

## Publicaciones en Ciencias y Tecnología

Vol 6, año 2012, N<sup>o</sup>1, pp.31–40, ISSN:1856-8890, Depósito Legal: pp200702LA2730PREDICCIÓN DEL RENDIMIENTO DE UN CULTIVO DE  
PLÁTANO MEDIANTE REDES NEURONALES  
ARTIFICIALES DE REGRESIÓN GENERALIZADA

\* Rita M. Ávila G. de Hernández    \*\* Vianel Rodríguez Pérez  
\*\*\* Edwin A. Hernández Caraballo

Recibido: 24/04/2012, Aceptado: 21/06/2012

## Resumen

El plátano es un cultivo importante en Latinoamérica, tanto a gran escala como para pequeños hacendados. En Venezuela, los principales centros productivos están en la zona del Sur del Lago de Maracaibo. Conocer el rendimiento de un cultivo obedece a la necesidad de maximizar la relación inversión-ganancia y la disponibilidad de esa información, con anticipación, permite tomar decisiones sobre el manejo de una plantación. La finalidad de esta investigación es evaluar la capacidad de las redes neuronales artificiales para predecir el rendimiento de un cultivo de plátano, empleando para ello el mejor grupo de datos predictores, determinado entre las características físicas del suelo y el perfil químico del tejido foliar. Se emplean redes de regresión generalizada y la estrategia *leave-one-out*, así como dos tipos de transformaciones de los datos. Se encuentra que las redes neuronales artificiales constituyen una excelente herramienta de predicción del rendimiento del cultivo de plátano y que los perfiles físico-químicos de suelo y del tejido foliar son adecuados descriptores para la variable respuesta. Entre los datos evaluados, en este estudio, se encontró que los datos físicos de suelo de 20-40 cm son el mejor grupo predictor, previo la estandarización de los datos de entrenamiento.

**Palabras clave:** red neuronal de regresión generalizada, predicción, rendimiento de cultivos, caracterización físico-química, *Musa* AAB.

PREDICTION OF THE YIELD OF BANANA CROP BY MEANS OF A GENERALIZED  
REGRESSION ARTIFICIAL NEURAL NETWORK

## Abstract

Banana is an important crop in Latin America, for both large scale and small farmers. In Venezuela, the main production centers are in the area of Sur del Lago de Maracaibo. Knowing the crop yield is vital due to the need to maximize the investment-profit, and availability of such information in advance helps in the decision-making process of the production unit. The purpose of this research is to evaluate the ability of artificial neural networks to predict the yield of a banana crop, employing the best predictors dataset, given the physical characteristics of the soil and the chemical profile of leaf tissue. Generalized regression networks trained with the leave-one-out strategy, and two types of data transformations were used in this study. It was found that neural networks were excellent tools for predicting the yield of banana crop. The physicochemical profiles of soil and leaf tissue were suitable descriptors of the response variable. Among the data evaluated in this study, it was found that the physical data of 20-40 cm of soil was, after standardization of the training data, the best predictor group.

**Keywords:** generalized regression neural network, prediction, crop yield, physical-chemical characterization, *Musa* AAB.

\* *Ingeniero Químico. Docente del Programa de Ingeniería Agroindustrial. Decanato de Agronomía, Universidad Centroccidental "Lisandro Alvarado", ritaavila@ucla.edu.ve*

\*\* *Doctor en Ingeniería Agronómica. Departamento de Fitotécnia. Decanato de Agronomía, Universidad Centroccidental Lisandro Alvarado, vianelr@ucla.edu.ve.*

\*\*\* *Doctor en Química. Departamento de Química y Suelos. Decanato de Agronomía, Universidad Centroccidental Lisandro Alvarado, ehernandez@ucla.edu.ve*

## Introducción

El plátano tiene una importancia económica y comercial en toda Latinoamérica (Martínez et al., 2009), tanto a gran escala como para pequeños hacendados. En Venezuela, los principales centros productivos están en la zona del Sur del Lago (estados Zulia, Trujillo, Mérida y Táchira), región que representa el 70 % de la producción nacional de este rubro (Machado-Allison y Rivas, 2004). Conocer el rendimiento de un cultivo en general, y del plátano en particular, obedece principalmente a la necesidad de maximizar la relación inversión-ganancia y la disponibilidad de esa información con anticipación permite tomar decisiones sobre el manejo de una plantación (NC State University, S/F).

Para la predicción del rendimiento la experiencia del agricultor es importante y cuando no se cuenta con métodos para la estimación, ésta se convierte en el único recurso; sin embargo, estas aproximaciones pueden ser insuficientes, y lo que se necesita es información sistemáticamente almacenada que contemple, por ejemplo: (i) registros históricos promedio de producción; y, (ii) variaciones de rendimiento por manejo agrícola o por factores climáticos entre otros, para disminuir cualquier sesgo o error (NC State University, Op. Cit.). Esta situación hace que se desarrollen nuevas metodologías y que se consideren otros datos *-por ejemplo físicos, climáticos, administrativos, perfiles químicos-* que puedan mejorar la calidad de las predicciones (Kumi-Boateng y Yakubu, 2010).

Datos	Herramienta	Cultivo	Referencia
Indices de Balance Nutricional DRIS (IBN-DRIS).	Ecuaciones de regresión.	Plátano Musa AAB Hartón	Rodríguez et al., 1999a.
Balance de nutrientes IBN-DRIS y las variables biométricas (número de hojas de la planta madre, largo, ancho y relación foliar de la hoja III).	Ecuaciones de regresión.	Plátano Musa AAB Hartón	Rodríguez et al., 1999b.
Veintidós variables asociadas al rendimiento (por ejemplo peso de la primera mano, diámetro del dedo interno de la segunda mano, peso de la segunda mano, número de manos, perímetro y diámetro del dedo externo de la segunda mano, longitud interna del dedo interno de la segunda mano).	Modelos de regresión: Forward y Stepwise.	Plátano Musa AAB Hartón	Chacín et al., 2005
Balance nutricional IBN-DRIS y número de hojas de la planta madre.	Ecuaciones de regresión.	Plátano Musa AAB Hartón	Rodríguez et al., 2005.
Datos de la planta (Ancho de la hoja, circunferencia de la planta, número de hojas, altura de la planta) y del fruto (dedos en racimo, número de manos, diámetro y longitud de dedos).	Modelos de regresión múltiple.	Plátano	Kumar et al., 2007.

Cuadro 1: Selección de algunos métodos que se emplean para predecir el rendimiento en el cultivo del plátano (Musa AAB, subgrupo ‘Hartón’).

Como se aprecia en el Cuadro 1, la predicción del rendimiento del cultivo de plátano ha sido abordada por diversos investigadores, quienes han empleado varias herramientas matemáticas convencionales que recurren a datos físicos o químicos relacionados con el suelo, las hojas y los frutos. Sin embargo, los ajustes

de los modelos desarrollados en esos trabajos (reportados en función del coeficiente de determinación) muestran que dichas herramientas no son suficientemente poderosas para modelar la complejidad del sistema en cuestión. En consecuencia, se requiere de alternativas matemáticas más apropiadas para ese fin.

Las redes neuronales artificiales se encuentran entre las opciones con las que cuentan los investigadores para el desarrollo de modelos de sistemas complejos, como es el caso de la predicción del rendimiento de un cultivo (Hernández-Caraballo, 2011) y puede emplearse junto con otras herramientas para el manejo científico de las unidades de producción. Los modelos basados en redes neuronales, que emulan el razonamiento humano, aventajan a los matemáticos en los casos en los que se desconoce la relación entre las variables dependientes e independientes, o éstas son tan complejas que se dificulta su formulación (Hernández et al., 2003). En efecto, entre las razones que las hace una alternativa de solución a estos sistemas, está que no es necesario conocer las relaciones funcionales entre variables dependientes e independientes. Existen investigaciones para predecir el rendimiento de cultivos con redes neuronales, por ejemplo en arroz (Ji et al., 2007), trigo (Khashei-Siuki, et al., 2011) y cebolla (Stastny et al., 2011); sin embargo, en lo que respecta al plátano ésta es escasamente reportada. De acuerdo con los argumentos anteriores se considera que las redes neuronales sean suficientemente poderosas para modelar la compleja relación que existe entre las variables físicoquímicas del suelo y el tejido foliar, y el rendimiento del cultivo de plátano.

La finalidad de esta investigación es evaluar la capacidad de las redes neuronales artificiales de regresión generalizada para predecir el rendimiento de un cultivo de plátano, empleando para ello el mejor grupo de datos predictores, determinado entre las características físicas del suelo y el perfil químico del tejido foliar.

## Medodología

1. *Base de datos: características y procesado.* La base de datos que se emplea en este estudio se construye a partir del muestreo de suelo y de tejido foliar, que realiza Rodríguez (2003) en una unidad de producción de plátano (*Musa* AAB Subgrupo plátano cv. Hárton) del Sur del Lago de Maracaibo, Venezuela. El lector interesado puede consultar el trabajo citado para los detalles del muestreo. Para los fines de este estudio, baste con expresar que la base de datos está constituida por 131 pares de datos, a saber, las variables predictoras (entradas de las redes neuronales), y el correspondiente rendimiento de la planta (kg de plátano/racimo). Las variables predictoras empleadas son las siguientes:

- Componente suelo (13 variables). Análisis físico: *pH*, *conductividad eléctrica (CE)*, *materia orgánica (MO)*, *arena (a)*, *limo (L)*, *arcilla (AA)* y análisis químico: *P*, *K*, *Mg*, *Mn*, *Cu*, *Fe*, *Zn*. Estas variables se repiten para dos profundidades u horizontes (0-20 y 20-40 cm).
- Componente foliar (11 variables). Análisis químico: *S*, *N*, *P*, *Na*, *K*, *Ca*, *Mg*, *Zn*, *Cu*, *Fe*, *Mn*.

Los valores de las variables predictoras se encuentran en distintos órdenes de magnitud y ello pudiera afectar el desempeño de los modelos desarrollados por las redes neuronales. Para evitar este problema, los datos se sometieron a dos transformaciones (Hernández-Caraballo, 2011), a saber, **Normalización** (1) y **Estandarización** (2), las cuales se describen a continuación:

$$D_{Ni} = \frac{D_i}{D_{maximo}} \quad (1)$$

Donde  $D_{Ni}$  es el dato normalizado;  $D_i$  es el dato original y  $D_{maximo}$  el valor máximo de la serie. Por su parte para estandarizar, se debe primero restar a cada valor la media de su serie y luego dividirlo entre la desviación estándar de la misma serie de datos. La expresión que se utiliza es la siguiente:

$$D_{Ei} = \frac{D_i - x}{s} \quad (2)$$

En este caso,  $D_{Ei}$  es el dato estandarizado;  $D_i$  es el dato original;  $x$  la media y  $s$  la desviación estándar de la serie.

2. *Redes neuronales, descripción, estrategia de entrenamiento y selección.* La red neuronal de regresión generalizada, es un tipo de red que se usa para desarrollar regresiones entre variables dependientes e independientes y su estructura se muestra en la figura 1.

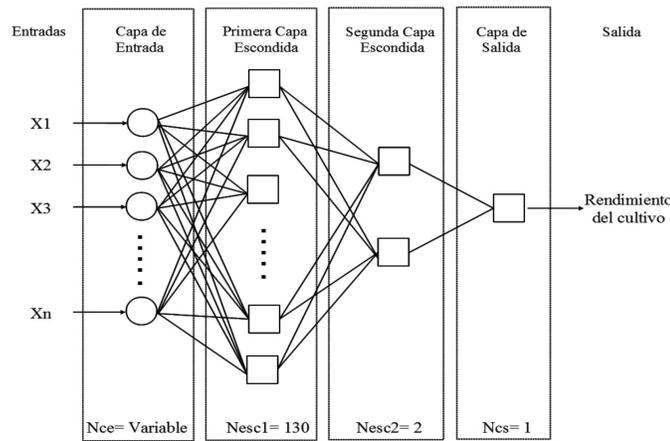


Figura 1: Arquitectura de una red neuronal de regresión generalizada. Nce: nodos en la capa de entrada. Nesc1: nodos en la capa escondida 1. Nesc2: nodos en la capa escondida 2. Ncs: nodos en la capa de salida

En la figura 1, se observa que la capa de entrada tiene los nodos que distribuyen los datos de las variables descriptivas (datos físico-químicos del suelo y del tejido foliar), hacia todas las neuronas presentes en la primera de las dos capas subsiguientes. A continuación se encuentran las capas escondidas; la primera tiene un número fijo de neuronas, el que es igual al de casos empleados en el desarrollo de un modelo determinado y las neuronas de esta capa procesan los datos recibidos mediante la función de base radial que se describe en la ecuación 3:

$$FBR = \exp \frac{\|c_i - p_i\|^2}{FS} \quad (3)$$

Donde  $c_i$  es el centroide (valor fijo determinado por los vectores que se emplean en la construcción de un modelo);  $FS$  es la amplitud de la función de base radial, o Factor de Suavizado, que se optimiza durante el entrenamiento de la red a fin de maximizar el desempeño del modelo); y  $p_i$  es el valor de un dato de prueba dado. La segunda capa escondida tiene dos neuronas y el proceso en ésta ocurre de la siguiente manera: una neurona calcula la suma de los centroides de las neuronas de la primera capa escondida, ponderándola a través de la cercanía entre  $c_i$  y  $p_i$  (numerador de la ecuación 3); y la otra, estima la suma de los factores de ponderación, i.e., la suma de los pesos asignados a las neuronas en función del grado de similitud mencionado anteriormente. En el extremo opuesto del diagrama se encuentra la capa de salida, la que consta de una única neurona responsable de arrojar el valor de la variable predicha (rendimiento del cultivo de plátano). Este valor es el cociente de los valores generados por las dos neuronas de la capa anterior (Bauer, 1995).

Se entrenan cuatro topologías de redes, una para cada grupo de datos, a saber, dos para suelo (una para cada horizonte), una para hoja y una combinada suelo-hoja. En el Cuadro 2, se muestra la

distribución de las variables de entrada por cada grupo. Los datos del horizonte de 0-40 cm -pH, a, L, K- se duplican para la profundidad de 0-20 cm para completar los datos.

El entrenamiento se realiza con el software Statistica bajo ambiente Windows y los datos se presentan a la red bajo la modalidad “leave one out” -deja uno afuera- que consiste en tomar n-1 datos para el entrenamiento en cada ciclo, lo que hace variarlos y causar una perturbación que permite estudiar cómo varía la respuesta del algoritmo de aprendizaje bajo los efectos de ésta (Elisseff y Pontil, S/F). En este caso son 131 datos y en consecuencia, se emplean cada vez 130 para entrenar.

<i>Grupo</i>		<i>Variables de entrada</i>	<i>Salida</i>
1 - A	Suelo (0 - 20 cm)	CE, MO, AA, P, Mg, Mn, Cu, Fe, Zn, K, pH, a, L	Rendimiento
1 - B	Suelo (20 - 40 cm)	CE, MO, AA, P, Mg, Mn, Cu, Fe, Zn, K, pH, a, L	
2	Hoja	N, P, K, S, Ca, Mg, Zn, Cu, Fe, Mn	
3	Combinación suelo-hoja	(CE, MO, AA, P, Mg, Mn, Cu, Fe, Zn) suelo a 0-20 cm, (CE, MO, AA, P, Mg, Mn, Cu, Fe, Zn, K, pH, a, L) suelo a 20-40 cm y (N, P, K, S, Ca, Mg, Zn, Cu, Fe, Mn) hoja.	

Cuadro 2: Clasificación de los datos para el entrenamiento de la red neuronal. CE: conductividad eléctrica. MO: materia orgánica. AA: arcilla. a: arena. L: limo.

El desempeño de las redes se monitoriza mediante la raíz del error cuadrático medio (RECM) que resume el error total de la predicción y se calcula con la expresión (4). Consiste en sumar el cuadrado de la diferencia del rendimiento real (Rreal) y el bruto (Rbruto), extraerle la raíz cuadrada y dividirlo entre el número de datos (n).

$$RECM = \frac{\sqrt{\sum(Rreal - Rbruto)^2}}{n} \quad (4)$$

Posteriormente se seleccionan las redes con menor RECM, se calcula el rendimiento que predice cada una de estas redes y luego éste se grafica versus el rendimiento real. Para finalizar, a cada gráfica se le determina la ecuación de correlación, el coeficiente de ajuste (R) y se calcula el error relativo de la raíz cuadrática media (ERRCM) que es el parámetro que se utiliza para hacer la selección final de la red que mejor predice el rendimiento. El ERRCM, resume el error total de la predicción y se calcula de acuerdo a la expresión 5. Se basa en sumar los errores individuales estandarizados, dividirlo entre los grados de libertad y extraerle la raíz cuadrada. El ERRCM varía en el intervalo [0, 1]. Mientras más cercano el valor a cero, mejor es la predicción de la red y al aproximarse a uno, ésta tiende a predecir el valor medio.

$$ERRCM = \frac{\sqrt{(\frac{1}{N})\sum[y_i - Y_i]^2}}{\sqrt{(\frac{1}{N})\sum[y_i - y]^2}} \quad (5)$$

Donde, N es el número de datos;  $y_i$ ,  $Y_i$ , e y son el rendimiento predicho por la red, el real del cultivo y el rendimiento promedio, respectivamente.

## Resultados y discusión

### 1. Análisis estadístico de los datos.

En los cuadros 3 y 4, se resume la estadística básica de los datos de suelo, hoja y rendimiento. Se

observa que los datos de hoja ( $11,56\% < CV < 80,67\%$ ) son menos dispersos que los de suelo ( $16,44\% < CV < 133,25\%$ ). Por su parte, el suelo tiene una textura muy variada desde limosa hasta franco arenosa, siendo en promedio franco limosa según el triángulo de las clases texturales Escala USDA (Marín et al., 2003), que se basa en el contenido de arcilla, limo y arena. En promedio es ligeramente ácido ( $6 < pH < 7$ ) y debido al contenido de materia orgánica (Walkley y Black, citados por Marín et al., Op. cit) es un suelo mineralizado ( $1\% < MO < 1,9\%$ ) y rico en hierro, fósforo y magnesio; mientras que las hojas de plátano son ricas en potasio, manganeso e hierro. Sin embargo, cotejando uno a uno los valores promedio de los nutrientes en hojas -del Cuadro 2- con los rangos adecuados de concentraciones foliares de nutrientes -se incluyen entre paréntesis- para el cultivo del banano (International Plant Nutrition Institute, 2001), se tiene que el N, está fuera del rango adecuado (2,8-4); el P está en el límite inferior del rango adecuado (0,2-0,25); el K supera el límite del rango adecuado (3-4); Ca, Mg, Zn y Cu están en concentraciones adecuadas, en ese mismo orden (0,5-1,2); (0,2-0,46); (21-35) y (7-20); y por último Fe y Mn están fuera del rango adecuado (70-200) y (100-2200), respectivamente.

2. *Estrategia de entrenamiento y selección de las redes con el mejor desempeño.*

Los resultados del desarrollo de los modelos basados en las redes neuronales se muestran en la figura 2. En general, el desempeño de las redes muestra una tendencia similar para los dos tipos de procesado al que se someten los datos, cuando el logaritmo del factor de suavizado es igual o menor que 0,1. Por encima de este valor, el desempeño de las redes que emplean datos normalizados se deteriora más rápidamente que el de aquellas que utilizan datos estandarizados.

	Concentración											Rend. (Kg)
	Macronutrientes primarios (g/kg)				Macronutrientes secundarios (g/kg)			Micronutrientes (mg/kg)				
	N	P	K	Na	S	Ca	Mg	Zn	Cu	Fe	Mn	
Mín.	1,11	0,12	2,80	0,01	0,11	0,22	0,17	8,43	4,75	21,56	17,52	8,60
Máx.	3,63	0,30	5,46	0,27	0,45	1,29	0,59	144,99	13,56	126,31	204,54	29,00
Media	2,61	0,20	4,24	0,06	0,23	0,70	0,30	20,95	8,52	54,20	74,65	15,86
DE	0,35	0,03	0,49	0,04	0,07	0,22	0,06	16,90	1,62	21,32	30,43	3,63
CV(%)	13,41	15,00	11,56	66,67	30,43	31,43	20,00	80,67	19,01	39,34	40,76	22,88

Cuadro 3: Estadística básica para los datos del perfil químico foliar y el Rendimiento (n=131). Min. y Máx: valores mínimo y máximo de la serie. DE: desviación estándar. CV: coeficiente de variación. Rend.: rendimiento.

	Min	Max	Media	DE	CV (%)
Profundidad 0 - 20 cm					
Conductividad Eléctrica, CE (mS/cm)	0,07	1,60	0,39	0,24	61,90
Materia Orgánica, MO (%)	0,10	3,35	1,62	0,64	39,52
Arcilla, AA (%)	9,00	37,00	20,68	6,16	29,81
P (g/kg)	1,00	133,00	26,77	26,52	99,06
Mg (g/kg)	5,83	54,75	18,74	9,69	51,71
Mn (mg/kg)	1,49	52,78	11,38	9,19	80,78
Cu (mg/kg)	0,40	3,93	1,75	0,66	37,47
Fe (mg/kg)	7,29	406,54	57,05	65,85	115,42
Zn (mg/kg)	0,39	9,35	1,80	1,32	73,11
Profundidad 20-40 cm					
Conductividad Eléctrica, CE (mS/cm)	0,06	1,40	0,32	0,22	66,93
Materia Orgánica, MO (%)	0,30	2,30	1,04	0,44	42,01
Arcilla, AA (%)	7,00	41,00	19,23	6,18	32,11
P (g/kg)	1,00	133,00	17,74	20,15	113,57
Mg (g/kg)	5,17	46,42	14,85	6,84	46,09
Mn (mg/kg)	1,49	45,65	8,78	7,75	88,18
Cu (mg/kg)	0,51	5,08	1,40	0,57	41,03
Fe (mg/kg)	4,81	361,63	39,96	53,25	133,25
Zn (mg/kg)	0,24	14,33	1,27	1,56	123,49
Profundidad 0-40 cm					
pH	3,60	8,45	6,64	1,09	16,44
Arena, a (%)	1,00	61,00	20,94	15,58	74,38
Limo, L (%)	23,00	77,00	59,16	11,89	20,09
K (g/kg)	0,52	9,27	1,41	1,18	83,58

Cuadro 4: Estadística básica para los datos físicos y el perfil químico de suelo (n= 131). Min. y Máx: valores mínimo y máximo de la serie. DE: desviación estándar. CV: coeficiente de variación. Rend.: rendimiento.

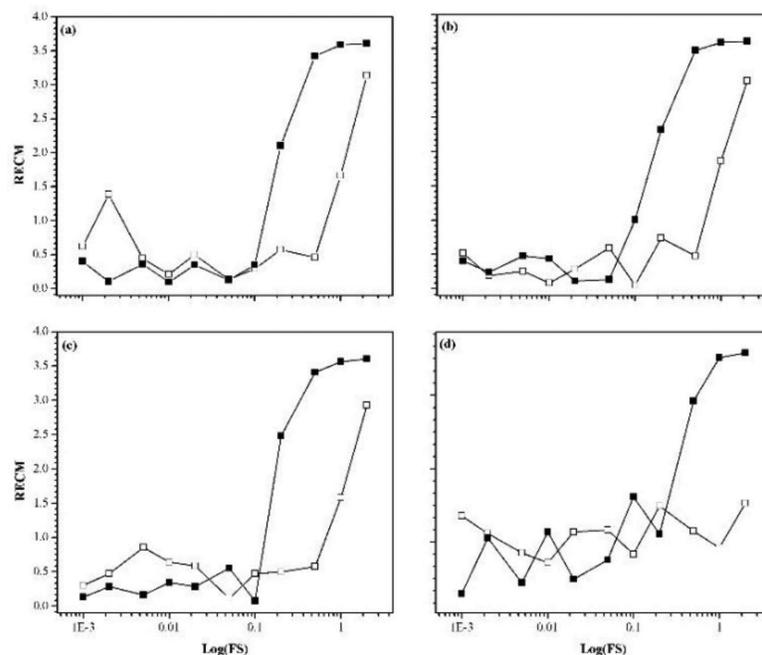


Figura 2: Variación de la raíz del error cuadrático medio (RECM) versus el logaritmo del Factor de Suavizado (FS). Redes: a) suelo de 0-20. b) suelo de 20-40. c) hoja. d) suelo-hoja. ( Datos estandarizados. Datos normalizados).

Datos de entrada de la red	FS	Ecuación de correlación	$R^2$	RECM	ERRCM
Suelo 0-20 (a)	0,01	$y=0,9425x + 0,9359$	1,0000	0,20929	0,05789
Suelo 0-20 (b)	0,01	$y=0,9749x + 0,3693$	1,0000	0,09426	0,02607
Suelo 20-40 (a)	0,1	$y=1,0018x + 0,0219$	1,0000	0,05037	0,01393
Suelo 20-40 (b)	0,02	$y=0,9859x + 0,3241$	1,0000	0,11355	0,03139
Hoja (a)	0,05	$y=1,0101x - 0,0563$	1,0000	0,10963	0,03031
Hoja (b)	0,1	$y=1,0161x - 0,211$	1,0000	0,07274	0,02012
Combinación suelo-hoja (a)	0,01	$y=1,1797x - 3,1501$	0,9999	0,71538	0,19721
Combinación suelo-hoja (b)	0,001	$Y=0,9778x - 0,6257$	0,9994	0,29722	0,08198

Cuadro 5: Variación del ERRCM en función del FS para varios grupos de datos. (a: entrada y salida estandarizada. b: entrada y salida normalizada. En negrita, las redes con el menor error global.). FS: factor de suavizado. R: coeficiente de ajuste. RECM: raíz del error cuadrático medio. ERRCM: error relativo de la raíz cuadrática media.

A partir de los resultados que se muestran en la figura 2, se seleccionaron los valores de FS asociados al mejor desempeño de cada red (menor ERRCM) y se evaluó la capacidad de cada una de estas redes para predecir el rendimiento real de este cultivo, mediante análisis de correlación. En el Cuadro 5 se observa que las que mejor predicen el rendimiento son las entrenadas con: datos de suelo 0-20 con la entrada y la salida normalizada, la de suelo 20-40 con la entrada y la salida estandarizada, la de hoja y la de combinación suelo-hoja, ambas con entrada y salida normalizada. En orden creciente de ERRCM, las redes se presentan de la siguiente manera: suelo 20-40 /hoja/suelo 0-20/combinación suelo-hoja. Considerando que a menor ERRCM, mejor es la predicción, significa que, para esta muestra de plátano, el grupo que mejor predice el rendimiento es el de suelo 20-40 cm. Este resultado es de mucha relevancia porque es del suelo que provienen los nutrientes que al final favorecen el desarrollo de los frutos, lo que determina el rendimiento. Sin embargo, al cotejar las concentraciones de nutrientes en los dos horizontes de suelo pareciera que la profundidad de 0-20 cm es más rica que la de 20-40 cm, lo que resulta interesante porque aunque el sistema radical del plátano puede alcanzar más de dos metros de longitud, el mayor número de raíces alimentadoras se distribuye en los primeros 20 cm de profundidad orientándose hacia las regiones del suelo más fértiles (Rodríguez, 2009). Esta situación sugiere una nueva interrogante en relación a cómo están correlacionados los datos de suelo a 20-40 cm con el rendimiento -kg racimo/planta- en comparación con los de la profundidad de 0-20 cm.

## Conclusiones

Las redes neuronales artificiales constituyen una excelente herramienta de predicción del rendimiento del cultivo de plátano, tal como lo revelan los distintos indicadores de desempeño ( $R^2$ , RECM, ERRCM). Los perfiles físicoquímicos de suelo y del tejido foliar son adecuados descriptores para la variable respuesta. Entre los datos evaluados, en este estudio, se encontró que los datos físicos de suelo de 20-40 cm son el mejor grupo predictor. El tipo de red empleado ofrece el atractivo de la rapidez con la que se generan los modelos porque solo requiere la optimización del Factor de Suavizado (además del tratamiento adecuado de los datos). Sin embargo, se debe dedicar particular atención a este factor, puesto que éste determina cuán efectivo es un proceso de transformación de los datos de entrenamiento. En este caso, el desempeño de la red presenta un comportamiento similar por debajo del logaritmo de 0,1 mientras que por encima, es muy diferente dependiendo del tipo de transformación.

## Agradecimientos

Al Fondo Lisandro Alvarado para la Investigación de la Universidad Centroccidental Lisandro Alvarado, cuyos recursos se emplearon para la realización de este trabajo y a los árbitros anónimos que revisaron este trabajo, quienes con sus valiosos aportes contribuyeron a la versión final de este manuscrito.

## Referencias

- [1] Bauer, M. M. 1995. Generalized regression neural network for technical use. Master's Thesis, University of Wisconsin-Madison, USA.
- [2] Chacín, F., Ascanio, M., Hernández, A., García, J., Cobo, M., Ascanio, A. (2005). Predicción de cosechas en plátano (*Musa AAB 'Hartón'*) mediante dos modelos de regresión (Forward y Stepwise). *Rev. Fac. Agron. (Maracay)*, Vol. 31. No. 1. Pp 21-35.
- [3] Elisseeff, A., Pontil, M. (S/A). Leave one out error and stability of learning algorithms with applications. Disponible en: <http://www.citeseerx.its.psu.edu>
- [4] Hernández-Caraballo, E. A. (2011). Predicción del rendimiento potencial de un cultivo de naranja 'Valencia' mediante redes neuronales de regresión generalizada. Trabajo de Ascenso (sin publicar). Universidad Centroccidental Lisandro Alvarado. Venezuela.
- [5] Hernández, E., Avila, R., Rivas, F. (2003). Las redes neuronales artificiales en química analítica. Parte I. Fundamentos. *Rev. Sociedad Venezolana de Química*. Vol. 26. No. 4. Pp 17-25.
- [6] International Plant Nutrition Institute (2001). Guía de Campo. Síntomas de deficiencias nutricionales y otros desórdenes fisiológicos en el cultivo del banano (*Musa AAA*). Descripción, causas, prevención y corrección. IPNI. Disponible en: [http://nla.ipni.net/articles/NLA0070-EN/\\$FILE/G%20Banano.pdf](http://nla.ipni.net/articles/NLA0070-EN/$FILE/G%20Banano.pdf).
- [7] Ji, B., Sun, Y., Yang, S., Wan, J. (2007). Artificial neural networks for rice yield prediction in mountainous regions. *J. of Agricultural Sci.* Vol. 145. Pp 249-261.
- [8] Khashei-Siuki, A., Kouchakzadeh, M., Ghahraman, B. (2011). Predicting dryland wheat yield from meteorological data, using expert system, Khorasan Province, Iran. *J. Agr. Sci. Tech.*, Vol. 13. Pp 627-640.
- [9] Kumar, R., Rajan, S., Singh, A. (2007). Multiple regression models for yield prediction in plantain. *Indian J. of Horticulture*, Vol. 64. No. 2. Pp 159-162.
- [10] Machado-Allison, C., Rivas, J.C., 2004. La agricultura en Venezuela. Ediciones IESA, Venezuela.
- [11] Marín, M., Aragón, P., Gómez, C. (2003). Análisis químico de suelo y aguas. Manual de laboratorio. Editorial UPV: Valencia. Pp 175.
- [12] Martínez, E., Delgado, E., Rey, J., Giménez, C., Pargas, R., Manzanilla, R. (2009). Producción del plátano en Venezuela y el mercado mundial. *INIA Hoy*, Vol. 5 mayo-agosto. Pp 125-138.
- [13] NC State University (S/F). North Carolina Winegrape Grower's Guide. Chapter 12. Crop Prediction. Disponible en: <http://cals.ncsu.edu/hortsci/extension/documents/winegrape/winegrapesada12.pdf>
- [14] Rodríguez, G. (2009). Aspectos sobre la salud radical del banano en suelos de Venezuela. *Producción Agropecuaria*. Vol. 2. No.1. Pp 49-52.

- [15] Rodríguez, V. (2003). Avaliação do Estado Nutricional e da Fertilidade do Solo na Cultura do Platano (Musa AAB Subgrupo Platano cv. Hartón). Tesis Doctoral. USP Piracicaba, Brasil.
- [16] Rodríguez, V., Rodríguez, O., Bravo, P. (1999a). Índice de balance de nutrimentos para la predicción del rendimiento del plátano (Musa AAB subgrupo plátano Hartón). Rev. Fac. Agron. LUZ, Vol. 16. Pp 488-494.
- [17] Rodríguez, V., Bautista, D., Rodríguez, O., Díaz, L. (1999b). Relación entre el balance nutricional y la biometría del plátano (Musa AAB subgrupo plátano Hartón) y su efecto sobre el rendimiento. Rev. Fac. Agron. LUZ, Vol. 16. Pp 425-432.
- [18] Rodríguez, V., Da Silva, A., Rodríguez, O. (2005). Balance nutricional y número de hojas como variables de predicción del rendimiento del plátano Hartón. Pesq. Agrop. Bras. Vol. 40. Pp 175-177.
- [19] Stastny, J., Konecny, V., Trenz, O. (2011). Agricultural data prediction by means of neural networks. Agric. Econ. Czech. Vol. 57. No. 7. Pp 356-361.