(a/\tan(b))$ (Beven y Kirkby, 1979) y por la vegetación. De esta manera, y a diferencia de las interpolaciones clásicas, se tienen en cuenta factores que alteran el principio de autocorrelación espacial tales como cambios bruscos en la cubierta vegetal debidos a cambios en los usos del suelo.

Ninguno de los factores anteriormente mencionados es por sí solo capaz de explicar la variación espacial de la humedad del suelo de manera satisfactoria, sin embargo sí pueden hacerlo parcialmente de forma combinada. Se ha podido ver en este estudio preliminar que la interrelación entre estos factores es de una forma compleja que no se ajustan correctamente a modelos lineales de regresión múltiple.

Nuestra propuesta trata de demostrar mediante este estudio preliminar que es posible realizar una predicción razonable de la distribución espacial de la humedad edáfica superficial, estudiar sus patrones espaciales y contrastar la importancia de cada variable a partir de un número limitado de mediciones mediante el uso de redes neuronales artificiales multicapa de tipo feedforward con aprendizaje mediante mecanismos de retropropagación del error utilizando principalmente variables derivadas de un modelo digital de elevaciones (MDE) y de fotografía aérea. Una vez entrenada y usada la red, se ha evaluado el error mediante un método *bootstrap* y se ha realizado un análisis de importancia (saliency analysis) de las variables utilizadas.

Este estudio ha usado el MDE con una resolución planimétrica de 5 metros y 79 medidas de humedad realizadas con TDR en una pequeña cuenca experimental

de dehesa en Extremadura con condiciones de humedad altas (invierno).

ABSTRACT. Spatial distribution of soil moisture has been traditionally determined using classical techniques such as interpolation or geostatistics whose predictions are usually based on neighbour points with known values. In this preliminary study, we consider that near-surface soil water content in semiarid catchments may be a function of topography (horizontal and vertical curvature and topographic wetness index) and vegetation. Unlike traditional interpolation techniques, this method is not based on neighbour points to forecast the unknown but it takes into account those local factors which should explain soil moisture. Thus, this method takes into account those factors which break the principle of spatial continuity such as changes in vegetation coverage or in land use.

Although none of these factors is able to explain spatial variations of soil moisture by itself, multivariate non-linear techniques such as artificial neural networks can improve the predictions.

The aim of this paper is to demonstrate the possibility of making good predictions of soil moisture distribution, to study its spatial pattern and to assess the influence of each variable using a limited number of sample points. This is done by using feedforward artificial neural networks with a backpropagation learning algorithm using topographical variables derived from a digital elevation model (DEM) and aerial photography. Once the network is trained, its accuracy was evaluated by a bootstrap method and a saliency analysis was performed in order to assess the relative importance of each variable in the explanation of the process.

The study has been developed using a 5 m resolution DEM and 79 moisture measurements of the near-surface soil layer with TDR in a small watershed. The wetness conditions in the catchment were high (winter)

1. Introducción

Conocer la distribución espacial de la humedad del suelo es de gran importancia para los estudios de hidrología, especialmente para los estudios centrados en la zona no saturada y en la superficie del terreno. Reflejo de esta relevancia es la gran cantidad de estudios realizados sobre este tema (Albertson y Kiely, 2001; Bárdossy y Lehmann, 1998; Ceballos et al., 1998; Gómez-Plaza et al., 2001; Herbst y Diekkruger, ; Jacques et al., 2001; Qiu et al., 2001a; Qiu et al., 2001b; Wilson et al., 2003).

A pesar de la abundante literatura generada, los procesos de distribución del agua en el suelo y los métodos para predecir el contenido de humedad edáfica siguen estando abiertos al estudio debido a la gran variabilidad espacial y temporal de este parámetro y a la cantidad de factores que lo determinan y que se influyen mutuamente. Además, estos factores varían en importancia dependiendo de la escala de trabajo, del estado de humedad del suelo, del dominio hidrológico que estemos tratando, de la vegetación, de la topografía, de las propiedades del suelo, de la profundidad de la zona saturada, etc (Ceballos y Schnabel, 1998; Florinsky et al., 2002; Gómez-Plaza et al., 2001). Hacer estudios a través de todo el rango de escalas, condiciones antecedentes y en el mayor número de lugares de condiciones climáticas e hidrológicas diferentes es importante para poder determinar con mayor exactitud la influencia real de cada elemento y poder así superar las conclusiones contradictorias a las que diferentes autores han llegado (Famiglietti et al., 1998).

La mayor parte de estos autores coinciden en considerar el relieve como un factor de gran valor explicativo para determinar los patrones de humedad del suelo, dando lugar al desarrollo de índices o métodos que ponen en relación las formas del terreno con propiedades del suelo (Beven y Kirkby, 1979; Florinsky et al., 2002; Qiu et al., 2001a; Western et al., 1999). Los modelos digitales del terreno están, desde los años 1950 siendo aplicados a estudios de ciencias de la tierra siendo una fuente de información de gran relevancia para el proceso de la información geográfica (Weibel y Heller, 1991). Una de las grandes ventajas que ofrecen estos modelos es que contienen información totalmente distribuida, espacialmente continua y fácilmente transformable. La mayoría de los parámetros topográficos más usados pueden derivarse de un modelo digital de elevaciones, pudiendo extraerse múltiples capas de nueva información desde la misma fuente, cosa que hace a esta fuente de información más atractiva y potente.

Dentro de los parámetros topográficos derivables de un modelo digital de elevaciones podemos distinguir entre los locales, que no traen implícita información sobre la ubicación de ese punto con respecto al conjunto del terreno y que se calculan a partir de un vecindario de puntos restringido tal como la pendiente, la orientación o los valores de las curvaturas locales horizontal o vertical. Son características del terreno que podríamos considerar hasta cierto punto, independientes de su posición dentro del conjunto de datos. Por otro lado tenemos parámetros del terreno que sí contienen de manera implícita información

de la posición de un punto respecto al conjunto de datos por cuanto el cálculo de estos valores implica conocer su situación con respecto al conjunto. Estas variables introducen en el modelo coherencia espacial y organización geográfica pues mantienen una relación espacial implícita entre los puntos del terreno. Ejemplos de este tipo de variables son las áreas de contribución de cada punto del terreno o el índice topográfico de Beven y Kirkby (1979) derivado de esta.

La vegetación es otro factor citado por los autores como de gran valor explicativo en la fragmentación espacial de la humedad del suelo (Breshears y Barnes, 1999; Pariente, 2002), sin embargo, obtener información espacialmente continua para todo el área de estudio es más complicado si no se pueden aplicar técnicas de teledetección. En nuestro caso usaremos fotografía aérea para tratar de introducir información continua sobre la cubierta vegetal en nuestro modelo.

Por otro lado, las técnicas usadas para estudiar las relaciones entre estos factores y el contenido de humedad del suelo han recaído casi siempre en relaciones lineales del tipo de regresiones simples o múltiples o en otros casos utiliza métodos geoestadísticos basados principalmente en los valores medidos cercanos más que en los factores explicativos subyacentes. Las redes neuronales artificiales son unas técnicas que se han venido aplicando con éxito en tareas de clasificación, reconocimiento y predicción en muchas áreas de conocimiento. Uno de los grandes potenciales de las redes neuronales es que son capaces de establecer relaciones funcionales entre variables dependientes e independientes aún cuando las relaciones son altamente no lineales debido a la variabilidad de la influencia de los factores y a su sensibilidad ante la influencia de otros factores no contemplados en el modelo, tal y como ocurre en los procesos de lluvia escorrentía (Hsu et al., 1993; Shamseldin, 1997).

Este estudio preliminar tiene dos objetivos: El primero será probar las capacidades de un modelo de red neuronal para determinar el patrón espacial de la humedad edáfica superficial a través de variables fáciles de obtener de manera distribuida y continua para todo el terreno. El segundo objetivo será comprobar su valor como herramienta explicativa estudiando la estructura de la red para averiguar la importancia relativa de cada variable en la explicación de los resultados.

2. Zona de estudio

El estudio se ha realizado en la cuenca experimental Parapuños (Fig. 1 y 2) situada 24 kilómetros al noroeste de la ciudad de Cáceres con una superficie de unas 100 ha bajo explotación de dehesa. Se asienta sobre la superficie de erosión finimiocena que forma la Penillanura Central Cacerense (Gómez-Amelia, 1985). El sustrato geológico es principalmente el complejo esquisto-grauváquico dominante en toda la penillanura al que se le superponen restos de antiguos pedimentos formados por acumulación de cantos rodados de cuarcita de tamaño variable en una

matriz limo-arenosa no cementada y rojiza. Morfológicamente, el sustrato esquisto-grauváquico coincide con las partes bajas de la cuenca, de topografía suave y escasa pendiente. Las zonas más elevadas y de mayor pendiente coinciden con los restos de los antiguos pedimentos.



Fig. 1. Ubicación de la zona de estudio (cuenca Parapuños)

Los suelos de la cuenca están en su mayoría escasamente evolucionados, sobre todo los que se asientan sobre los depósitos aluviales de casi 1 metro de espesor situados en los fondos de vaguada. En general, dominan los leptosoles y cambisoles cuyos espesores no superan los 50 centímetros. Los suelos más desarrollados se sitúan sobre los pedimentos y pueden ser clasificados como acrisoles crómicos (FAO, 1990).



Fig. 2. Cuenca experimental Parapuños (Cáceres)

El principal uso de la cuenca es el ganadero junto con alguna zona cultivada con forrajes para abastecer el ganado cuando no hay pastos. El arbolado disperso y formado por encinas (*Quercus Rotundifolia*). El estrato arbustivo es abundante en las zonas de mayor pendiente y está principalmente formado por retama (*Retama Sphaerocarpo*), escoba (*Cystisus Multiflora*) y cantueso (*Lavandula Stoechas*).

Respecto al régimen de lluvias y temperaturas, la zona está dentro del clima mediterráneo pero matizado por

influencias atlánticas y continentales. La temperatura media anual está en torno a los 16° y la precipitación es de 512 mm (Schnabel, 1998). Ambos parámetros están sujetos a fuertes variaciones tanto estacionales como interanuales. Utilizando el índice de aridez de la UNESCO (1979), el área de estudio se clasifica como semi-árida.

La cuenca no tiene cursos permanentes, el principal drenaje lo forman dos cauces que vienen a unirse unos 300 metros antes de la salida de la cuenca. En general la zona está bien drenada aunque hay zonas hacia la parte baja de la cuenca tendentes a quedar encharcadas durante las estaciones húmedas.

3. Material y métodos

El problema planteado en este estudio se ha abordado en términos equiparables a los de cantidad de masa y cantidad de movimiento (*momentum*) del agua. Las zonas tendentes a acumular mayor cantidad de agua con menor energía serán zonas con mayor probabilidad de tener altos contenidos de humedad. Similarmente, las zonas donde el agua diverge y el terreno le imprime aceleración serán zonas con menor probabilidad de acumular humedad. Para ello se han utilizado tres modelos digitales del terreno. Los dos primeros modelos reflejan las curvaturas locales del terreno. La curvatura vertical estima el grado de concavidad o convexidad del terreno en la dirección de la máxima pendiente. Refleja así la aceleración local del agua en ese punto. Valores mayores de cero indican zonas convexas de aceleración local del flujo y valores negativos indican concavidades que frenan el flujo facilitando la infiltración. La curvatura horizontal, por otro lado, es el grado de concavidad o convexidad en la dirección perpendicular a la de máxima pendiente por lo que detecta las concavidades (valores negativos) o convexidades (valores positivos) locales de convergencia o divergencia del flujo. Con estas dos variables incluimos en nuestro modelo una manera de evaluar la cantidad de movimiento (curvatura vertical) y la concentración de masa (curvatura horizontal) locales.

La tercera variable es el índice topográfico de humedad, que es una función del área de contribución (que refleja por tanto la cantidad de agua que convergerá en el punto en cuestión) y la pendiente del terreno (que sería expresión de la cantidad de movimiento de agua en ese punto), pero esta vez con una expresión que tiene en cuenta un entorno más amplio, introduciendo así el factor de coherencia geográfica que comentábamos en la introducción.

Todos los modelos digitales del terreno serán explicados de manera más formal en la sección 3.1.

Finalmente se ha incluido una variable que trata de reflejar la influencia de la cubierta vegetal, factor que hay que tener muy en cuenta ya que altera notablemente todo el planteamiento teórico anterior. Dada la dificultad para obtener fácilmente información fiable y totalmente distribuida sobre el terreno de la cubierta vegetal, hemos optado por aproximarla usando una ortofotografía aérea en tonos de gris de 8 bits (256 tonos de gris) de manera que

los suelos desnudos se ven claros (valores digitales menores) y las zonas con mayor cubierta vegetal aparecen más oscuras en función de la densidad de la vegetación (valores digitales más altos). Obviamente, se trata de una simplificación de la cobertura dado que existen tres estratos vegetales superpuestos, pero es un sustituto que da una aproximación útil para el propósito empleado.

Para establecer la relación entre los parámetros predictores anteriores y el contenido volumétrico de humedad de la superficie del suelo, se han realizado medidas de los primeros 15 cms del suelo siguiendo un sistema de transectos de longitudes variables y orientados siguiendo la dirección de máxima pendiente del terreno (fig. 5). En la fig. 3 se puede ver la distribución de

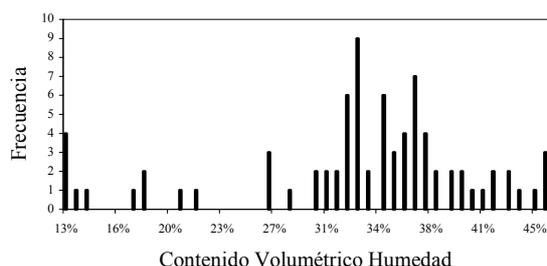


Fig. 3. Histograma con la distribución de la muestra de 79 puntos medidos

frecuencias de la muestra. Se midieron un total de 79 puntos con un TDR Tektronix 1502C (Dirksen, 1999). Las medidas se hicieron en febrero de 2003, cuando la cuenca se encontraba en estado húmedo tras un periodo de lluvias, aunque no había llovido inmediatamente antes de la campaña de muestreo.

3.1. Modelos digitales del terreno

El modelo digital de elevaciones (MDE) que ha servido de base para derivar el resto de los parámetros topográficos se ha realizado a partir de la restitución de fotografías de un vuelo a escala 1:6000 realizado en junio de 2002. El MDE se ha interpolado usando el inverso del cuadrado de la distancia a los datos vecinos más próximos. La resolución para el modelo usado ha sido de 5 m. A este modelo se ha aplicado un filtro de media para generalizar las formas y suavizar pequeñas singularidades del terreno que no son significativas para el objetivo del estudio (figura 5).

A partir de este modelo de elevaciones se han derivado el resto de los parámetros topográficos utilizados generando así tres modelos digitales del terreno: modelo de curvatura vertical (K_v), modelo de curvatura horizontal (K_h) e índice topográfico (TI).

El modelo K_v refleja la curvatura local del terreno en la dirección de la pendiente. Esta curvatura expresa la relativa aceleración o deceleración local del flujo. Su cálculo se define como (Florinsky et al., 2002):

$$K_v = \frac{p^2 r + 2 p q s + q^2 t}{(p^2 + q^2) \sqrt{(1 + p^2 + q^2)^3}} \quad (1)$$

donde r, t, s, p y q son las derivadas parciales de la altura (z) con respecto a las direcciones x e y calculadas en un kernel móvil de 3x3 sobre el MDE:

$$\begin{aligned} r &= \frac{\partial^2 z}{\partial x^2} & t &= \frac{\partial^2 z}{\partial y^2} & s &= \frac{\partial^2 z}{\partial x \partial y} \\ p &= \frac{\partial z}{\partial x} & q &= \frac{\partial z}{\partial y} \end{aligned} \quad (2)$$

El modelo K_h sería la curvatura del terreno en la dirección de la orientación (ortogonal por tanto a K_v). Esta curvatura es expresión de la convergencia o divergencia del flujo. Su cálculo viene definido por:

$$K_h = \frac{q^2 r + 2 p q s + p^2 t}{(p^2 + q^2) \sqrt{1 + p^2 + q^2}} \quad (3)$$

donde r, t, s, p y q están definidas igual que anteriormente en (2). Ambas expresiones, K_v y K_h tienen dimensiones de $[L^{-1}]$.

El modelo TI viene dada por la conocida expresión (Beven y Kirkby, 1979)

$$TI = \ln(CA / \text{tg}(\beta)) \quad (4)$$

donde CA es el área de contribución para el punto de cálculo y $\text{tg}(\beta)$ es la pendiente local del terreno. Este modelo expresa parcialmente el volumen relativo y la cantidad de movimiento del flujo que pasará por cada punto del terreno.

3.2. Redes neuronales

Las redes neuronales multicapa de tipo *feedforward* con aprendizaje por algoritmo de retropropagación (Rumelhart et al., 1986) son un tipo de estructura de computación paralela, donde muchas pequeñas unidades de cálculo llamadas neuronas están masivamente interconectadas con la capa anterior, de donde reciben la información, y con la capa posterior, hacia donde la transmiten. Es una forma de computación que trata de emular el funcionamiento de las neuronas biológicas en el sentido de que cada neurona tiene unos impulsos de entrada que estimula la neurona y la hace disparar por su axón un impulso proporcional al

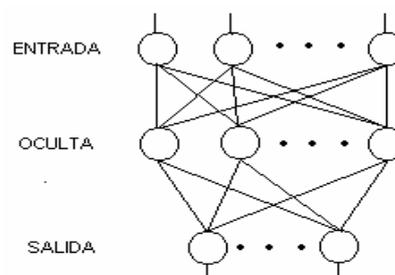


Fig. 4. Estructura de una red neuronal artificial feedforward

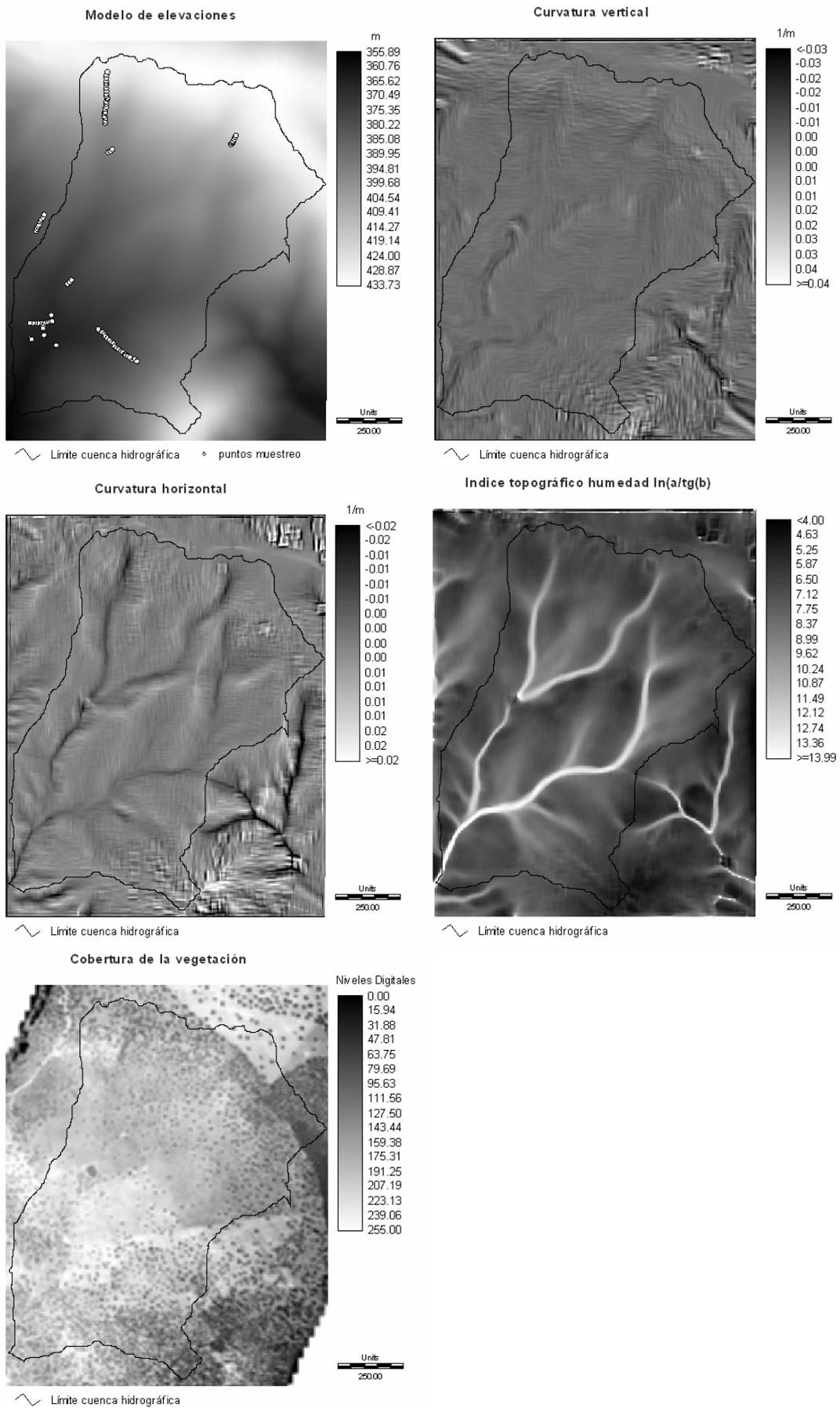


Fig. 5. Modelos digitales de elevación con puntos de muestreo y variables utilizadas en el modelo de red neuronal: Curvatura

estímulo recibido. Las redes neuronales artificiales de este tipo (fig.4) están constituidas por una primera capa de entrada (neuronas perceptoras) con tantas unidades como elementos tenga el vector de información que se va a introducir. Esta capa está conectada normalmente a una o más capas ocultas donde se realizan los cálculos y finalmente la información transformada llega una capa de salida de donde se obtienen los resultados. El almacenamiento del conocimiento adquirido por la red no está ubicado en ningún elemento concreto sino que está distribuido entre todas las neuronas a través del ajuste de unos coeficientes (pesos numéricos o pesos sinápticos). Esta red comienza con unos pesos aleatorios y aprende ajustando esos valores iterativamente hasta que llega a un estado estable donde es capaz de establecer una relación funcional dentro de un error objetivo prefijado entre la información de entrada y el conjunto de ejemplos que se le muestran a la red en su fase de aprendizaje. La versatilidad de este tipo de estructuras reside en su capacidad para detectar y establecer relaciones no lineales complejas entre las variables independientes (entradas) y las dependientes. El ajuste de los pesos que permite esto se hace a través de un mecanismo de retropropagación del error desde la capa de salida y hacia las capas ocultas después de comparar el resultado obtenido por la salida de la red con el valor correcto dado por el vector de valores del conjunto de ejemplos durante la fase de entrenamiento. El reparto del error a través de las neuronas se hace en función de la *culpa* de cada una de ellas en ese error, es decir, las neuronas con mayores pesos sinápticos y con mayor sensibilidad en la relación entre estímulo y respuesta serán a las que se les asigne una mayor parte del error. Este proceso, durante la fase de aprendizaje, se repite presentando sucesivamente a la red el conjunto de ejemplos de manera que se minimice una función objetivo como por ejemplo la raíz del error cuadrático medio (RMSE).

Formalmente el mecanismo de proceso de cada neurona es simple:

$$\text{Out} = f(\mathbf{w}^t \mathbf{x}) \quad (5)$$

donde Out es la salida de la neurona, $\mathbf{w}^t \mathbf{x}$ es el producto escalar entre el vector transpuesto de pesos sinápticos

$$\mathbf{w} = [w_0, w_1, w_2, \dots, w_m]^t \quad (6)$$

y el vector de entradas a la neurona

$$\mathbf{x} = [-1, x_1, x_2, \dots, x_m] \quad (7)$$

El subíndice m indica el número de entradas a la neurona (igual al número de neuronas de la capa anterior) y w_0 es un valor llamado de umbral y que da mayor flexibilidad a la red ya que permite ajustarse para disminuir el sesgo. En la práctica este valor es tratado como un peso más que va multiplicado siempre por una entrada de -1 (indicado como primer elemento en (7)). Viene a ser semejante al término que marca el corte con el eje y en una regresión lineal.

Finalmente, la función f que transforma el escalar resultante en la salida de la neurona normalmente es una función sigmoideal

$$f(w^t x) = \frac{1}{1 + e^{-(w^t x)}} \quad (8)$$

Esta función produce salidas dentro del rango [0,1]. La salida de cada neurona es una de las entradas para cada neurona de la siguiente capa tal y como se ve en la fig. 4 excepto en la primera capa donde la entrada de cada neurona es el elemento correspondiente del vector con las variables independientes y la salida es la función identidad. La capa de salida también es excepción ya que el resultado del cómputo neuronal no es entrada de otra neurona sino que es la salida de la red.

La estructura de la red neuronal aplicada en este estudio tiene una capa de entrada de cuatro neuronas (para los parámetros K_v , K_h , TI y veget) una capa oculta con 13 neuronas y una capa de salida con una sola neurona que nos dará la predicción del contenido volumétrico de humedad del suelo. Tanto el conjunto de datos de entrada (parámetros) como el conjunto de datos de ejemplo ha sido normalizado en el rango [0.1, 0.95]. La normalización de los datos en el rango [0, 1] es necesaria para evitar desbordamientos en los cálculos exponenciales y para mejorar la eficiencia de la red. Si además se normalizan en el rango propuesto se evitan problemas de saturación en la función sigmoideal para los valores extremos y se mejora la eficiencia del algoritmo de retropropagación, ya que al estar basado en cálculo de derivadas, trabaja muy lento en las colas de la función sigmoideal donde la pendiente es pequeña (Shamseldin, 1997).

3.3. Validación y cálculo del error: *bootstrap*

Para estimar el error real de predicción de una red neuronal normalmente se prueba la red sobre un conjunto de valores conocidos pero no usados en la fase de entrenamiento (Twomey y Smith, 1995). En nuestro caso, el número de datos para optimizar la red es demasiado limitado como para no usar todos en la fase de entrenamiento, por lo que no es viable extraer un subconjunto de datos y usarlo como conjunto de validación.

Como no conocemos la población total, el error que podamos calcular sobre una muestra será el error aparente ($\bar{\text{err}}$). La diferencia entre el error aparente y el error real (err) es el sesgo (β) de manera que

$$\text{err} = \bar{\text{err}} + \beta \quad (9)$$

Para hallar este sesgo usaremos un método llamado *bootstrap* (Efron y Tibshirani, 1993). Es un método no paramétrico, por lo que no debemos preocuparnos por conocer la distribución de la población. El método infiere la población F a partir de la muestra f realizando una simulación Monte Carlo con repetidos remuestreos sobre f.

La ventaja del método es que utiliza todo el conjunto n de datos de la muestra f para construir la red definitiva y estima el sesgo haciendo remuestreos *con reposición* sobre toda la muestra. Se remuestrea aleatoriamente hasta que se obtienen B ($T^{*1}, T^{*2} \dots T^{*B}$) muestras de tamaño n. Para cada muestra bootstrap (T^{*b}) se construye un modelo de predicción $f(T^{*b}, x_i)$, por lo que deberemos construir B+1 redes neuronales diferentes (B modelos bootstrap más un modelo final con todos los datos del conjunto). Así, estimamos el sesgo a través de B muestras independientes de tipo bootstrap usando la siguiente expresión (adaptada de Twomey y Smith (1993))

$$\beta_b^* = \sqrt{\frac{\sum_{i=1}^n (y_i - f(T^{*b}, x_i))^2}{n}} - \sqrt{\frac{\sum_{i=1}^n (y_i - f(T^{*b}, x_i^*))^2}{n}} \quad (10)$$

donde x_i indica que se pasan por la red todas las variables independientes de la muestra f, x_i^* indica que se pasan solo las variables independientes de la muestra T^{*b} e y_i es el valor esperado conocido. El resto de los términos están definidos más arriba. Finalmente

$$\hat{\beta}_{boot} = \frac{1}{B} \sum_{b=1}^B \beta_b^* \quad (11)$$

Conforme $B \rightarrow \infty$, el sesgo estimado va tendiendo a su valor verdadero. Para propósitos prácticos, el número de muestras B recomendado varía según los autores pero suele estar entre 25 y 200. En este estudio se han construido 20 muestras de este tipo construyendo 20 modelos de validación más 1 modelo final, número que para este estudio preliminar se ha considerado suficiente.

3.4. Análisis de importancia de las variables: *saliency análisis*

Los modelos de redes neuronales han sido poco usados por los científicos debido a que son considerados modelos *black-box* sin valor explicativo y son más considerados como modelos predictivos para aplicaciones de ingeniería, donde interesa obtener un valor ajustado en la predicción sin preocuparse de los aspectos físicos que explican el funcionamiento del proceso (Hsu et al., 1993). Sin embargo, es posible abrir la caja negra y obtener información de carácter explicativo de estos modelos a través de los análisis de importancia o significación de las variables [*saliency análisis*] (Abrahart et al., 2001). La manera de desagregar la red para estudiar la influencia de cada variable sobre la predicción es haciendo un estudio de los pesos sinápticos (Garson, 1991; Nath et al., 1997). El método divide los pesos sinápticos de la capa oculta hacia la capa de salida en componentes asociados con cada nodo de entrada de la capa de neuronas sensoriales (capa inicial de entrada de variables independientes). El peso resultante asociado con cada entrada sería reflejo de su importancia. Para ello, el valor absoluto de cada peso que conecta la

capa oculta con la de salida es incorporado a los pesos que conectan la capa de entrada con la capa oculta a través de la expresión (Nath et al., 1997)

$$w^*_{ij} = \{|w_{ij}|/S_j\} \times (|w_{j0}|) \quad (12)$$

donde w^*_{ij} es el peso asociado a cada entrada y que es proporcional a su importancia, w_{ij} son los valores absolutos de los pesos de la capa de entrada a la capa de

salida, $S_j = \sum_{i=1}^p |w_{ij}|$ y w_{j0} es el absoluto de los pesos

de la capa oculta a la capa de salida. Los subíndices i y j son los elementos en w_{ij} : i = 1, 2...p número de entradas para cada j = 1,2 ...h nodo en la capa oculta. Estos valores w^*_{ij} se pasan a porcentajes y esto será la estimación de la relevancia de cada variables con respecto a las demás.

4. Resultados y discusión

El método propuesto se comparará con los resultados de aplicar un análisis de regresión a cada variable y con la aplicación de un modelo de regresión múltiple con todas las variables propuestas. En la tabla 1 podemos ver los resultados de este análisis de regresión simple y en la tabla 2 podemos ver los resultados de la regresión múltiple.

Tabla 1. Estadísticos de regresión lineal para cada variable. r = coef de correlación paramétrico de Pearson. R2 = Coeficiente de determinación. SEE = Error estandar de la estimación. t = t ratio para r. 77 grados de libertad. n = 79.

Variable	r	R ²	SEE	t
Kv	-0.41	16.82%	0.074	-3.95
Kh	-0.50	25.00%	0.070	-5.07
IT	0.46	21.32%	0.072	4.57
Veget	0.48	22.97%	0.071	4.79

Tabla 2. Estadísticos de regresión múltiple. t = t ratio. 74 grados de libertad. n = 79. coeficiente de correl ajustado r = 0.62. R2 múltiple ajustado = 38.46%.

Variable	coeficiente	t
Corte eje y	-0.0015	-0.019
Kv	-4.8076	-1.300
Kh	-6.5914	-1.211
IT	0.0093	1.584
Veget	0.0013	3.652

Podemos observar que las regresiones simples dan resultados satisfactorios, todos los r son significativos al 0.05 iguales o mayores que los reportados por otros autores (Famiglietti et al., 1998; Florinsky et al., 2002) y los coeficientes de determinación son relativamente altos. Podemos observar cómo la relación entre k_v y k_h es inversa debido a que cuando sus valores son menores que cero el flujo decelera o converge respectivamente y cuando son mayores que cero el flujo acelera o diverge según el tipo de curvatura.

Sin embargo, la explicación conjunta a través de regresión múltiple no da los resultados que se esperaban, donde tan solo el coeficiente de la vegetación supera el test t al 0.05 y el coeficiente de determinación múltiple no es demasiado satisfactorio. El RMSE calculado es 0.061 (6.1%).

En la fig. 6 podemos observar la regresión entre los datos obtenidos a través del modelo de red neuronal artificial propuestos y los datos observados. El RMSE de la estimación es de 0.0269 (2.7%), aunque éste es el error aparente estimado sobre el mismo conjunto de datos que se ha utilizado para la construcción del modelo. Aplicando (9), (10) y (11), haremos una estimación del sesgo y podremos acercarnos al error real. La estimación se ha hecho construyendo 20 modelos de validación que han arrojado un sesgo de 0.027 (2.7%) con una desviación estándar en la distribución del sesgo de 0.013 (1.3%). Finalmente, a través de (9), tenemos un RMSE de la predicción estimado de 0.054 (5.4%). En la fig. 7 se puede ver el mapa de la predicción generado por el modelo de red neuronal. Las zonas negras son zonas sin predicción debido a que alguno de los parámetros predictores estaba fuera del rango establecido por los parámetros usados en el entrenamiento y se ha evitado extrapolar estos puntos. En la fig. 8 observamos el histograma de la distribución de la humedad superficial estimada para toda la cuenca, que tiene una forma similar a la de la muestra (fig. 3), donde destaca la elevada proporción de valores extremos. En este mapa y en el histograma de distribución de la humedad se puede observar como hay una gran proporción de zonas de bajo contenido de humedad que coinciden con las zonas de mayor pendiente y una con altos valores de humedad que coincide con los bordes de los cauces. El resto del terreno se agolpa en humedades en torno al 35%, que son también humedades muy altas siguiendo una distribución bastante normal. En el mes del muestreo, el estado de humedad de la cuenca aún era relativamente alto, había grandes zonas de la cuenca encharcadas. Esto ocurría en todas las zonas

Datos simulados con ANN vs datos medidos

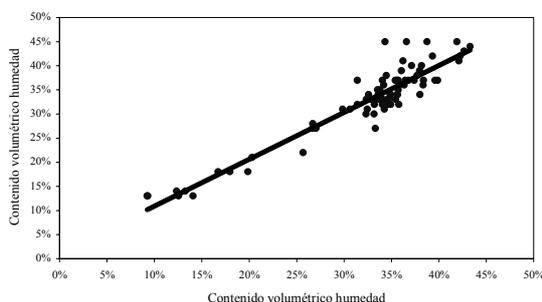


Fig. 6. predicción mediante red neuronal del contenido volumétrico de humedad vs mediciones

bajas o con drenaje más deficiente, zonas que contrastaban con las zonas secas de las pendientes de la cabecera. Así, se distinguen tres unidades diferenciables en las etapas húmedas de la cuenca coincidiendo con las zonas de mayor

pendiente, las zonas más bajas cercanas a los cauces y fondos de vaguada y la zona intermedia que sería el resto de la cuenca.

Una vez que tenemos cierta confianza sobre la estimación del modelo de red neuronal y sobre los parámetros elegidos como predictores podemos hacer un estudio de los pesos sinápticos para evaluar su grado de participación en la predicción. El número total de pesos de nuestra red, incluyendo los valores de umbral es de 79. El método se ha seguido tal y como se ha explicado en la sección 3.4 para las cuatro variables introducidas más la influencia del valor de umbral o sesgo, que como hemos visto, se trata igual que un peso ordinario asociado a una entrada de -1. La tabla 3 muestra los resultados del análisis expresando el porcentaje de participación o importancia de cada variable en la explicación de la distribución de la humedad para este caso.

Tabla 3. Resultados del análisis de relevancia de las variables en la explicación del resultado del modelo

Variable	Relevancia (%)
Kv	17.30
Kh	28.55
IT	17.02
Veget	20.22
Sesgo o valor Umbral	16.91
Total	100

En este caso parece que la capacidad local del terreno para concentrar la humedad es una de las variables que más influyen en el proceso.

Predicción del contenido volumétrico de humedad

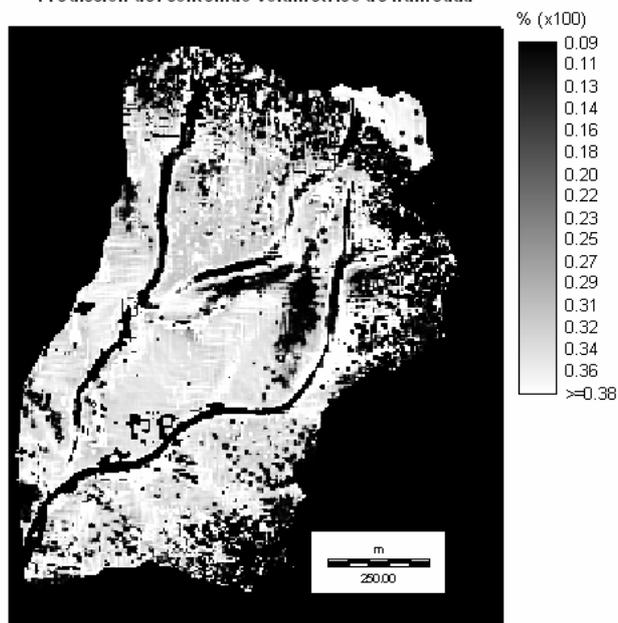


Fig. 7. Predicción del contenido de humedad del suelo por red neuronal artificial

Nótese que esta variable es también la que ha tenido el mayor coeficiente de determinación en la regresión lineal.

La vegetación tiene también un papel relevante en la explicación del proceso como era esperable. Con relevancias similares estarían el componente local de aceleración del agua (K_v) y el índice topográfico (IT), siendo este último ligeramente más bajo que K_v . El índice topográfico tiene una influencia moderada, siendo menos explicativa que otras variables como ya han encontrado otros autores (Famiglietti et al., 1998; Gómez-Plaza et al., 2001). Los estudios de estos autores se han desarrollado también en entornos semiáridos y han encontrado que el índice tiene menor capacidad predictiva que otros atributos topográficos, sobre todo durante condiciones de retraimiento de la humedad (secado) y cuando el índice se aplica para estimar la humedad en los primeros centímetros del suelo y no en toda la zona no saturada. Para Gómez-Plaza et al. (2001), el índice topográfico ha sido desarrollado en zonas húmedas donde el mecanismo de generación de escorrentía es por exceso de saturación del suelo y existe una conexión hidrológica permanente entre un punto del terreno y los que se encuentran vertiente arriba. Es por ello que proponen la creación y aplicación de otros índices para condiciones semiáridas incluyendo, por ejemplo, valores de radiación.

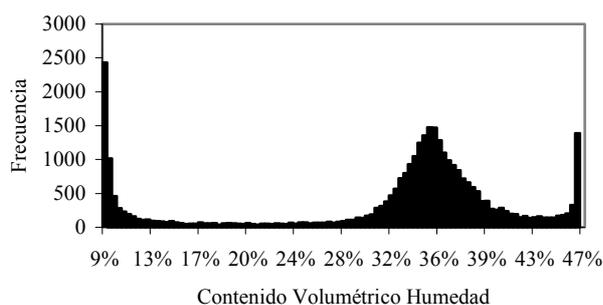


Fig. 8. Histograma de distribución de humedad para toda la cuenca

Como se indicó en la introducción, parece que en regiones semiáridas, el mecanismo de redistribución del agua en el suelo a través de flujo lateral subsuperficial solo operaría en momentos cortos cuando la cuenca está saturada y el resto del tiempo, los movimientos dominantes del agua en el suelo son los verticales hacia abajo a través de infiltración y hacia la atmósfera por evapotranspiración. Los parches de humedad suelen estar desconectados entre sí y están por tanto más influidos por la vegetación y por factores topográficos locales tales como K_v y K_h . La capacidad de los índices de humedad variará su representatividad sobre la predicción de la distribución de la humedad del suelo en función del estado de humedad de la cuenca.

En el caso de este estudio preliminar, la cuenca estaba en condiciones de humedad elevadas, aunque entrando en la fase de retraimiento, serán necesarios estudios más detallados, ocupando diferentes estadios en el proceso de humedecimiento-deseccación de la cuenca e incluyendo

más parámetros para obtener conclusiones más robustas y respaldadas.

5. Conclusiones

Los resultados de este estudio preliminar muestran que la aplicación de las redes neuronales artificiales para predecir la distribución espacial de la humedad da resultados satisfactorios y tienen mejor rendimiento que los modelos de regresión múltiple aún con un muestreo limitado. Esto puede ser debido a la no linealidad en las relaciones entre los parámetros predictores y el contenido volumétrico de humedad en la superficie del suelo y al hecho de que para las redes neuronales artificiales, al ser un método no paramétrico, las asunciones necesarias son más relajadas que para las regresiones múltiples.

Al mismo tiempo, se puede ver una cierta capacidad de la red para ofrecer información que ayuda a la explicación del proceso y permite valorar la importancia de cada variable introducida en el modelo.

Es un método que si bien requiere un mayor gasto computacional y de tiempo durante el proceso de entrenamiento, una vez calibrada la red procesar un volumen grande de datos para generar predicciones es sencillo y muy rápido. Tiene además la ventaja de ser un sistema bastante robusto ante información ruidosa o incompleta siendo capaz de seguir generando predicciones razonables.

Parece acertado pensar que el modelo puede aumentar su capacidad explicativa si se introducen nuevos parámetros elegidos de manera cuidadosa. En épocas donde la cuenca esté en proceso de secado, la introducción de un mapa de la distribución de la insolación o la radiación solar puede ser de gran valor. Información sobre la profundidad del suelo o la conductividad hidráulica para explicar la distribución de la humedad son también importantes variables a tener en cuenta debido a su probada capacidad explicativa (Gómez-Plaza et al., 2001) aún cuando esta información es más costosa de obtener de forma espacialmente distribuida.

Seguir realizando pruebas en diferentes épocas del año, introduciendo un mayor número de variables y sobre todo con un muestreo más extensivo debe ser motivo de más investigación para valorar de una manera más rigurosa los resultados de este estudio preliminar.

Agradecimientos. Este estudio ha sido realizado dentro del proyecto PROHISEM REN2001-2268-CO2-02 (MCYT) y RESEL (MMA). Los autores quieren expresar su agradecimiento al Dr Ángel M. Felicísimo por la revisión realizada sobre el primer borrador de este artículo.

Referencias

- Abrahart, R.J., See, L. y Kneale, P.E., 2001. Investigating the role of saliency analysis with a neural network rainfall-runoff model. *Journal of Computers and Geosciences*, 27: 921-928.
- Albertson, J.D. y Kiely, G., 2001. On the structure of soil moisture time series in the context of land surface models. *Journal of Hydrology*, 243(1-2): 101-119.

- Bárdossy, A. y Lehmann, W., 1998. Spatial Distribution of Soil Moisture in a Small Catchment. Part I: Geostatistical Analysis. *Journal of Hydrology*, 206: 1-15.
- Beven, K.J. y Kirkby, M.J., 1979. A physically-based variable contributing area model of basin hydrology. *Hydrological Science Bulletin*, 24(1): 43-69.
- Breshears, D.D. y Barnes, F.J., 1999. Interrelationships between plant functional types and soil moisture heterogeneity for semiarid landscapes within the grassland/forest continuum: a unified conceptual model. *Landscape Ecology*, 14: 465-578.
- Ceballos, A. y Schnabel, S., 1998. Comportamiento de la humedad del suelo en una pequeña cuenca hidrográfica de la dehesa extremeña (Guadalperalón, Cáceres). *Cuadernos de investigación geográfica. Universidad de la Rioja*, XXIV: 25-38.
- Ceballos, A., Schnabel, S., Gómez-Amelia, D. y Cerdá, A., 1998. Relación entre la escala espacial y escorrentía superficial en una pequeña cuenca hidrográfica semiárida ante condiciones contrastadas de humedad del suelo (Extremadura, suroeste de España). *Cuaternario y Geomorfología*, 12(1-2): 63-75.
- Dirksen, C., 1999. *Soil Physics Measurements*. GeoEcology. Catena Verlag, Reiskirchen, Germany, 154 pp.
- Efron, B. y Tibshirani, R.J., 1993. *An introduction to the bootstrap*. Chapman & Hall.
- Famiglietti, J.S., Rudnicki, J.W. y Rodell, M., 1998. Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. *Journal of Hydrology*, 210(1-4): 259-281.
- FAO, 1990. *Soil maps of the world, revised legend*, Roma.
- Florinsky, I.V., Eilers, R.G., Manning, G.R. y Fuller, L.G., 2002. Prediction of Soil Properties by Digital Terrain Modelling. *Journal of Environmental Modelling & Software*, 17: 295-311.
- Garson, D.G., 1991. Interpreting neural networks connection weights, *AI Expert*, pp. 47-51.
- Gómez-Amelia, D., 1985. *La penillanura cacereña. Estudio geomorfológico*. Serv Public Universidad Extremadura, Cáceres.
- Gómez-Plaza, A., Martínez-Mena, J., Albadalejo, J. y Castillo, V.M., 2001. Factors regulating spatial distribution of soil water content in small semiarid catchments. *Journal of Hydrology*, 253: 211-226.
- Herbst, M. y Diekkruger, B., Modelling the spatial variability of soil moisture in a micro-scale catchment and comparison with field data using geostatistics. *Physics and Chemistry of the Earth, Parts A/B/C*, In Press, Corrected Proof.
- Hsu, K.I., Gupta, H.V. y Sorooshian, S., 1993. Artificial neural network modeling of the rainfall-runoff process. *Water resources research*, 29(4): 1185-1194.
- Jacques, D., Mohanty, B., Timmerman, A. y Feyen, J., 2001. Study of time dependency of factors affecting the spatial distribution of soil water content in a field-plot. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(7-8): 629-634.
- Nath, R., Rajagopalan, B. y Ryker, R., 1997. Determining the saliency of input variables in neural network classifiers. *Journal of Computers Ops Res*, 24(8): 767-773.
- Pariente, S., 2002. Spatial patterns of soil moisture as affected by shrubs, in different climatic conditions. *Environmental Monitoring and Assessment*, 73: 237-251.
- Qiu, Y., Fu, B., Wang, J. y Chen, L., 2001a. Soil moisture variation in relation to topography and land use in a hillslope catchment of the Loess Plateau, China. *Journal of Hydrology*, 240(3-4): 243-263.
- Qiu, Y., Fu, B., Wang, J. y Chen, L., 2001b. Spatial variability of soil moisture content and its relation to environmental indices in a semi-arid gully catchment of the Loess Plateau, China. *Journal of Arid Environments*, 49(4): 723-750.
- Rumelhart, D.E., Hinton, G.E. y Williams, R.J., 1986. Learning internal representations by error propagation. In: D.E. Rumelhart y J.L. McClelland (Editors), *Parallel Distributed Processing*, vol 1. MIT Press, Cambridge.
- Schnabel, S., 1998. La precipitación como factor en los procesos hidrológicos y erosivos. In: S. Schnabel, D. Gómez-Amelia y A. Ceballos (Editors), *Norba. Revista de Geografía. Universidad de Extremadura, Cáceres*.
- Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall-runoff modelling. *Journal of Hydrology*, 199: 272-294.
- Twomey, J.M. y Smith, A.E., 1993. Nonparametric error estimation methods for evaluating and validating artificial neural network prediction models. In: Dagli, Burke, Fernández y Ghosh (Editors), *Intelligent Engineering Systems through Artificial Neural Networks*. ASME Press, pp. 233-238.
- Twomey, J.M. y Smith, A.E., 1995. Performance measures, consistency, and power for artificial neural network models. *Mathematical Computer Modelling*, 21(1/2): 243-258.
- UNESCO, 1979. *Carte de la répartition mondiale des régions arides*. UNESCO, Paris.
- Weibel, R. y Heller, M., 1991. Digital terrain modelling. In: D.J. Maguire, M.F. Goodchild y D. Rhind (Editors), *Geographical Information Systems: Principles and Applications*. Vol 1: Principles. Longman, Harlow, pp. 269-297.
- Western, A.W., Grayson, R.B., Blöschl, G., Willgoose, G.R. y McMahon, T.A., 1999. Observed spatial organization of soil moisture and its relation to terrain indices. *Water resources research*, 35: 797-810.
- Wilson, D.J. et al., 2003. Spatial distribution of soil moisture over 6 and 30 cm depth, Mahurangi river catchment, New Zealand. *Journal of Hydrology*, 276(1-4): 254-274.