

MODELIZACIÓN DE SISTEMAS ECOLÓGICOS MEDIANTE REDES NEURONALES

A. Bravo-Oviedo ¹ y G. Kindermann ²

¹ Dpto. Producción Vegetal y Recursos Forestales. Universidad de Valladolid. Avda. Madrid, 44. 34004-PALENCIA (España). Correo electrónico: abravo@pvs.uva.es

² Institut für Waldwachstumsforschung. Universität für Bodenkultur. Peter Jordan-strasse, 82. A-1190 VIENA (Austria).

Resumen

Las relaciones que existen en los sistemas ecológicos son objeto de estudio por parte de investigadores de todo el mundo. Los ecólogos han utilizado tradicionalmente en dichos estudios las relaciones lineales y las no lineales. En algunas ocasiones las relaciones no lineales no son obvias y presentan una gran dificultad a la hora de la modelización, dando lugar a complejos modelos cuya precisión no es siempre la deseada. En los últimos años, la inteligencia artificial se ha convertido en una técnica que permite tanto la gestión de grandes bases de datos como la utilización de algoritmos que, aunque complejos en su estructura, arrojan resultados fácilmente interpretables. El trabajo que se presenta ofrece, por una parte, una introducción a la teoría de las redes neuronales y, por otra, la aplicación de las mismas en la modelización ecológica.

Palabras clave: *Inteligencia artificial, Aprendizaje, Relaciones no lineales, Algoritmo Backpropagation*

INTRODUCCIÓN

La modelización de sistemas ecológicos es compleja y muchas veces parcial debido, en parte, a que no todas las variables independientes pueden ser medidas. En ocasiones la interacción entre variables o entre sistemas es tal que la modelización completa y precisa no es posible, por lo que se da un margen de error que consideramos aceptable en cada caso particular.

Para evitar los problemas de dificultad en la modelización, la inteligencia artificial (IA) nos proporciona una serie de técnicas capaces de hallar relaciones predictivas cuando no se conoce ninguna fórmula matemática que relacione variables dependientes e independientes. Entre esas técnicas las redes neuronales han sido utilizadas con más o menos fortuna en la modelización de sistemas ecológicos. El trabajo que se

presenta trata de explicar en qué consisten dichas redes y muestra algunas aplicaciones de las mismas en las ciencias de la vida.

LAS REDES NEURONALES ARTIFICIALES

Existen muchas definiciones de redes neuronales, casi tantas como trabajos en la materia. Algunas de ellas son:

- Redes interconectadas en paralelo y con organización jerárquica (KOHONEN, 1988a).
- Un modelo matemático compuesto por un gran número de elementos procesales organizados en niveles (HILERA & MARTÍNEZ, 1995).
- Ordenación secuencial de tres tipos de nodos o capas: nodos de entrada, nodos intermedios y nodos de salida que utilizan el cerebro huma-

no y su estructura para desarrollar una estrategia de procedimiento (HAIR et al., 1999).

- Intento de implementar en un ordenador estructuras y funciones similares a las capacidades de procesamiento de datos del cerebro humano (RECKNAGEL, 2001).
- Algoritmos generales de análisis de datos basados en un uso intensivo del ordenador (PEÑA, 2002).

Las redes neuronales (RN) constan de unidades que se agrupan en tres tipos de capas; la capa de entrada, asimilable a las variables independientes, la capa de salida que corresponde a la variable dependiente y una serie de capas intermedias denominadas capas ocultas. Cada unidad de una capa está unida a otra unidad de la siguiente capa mediante conexiones, que pueden ser hacia adelante, hacia atrás o mixtas. La figura 1 muestra la estructura de una red neuronal con conexiones hacia adelante. En un instante t cada unidad posee un estado de activación. Cuando una unidad de entrada i envía información x_i a una unidad j de una capa oculta, dicha información es ponderada por w_{ij} . La suma de todas las unidades de entrada ponderadas modifican el estado de activación de la unidad oculta mediante una función, conocida como función

de activación que, en la mayoría de los casos, es una función sigmoide del tipo logístico. La información almacenada ahora en la unidad oculta pasa a otra unidad k de otra capa oculta, este paso de información se realiza mediante una función de salida, que suele ser la función identidad. La información es ponderada ahora por w_{jk} y es incorporada a la función de activación de la unidad k que envía la señal a la siguiente unidad y así sucesivamente hasta la unidad de salida (HILERA & MARTÍNEZ, 1995). El tipo de red explicado corresponde a un modelo MLP (*Multilayer perceptron*).

El número de capas y el de unidades constituye la arquitectura o topología de la red. El concepto de red hace referencia a las conexiones que existen entre diferentes unidades o nodos. El término neuronal hace referencia a la estrategia de procedimiento de la red. El cerebro humano está formado por aproximadamente 10^{11} neuronas que se unen fisiológicamente con otras neuronas mediante pulsos electroquímicos, estas uniones o sinapsis son continuas las 24 horas del día, cuando descansamos, cuando hacemos deporte o cuando leemos un artículo miles de sinapsis se producen en nuestro cerebro. HEBB, (1949) señala que el aprendiza-

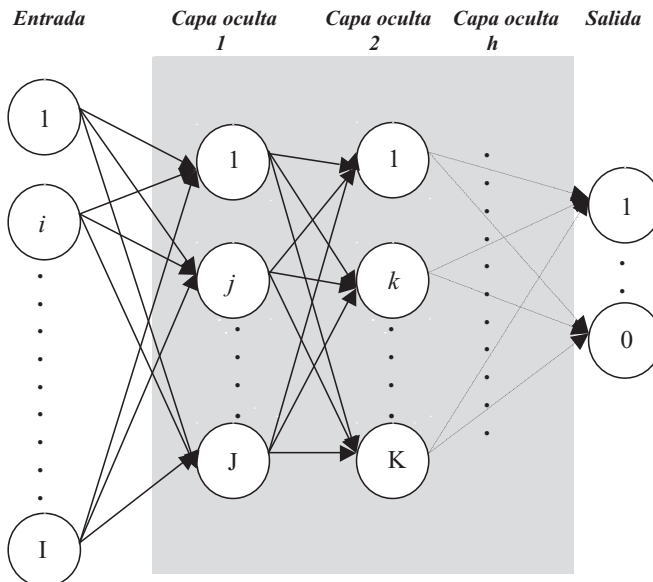


Figura 1. Topología de red neuronal con conexiones hacia adelante

je consiste en fortalecer la unión entre neuronas, para lograrlo nos entrenamos, por ejemplo cuando estudiamos repetidamente un tema de una oposición. Pues bien, las redes neuronales siguen este proceso. Cuanto más fuerte es la conexión entre unidades mejor es el resultado obtenido, es decir, cuanto mejor haya “aprendido” la red durante el “entrenamiento” al que la sometemos, menor error se obtendrá. La conexión entre unidades en una red neuronal se cuantifica mediante pesos, cuyo valor es asimilable al de los coeficientes de las variables independientes (HANGOS & CAMERON, 2001)

Ya hemos citado dos conceptos fundamentales de la construcción de redes neuronales, el aprendizaje y el entrenamiento. Hablaremos respectivamente de aprendizaje supervisado o no supervisado según se conozcan, o no, los valores de salida, *outputs*, de la red. Para entrenar necesitamos datos que, para su utilización en la red neuronal han de dividirse en dos grupos aunque algunos autores recomiendan tres (ZELL, et al., 1999): datos de entrenamiento (training set), datos de validación (validation set) que se utilizarán para determinar el funcionamiento de la red con patrones no entrenados y datos de comprobación (test set) para evaluar el comportamiento global de la red. La labor de los modelizadores es encontrar la arquitectura y el algoritmo más apropiado para los datos que se están analizando, cuidando que no se produzca un sobreentrenamiento de la red que produciría un incremento de error en los datos de validación.

El algoritmo de aprendizaje supervisado de una red con conexiones hacia adelante más utilizado por su sencillez y por resultados es, sin duda, el de retropropagación de errores conocido por BP (*Backpropagation*).

El algoritmo BP fue diseñado por RUMELHART et al. (1986) y busca minimizar el error medio cuadrático de la salida predicha de una red MLP con conexiones hacia adelante. Requiere que la función de activación sea no lineal y diferenciable (WEINGARTNER, 1999). A continuación se pone un ejemplo de cómo funciona el algoritmo BP siguiendo la notación de LEK & GUÉGAN (1999) para una red con una unidad de salida.

1. Diseñar la arquitectura de la red y el número máximo de iteraciones para evitar sobreentrenamiento

2. Proponer la tasa de aprendizaje η y el valor inicial de los pesos mediante números aleatorios pequeños.
3. Introducir el vector de valores de entradas (*inputs*) y el vector de valores de salida (*outputs*)
4. Calcular el input de las unidades ocultas:

$$a_j^h = \sum_{i=1}^n W_{ij}^h x_i \quad \text{Calcular el output de las unidades ocultas mediante la función de}$$

$$\text{activación } x_j^h = f(a_j^h) = \frac{1}{1 + e^{-a_j^h}}$$

5. Calcular el input de las unidades de salida:

$$a_k = \sum_{j=1}^L W_{jk} x_j^h \quad \text{Calcular el output final}$$

$$\hat{Y}_k = f(a_k) = \frac{1}{1 + e^{-a_k}}, \quad \text{donde L es el número}$$

de unidades ocultas y $k=1$.

6. Calcular el término de error para la salida:

$$\delta_k = \left(Y - \hat{Y} \right) f'(a_k) \quad \text{y el error para las unidades}$$

$$\text{de ocultas } \delta_j^h = f'(a_j^h) \sum_k \delta_k W_{jk}$$

7. Actualizar los pesos comenzando desde la

$$\text{capa de salida } W_{jk}(t+1) = W_{jk}(t) + \eta \delta_k x_j^h$$

Posteriormente y hacia atrás (*backpropagation*) se actualizan los pesos de la capa oculta

$$W_{ij}(t+1) = W_{ij}(t) + \eta \delta_j^h x_i$$

8. Repetir los pasos 3-7 hasta que el error

$$E_p = \frac{1}{2} \sum_{k=1}^M \delta_{pk}^2 \quad \text{sea aceptable donde p es}$$

cada uno de los vectores de entrenamiento presentados y M el número de unidades de salida, que en nuestro caso $M=k=1$

Nota: La tasa de aprendizaje del algoritmo, η , es la constante de proporcionalidad del gradiente descendiente de la función del error, que oscila entre 0 y 1. A mayor tasa de aprendizaje mayor es la modificación de los pesos en cada interacción, acelerando el proceso de aprendizaje (HILERA & MARTÍNEZ, 1995).

Existen multitud de redes neuronales. SARLE (1999) señala que “nadie sabe exactamente cuantas hay”, ya que cada semana se inventan nuevas redes y cita 47 algoritmos de aprendiza-

je combinados principalmente con dos tipos de arquitecturas: hacia adelante y hacia atrás.

APLICACIONES DE LAS REDES NEURONALES

En esta sección se muestran algunas de las aplicaciones de las redes neuronales en el estudio de sistemas ecológicos.

LEK et al. (1996) construyeron una red utilizando el algoritmo BP compuesta de tres capas. La capa de entrada constaba de 10 unidades *inputs* referidas al hábitat de la trucha en el estudio de su reproducción, 8 unidades ocultas en la capa oculta, determinadas de manera empírica y una única unidad de salida *output* que representaba la densidad de los alevines de trucha por metro lineal de corriente. La red fue comparada con un modelo de regresión múltiple lineal y encontraron mayor precisión con el modelo neuronal ($R^2=0.785$ para el modelo de red frente a 0.371 en la regresión múltiple con datos de validación).

En el caso de modelización forestal la mortalidad regular, aquella producida por factores no catastróficos, ha sido ampliamente estudiada mediante redes neuronales (GUAN & GERTNER, 1991a, b; 1995; HASENAUER & MERKL, 2000; KING et al., 2000). En uno de estos trabajos (GUAN & GERTNER, 1991a) se compara el modelo logístico de mortalidad presentado por BUCHMAN et al., (1983) para *Pinus resinosa* Ait. y una red neuronal con algoritmo BP, dos variables independientes del modelo logístico (diámetro normal e incremento diametral anual) y 11 unidades ocultas. Además prueban 5, 7, 10 y 20 unidades en la capa oculta y 7 unidades en la capa de entrada. Las unidades en la capa de entrada son el resultado de combinar las variables independientes en lo que llaman “patrones de actividades”, basados en la distribución Gaussiana utilizando cuatro clases diamétricas y tres clases de incremento. El resultado es una mejor estimación de la mortalidad con la red neuronal de 5 unidades ocultas y 7 *inputs*.

No siempre la red neuronal es el mejor modelo posible. Para el estudio de la ocurrencia del ave *Rhyacornis fuliginosus* en los ríos del Himalaya, MANEL et al., (1999), diseñaron una red con 32 unidades referidas a variables medioambientales

en la capa de entrada, 5 en la capa oculta y una en la capa de salida que indicaba presencia o ausencia de la especie. El algoritmo utilizado fue el BP y lo compararon con análisis discriminante y regresión logística. Los resultados más satisfactorios fueron encontrados para el modelo logístico seguido del modelo neuronal. Hay que señalar que en este caso sólo un tipo de red neuronal fue comparado pudiendo existir otro tipo de arquitectura o de algoritmo más apropiado para modelizar la presencia del ave en el río. Este tipo de comparación con varios algoritmos y distintos modelos neuronales fue llevado a cabo por HASENAUER et al. (2001) para el estudio de la mortalidad en masas de *Picea abies* en Austria. Los algoritmos utilizados fueron el BP, Rprop, y el SCG como modelos perceptrón multicapa, BP y Rprop en modelos de correlación en cascada y el LVQ como modelo de aprendizaje competitivo. Todas estas redes se compararon con el modelo logístico desarrollado por MONSERUD & STERBA (1999). Información sobre estos algoritmos puede encontrarse en RIEDMILLER & BRAUN (1993) para el Rprop, MOELLER (1993) para SCG y KOHONEN (1988b) para LVQ. Los modelos de correlación en cascada CC son explicados en FAHLMAN & LEBIERE (1990) La diferencia encontrada respecto el número predicho y observado de pies muertos fue más pequeño en todos los modelos neuronales respecto a la logística, siendo el mejor de ellos el LVQ.

Por último, otro ejemplo de red neuronal con aplicación forestal es el propuesto por HASENAUER & MERKL (2000), para predecir el establecimiento de regenerado en masas mixtas. La red propuesta está formada por 6 unidades de entrada, 7 unidades ocultas y una de salida con algoritmo Rprop hallando un coeficiente de determinación de 0.67.

La mortalidad natural es el proceso más ampliamente estudiado mediante redes neuronales en el ámbito forestal debido a la sencillez de su salida (*output*). Un árbol está muerto o vivo al final del periodo considerado. Sin embargo, pueden existir otro tipo de aproximaciones con redes neuronales como la predicción de clases diamétricas (LEDUC et al., 2001), que abre el campo de la modelización de eventos forestales, mediante herramientas basadas en inteligencia artificial, a aspectos aún poco explorados, como la clasifica-

ción en calidades de estación, separación de especies en función de parámetros ecológicos, clasificaciones tecnológicas de la madera o la probabilidad de ocurrencia de ataque de insectos. Allí donde son más débiles los modelos matemáticos no lineales puede existir una red neuronal que funcione. Sólo hay que encontrarla.

Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia y Tecnología con cargo al proyecto AGL2001-1780 en el marco del Plan Nacional de Investigación y Desarrollo Tecnológico 2001-2003.

BIBLIOGRAFIA

- BUCHMAN, R.G.; PEDERSON, S.P. & WALTERS, N.R.; 1983. A tree survival model with application to species of the Great Lakes region. *Can. J. For. Res.* 13: 601-608.
- FAHLMAN, S.E. & LEBIERE, C.; 1990. *The cascade-correlation learning architecture*. Technical Report CMU-CS-90-100. Dept. of Computer Science. Carnegie Mellon University. Pittsburgh.
- GUAN, B.T. & GERTNER, G.; 1991a. Using a parallel distributed processing system to model individual tree mortality. *For. Sci.* 37(3): 871.
- GUAN, B.T. & GERTNER, G.Z.; 1991b. Modeling red pine tree survival with an artificial neural network. *For. Sci.* 37(5): 1429-1440.
- GUAN, B.T. & GERTNER, G.Z.; 1995. Modeling individual tree survival probability with a random optimization procedure: An artificial neural approach. *AI Applications* 9(2): 39.
- HAIR, J.F.; ANDERSON, R.E.; TATHAM, R.L. & BLACK, W.C.; 1999. *Análisis multivariante*. Prentice Hall Iberia. Madrid.
- HANGOS, K & CAMERON, I.; 2001. *Process Modelling and Model Analysis. Vol. 4 Process system engineering*. Academic Press. Londres.
- HASENAUER, H. & MERKL, D.; 2000. Predicting regeneration establishment within mixed species stands using neural networks. In: B. Koch (eds.), *Spatial technologies for forest inventories. Contributions to the XXI IUFRO World Congress*: 145-154. Freiburger Forstliche Forschung. Kuala Lumpur.
- HASENAUER, H.; MERKL, D. & WEINGARTNER, M.; 2001. Estimating tree mortality of Norway spruce stands with neural networks. *Adv. Env. Res.* 5(4): 405-414.
- HEBB, D.O.; 1949. *The organization of Behavior*. Wiley. New York
- HILERA, J.R. & MARTÍNEZ, V.J.; 1995. *Redes Neuronales Artificiales. Fundamentos, modelos y aplicaciones*. RA-MA. Madrid
- KING, S.L., KRISTIN, P.B. & LIST, S.; 2000. Modeling noncatastrophic individual tree mortality using logistic regression, neural networks, and support vector methods. *Computers and electronics in agriculture* 27: 401.
- KOHONEN, T.; 1988a. An introduction to neural computing. *Neural Networks* 1: 3-16.
- KOHONEN, T.; 1988b. Learning Vector Quantization. *Abstracts of the First Annual INNS Meeting*. 303.
- LEDUC, D.J.; MATTNEY, D.G., BELL, K.L., BALDWIN JR. V.C. 2001. *Predicting diameter distributions of longleaf pine plantations: a comparison between artificial neural networks and other accepted methodologies*. Res. Pap. SRS-25-USDA Forest Service.
- LEK, S.; DELACOSTE, M.; BARAN, P.; DIMOPOULOS, I.; LAUGA, J. & AULAGNIER, S.; 1996. Application of neural networks to modelling nonlinear relationships in ecology. *Ecol. Model.* 90: 39-52.
- LEK, S. & GUÉGAN, J.F.; 1999. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecol. Model.* 120: 65-73.
- MANEL, S.; DIAS, J.M. & ORMEROD, S.J.; 1999. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecol. Model.* 120: 337-347.
- MOELLER, M.F.; 1993. A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks* 6: 525-533.
- MONSERUD, R.A. & STERBA, H.; 1999. Modeling individual tree mortality for Austrian forest species. *For. Ecol. Manage.* 113: 109-123.
- PEÑA, D.; 2002. *Análisis de datos multivariantes*. McGraw-Hill. Barcelona

- RECKNAGEL, F.; 2001. Applications of machine learning to ecological modelling. *Ecol. Model.* 146: 303-310.
- RIEDMILLER, M.; BRAUN, H.; 1993. A direct adaptive method for faster backpropagation learning: The Rprop algorithm. *En: H. Ruspini (ed.). Proceedings of the IEEE International Conference on Neural Networks: 586-591.* ICNN.
- RUMELHART, D.E.; HINTON, G.E. & EILLIAMS, R.J.; 1986. Learning representations by backpropagation errors. *Nature* 323: 533-536.
- SARLE, W.S. (Editor). *Neural Network FAQ, En: Usenet newsgroup comp.ai.neural-nets, URL: ftp://ftp.sas.com/pub/neural/FAQ.html*
- VERIKAS, A. & BACAUSKIENE, M.; 2003. Using artificial neural networks for proces and system modelling. *Chemom. Intel. Lab. Sys.* 67: 187-191.
- WEINGARTNER, M.; 1999. *Using neural networks to predict individual tree mortality.* Diplomarbeit. Institut für Softwaretechnik E188 der Technischen Universität Wien.Wien.
- ZELL, A. (Coordinador). 1999. *SNNS. Stuttgart Neural Network Simulator. User Manual, V. 4.2.* Universidad de Stuttgart-Universidad de Tübingen. Stutgart.