

## **Simulación de la producción de café (Coffea) en Brasil con redes neuronales (RNA)**

BEDOYA-CARDOSO, Marlio, SALAZAR, Raquel, PÉREZ, Francisco y PORTILLO, Marcos

M. Bedoya, R. Salazar, F. Pérez y M. Portillo

^ Universidad Autónoma Chapingo, México.  
Marlio.bedoya@gmail.com

D. Sepúlveda, F. Rérez, D. Sepúlveda, E. Figueroa, R. Salazar, L. Godínez (eds.) Matemáticas Aplicadas. Handbook T-I. -©ECORFAN, Texcoco de Mora-México, 2016.

## Abstract

Coffee (*Coffea*) has an important role in the Brazilian economy, it is the primary sector activity that most employs rural labor, it has also promoted the development of enterprises in industry and services, likewise, it is an important source in the entry of foreign exchange (Sugai *et al.*, 2004). An important element in the management of coffee farms in the harvest forecast to planning activities and through which the number of contracted employees can be estimated, this paper the results of the simulation of coffee production in Brazil are presented using networks neural, in order to predict the production of the following years, using different methodologies maximum number of neurons in the hidden layer and proportion of data for cross-validation, involving economic and non-economic variables that influence grain production. Dynamic neural models achieved a better performance in the simulation of coffee production in Brazil, as the difference between the mean absolute error (MAE) with static neural networks was 3000 tons of green coffee per year.

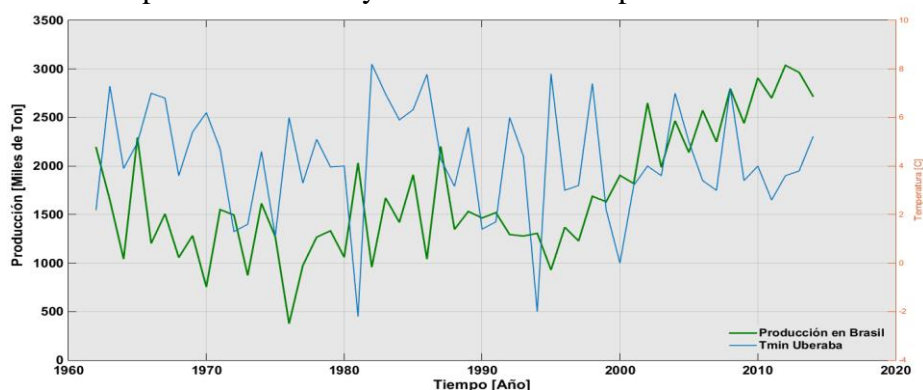
## 10 Introducción

El cultivo de café es el más importante y comercializable para 25 millones de pequeños agricultores de 60 países tropicales (Pelupessy, 2007, p. 190). En América Latina, se cultiva principalmente en Brasil, Colombia, Honduras, México, Perú, Costa Rica, Nicaragua, El Salvador, Ecuador y Venezuela.

El cultivo del café es considerado un dinamizador de la economía en las regiones donde se cultiva y cosecha, además es el producto agrícola más importante que se comercializa en el mercado de futuros de Nueva York y Londres. En los años cafetaleros 2010, 2011 y 2012, Brasil, Vietnam, Colombia e Indonesia proporcionaron más del 62% del producto a nivel internacional (ICO, 2015). Por lo tanto, lo que pase con la producción del aromático en cualquiera de estos cuatro países (principalmente Brasil) puede influir fuertemente en su cotización. En Brasil algunos consultores emplean personal capacitado y un método volumétrico para estimar la zafra o cosecha de café; en Colombia la Federación Nacional de Cafeteros utiliza un método destructivo para el pronóstico de la cosecha lo cual requiere de tiempo, y dinero.

Desde el año 1962 hasta la cosecha cafetalera del ciclo 2013/14, la oferta internacional del café ha estado sometida a fuertes variaciones, debido principalmente a condiciones climáticas extremas en la zona cafetalera de los países productores del aromático. En este periodo en Brasil se presentaron diversos eventos climáticos relacionados con heladas y sequías, destacándose principalmente los ocurridos en 1965, 1975, 1981 y 1994, los cuales fueron responsables de reducir la oferta mundial del grano en 18.1, 14.9, 17.9 y 6.3 millones de sacos de 60 kilogramos respectivamente (Figura 10).

**Figura 10** Temperatura mínima y su relación con la producción del café en Brasil



En las variaciones de la oferta del grano en Brasil desde 1962 hasta el año 2014, se identifica una estrecha relación entre la producción de café y la temperatura mínima, pues temperaturas entre 2 y -2 grados centígrados en la zona cafetalera, afecta significativamente la cosecha para el año siguiente entre un 28 y 70% dependiendo de la duración e intensidad (FAO, 2014; INMET, 2015).

En el año 2005 la producción agrícola en Brasil representó el 10% de su producto interno bruto (PIB) y la participación del sector agroalimentario fue las dos terceras partes del superávit comercial. En este mismo año el sector agroalimentario de Brasil fue el más grande del mundo con 27.5 billones de dólares. Asimismo, las exportaciones de café alcanzaron una participación del 26% del mercado mundial por un valor de 2533 millones de dólares (Valdes, 2006).

El Agronegocio del café en Brasil está pasando un momento favorable dado la producción, consumo y mercado internacional, con buenas perspectivas en el corto y mediano plazo (Santos, 2011). Debido a la importancia económica de la actividad cafetalera en Brasil y muchos países productores del grano, se hace necesario implementar herramientas que simulen y predigan con exactitud la oferta del café; una alternativa para ello son las Redes Neuronales Artificiales (RNA). Las RNA, representan una tecnología que tiene sus raíces en muchas disciplinas: neurociencia, matemáticas, estadística, física, ciencias de la computación e ingeniería (Haykin, 1999). Se han empleado en áreas tan diversas como la neurofisiología, física, ingeniería biomédica, ingeniería electrónica, ciencias de la computación, acústica, cibernética, robótica, procesamiento de imágenes, finanzas y otras (Govindaraju, 2000).

El enfoque de RNA es uno de los métodos usados para hacer mejores predicciones del futuro y es importante para la toma de decisiones, particularmente en los últimos años, como un método con alto nivel de validación en el campo de la economía y las predicciones financieras (Yildirim *et al.*, 2011).

Bahrammirzaee (2010) afirma que las RNA tienen la capacidad de actualizar los datos y ser un modelo estimador libre, además las redes neuronales son superiores a la naturaleza nominal de técnicas de manipulación simbólica porque en estas técnicas numéricas, los datos deben ser convertidos en valores nominales antes de ser utilizados como entrada, por lo tanto, hay problemas con la pérdida de información. Con las RNA se puede introducir datos numéricos directamente como entrada para el procesamiento. Las redes neuronales podrían jugar un papel importante en el esfuerzo de construir modelos cuantitativos más exactos en economía y gestión. (Vesely, 2011).

Las redes multicapas de alimentación directa o hacia adelante (multilayer feedforward) son empleadas en la construcción de redes neuronales estática, compuestas de un conjunto de neuronas organizadas lógicamente en una o más capas (Masters, 1993).

Las redes neuronales dinámicas o recurrentes, también utilizan la estructura multicapas hacia adelante, pero con uno o más circuitos de retroalimentación (Haykin, 1999). Predecir los valores futuros de una serie de tiempo es un problema común, y cualquier red neuronal que sea capaz de aceptar vectores de valor real como entradas y producir un vector de salida real, puede ser usada para predicciones de series de tiempo (Masters, 1993).

Una vez seleccionado el modelo, se debe considerar como evaluar su desempeño para determinar que tan bien el modelo representa las respuestas del sistema, en este caso la serie de tiempo. Hay muchas medidas estadísticas que describen que tan bien se ajusta el modelo a una muestra de datos, y cuando se usa más de una técnica de predicción, las medidas de desempeño ayudan a discriminar entre los modelos usados (Montgomery *et al.*, 2011).

La evaluación es importante para el uso potencial del modelo, ya que, el usuario necesita información acerca de la calidad del modelo para decidir qué tanta credibilidad debe darle al modelo. La evaluación es una de las principales actividades en todo trabajo de modelación, pero particularmente para modelos en agronomía (Wallach *et al.*, 2013).

El coeficiente de determinación ( $R^2$ ) es otro criterio estadístico, en el cual el denominador es simplemente la suma total al cuadrado de las observaciones con respecto a la esperanza, el cual es constante y no depende del modelo, y el numerador es la suma de los cuadrados residuales, por lo tanto seleccionar el modelo que maximiza el  $R^2$ , es equivalente a seleccionar el modelo que minimiza la suma al cuadrado de los residuales, por lo tanto un valor alto del  $R^2$ , sugiere un buen ajuste de los datos históricos (Ecuación 1); pero un valor alto del  $R^2$ , no asegura que los errores de predicción hacia a delante fuera de la muestra serán pequeños (Montgomery *et al.*, 2011).

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (10)$$

Wallach *et al.*, (2013), manifiestan que, los figuras son ayudas visuales invaluable en la evaluación del modelo, pero se requiere un resumen cuantitativo de la bondad de ajuste de los datos, podría decirse que el cuadrado medio del error (MSE) es una de las medidas más usadas (Ecuación 2).

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (10.1)$$

Dónde  $Y_i$  es el valor medido,  $\hat{Y}_i$  es el correspondiente valor simulado, y  $n$  es número de mediciones. El MSE mide la variabilidad (Varianza) en los errores de predicción, lo deseable es obtener la menor variabilidad en los errores de predicción (Montgomery *et al.*, 2011; Wallach *et al.*, 2013).

El MSE es la diferencia entre el valor observado y el valor simulado o predicho al cuadrado, esta medida es similar a la del error medio absoluto (MAE), excepto que la función de cuadratura es utilizada en lugar de la función del valor absoluto. Como el MSE es calculado con los errores al cuadrado de la predicción, será más sensible a grandes errores y valores atípicos en comparación con el MAE (Wilks, 2011).

Una medida alternativa para determinar el error del modelo, que al igual que el MSE evitan la compensación entre la sub o sobre predicción, es el error medio absoluto (Ecuación 3). La unidad del MAE es la misma que la de ‘Y’ o variable a predecir, por lo tanto, no hay problema de sobre ponderación de grandes diferencias empleando MAE (Wallach *et al.*, 2013).

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| \quad (10.2)$$

El MAE es la media aritmética de valores absolutos de la diferencia entre los valores observados y predichos. Claramente el MAE es cero si las predicciones son perfectas o puede incrementar si las diferencia entre los valores predichos y observados son grandes, por lo tanto, MAE es una magnitud típica para los errores de predicción en un conjunto de datos de verificación dados (Wilks, 2011).

Para modelos de agronomía el predictor genuino (naive) será a menudo el promedio de los valores observados, probablemente el predictor más ampliamente utilizado es el que determina la eficiencia del modelo definido como (Wallach *et al.*, 2013):

$$EF = 1 - \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10.3)$$

En el presente trabajo se emplearon modelos neuronales estáticos y dinámicos con el objetivo de simular la producción cafetalera en Brasil y seleccionar la mejor arquitectura y modelo basado en el desempeño estadístico de los mismos, empleando diferente número de neuronas en la capa oculta y proporción en la asignación de datos para el proceso de validación cruzada (entrenamiento, validación y prueba).

### 10.1 Metodología

La información para la simulación de la producción cafetalera en Brasil fue obtenida por la Organización Internacional del Café (ICO), el Instituto Nacional de Meteorología en Brasil (INMET), la Organización de las Naciones Unidas para la Alimentación (FAO) y corresponde al periodo entre los años 1962 a 2014; posteriormente se organizó y tabuló la información en una hoja de cálculo del programa Microsoft Excel 2013.

Las variables de entrada (31) de los modelos neuronales fueron: el área cosechada, cantidad de fertilizante empleado, producción anual de café, indicador de heladas, y sequías, la precipitación y las temperaturas máximas y mínimas de 8 estaciones climatológicas ubicadas en la zona cafetalera de Brasil. Todas las estaciones climáticas seleccionadas contaban con más del 98% de la base de datos en el periodo seleccionado. El relleno de la información faltante se realizó empleando la herramienta knnimpote del programa Matlab14b.

En los modelos neuronales estáticos y dinámicos, se emplearon tres metodologías para establecer el número de neuronas en la capa oculta, la primera es propuesta por Hecht-Nielsen (1987) ( $h = 2n + 1$ ) citada por Kůrková (1992); la segunda metodología es propuesta Masters (1993) ( $h = \sqrt{mn}$ ); finalmente se probó un criterio adicional intermedio ( $h = n$ ); donde  $n$  es el número de variables de entrada y  $m$  es el número de salidas, para este caso  $m = 1$

Para la asignación de los datos y realizar la técnica de validación cruzada (entrenamiento validación y prueba) se empleó la rutina de división aleatoria de Matlab en las siguientes proporciones (50 25 y 25%), (60 20 y 20%) y (70 15 y 15%).

El total de datos disponibles para realizar la estimación de la producción cafetalera se dividió en dos tandas, la primera con el 80% de los datos para la validación cruzada y el 20% restante se utilizó en la fase de simulación.

El mejor modelo neuronal en el proceso de simulación, fue escogido teniendo en cuenta los criterios estadísticos del mayor coeficiente de determinación ( $R^2$ ) y la eficiencia del modelo (EF) así como, el menor error medio absoluto (MAE) y cuadrado medio del error (MSE).

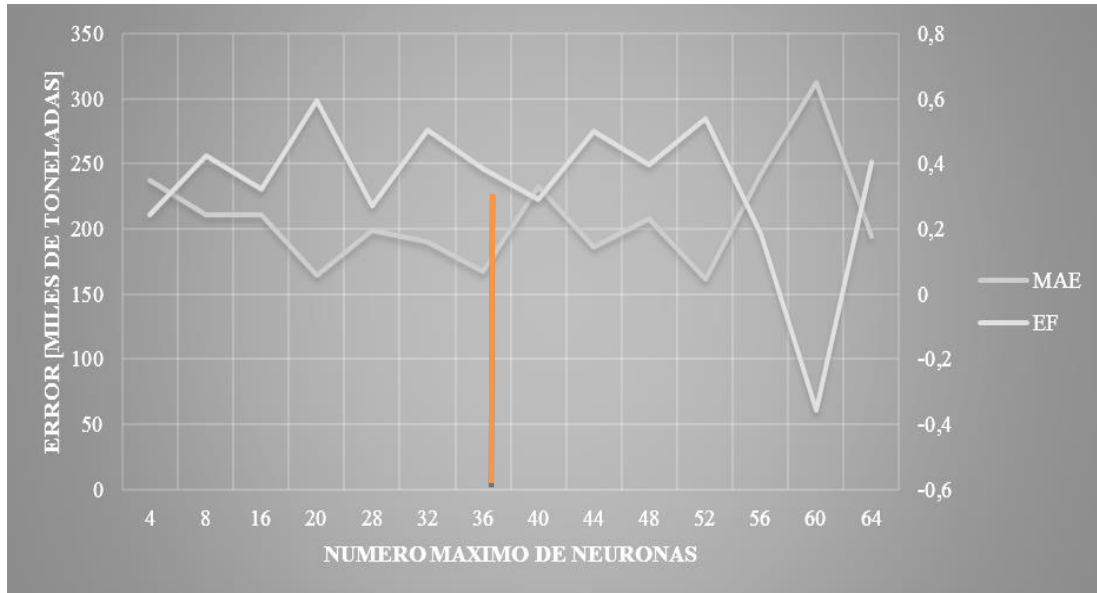
## 10.2 Resultados

De las tres metodologías para establecer el número máximo de neuronas en la capa oculta se observa que en las metodologías  $h = \sqrt{mn}$  y  $h = n$  se encuentran dentro de la metodología  $h = 2n + 1$ .

Sin embargo, el empleo de esta última metodología resulta muy compleja al aplicarlo a redes con más de 20 variables de entrada, pues se tendrían que hacer simulaciones desde 2 hasta 81 neuronas en la capa oculta, lo cual requiere de mucho tiempo de simulación por la cantidad de cálculos que tendría que hacer el ordenador.

Se determinó el número máximo de neuronas en la capa oculta donde se cruzaron las líneas del error medio absoluto (MAE) y la eficiencia del modelo (EF), para minimizar los cálculos y tiempo de simulación en las redes neuronales (Figura 10.1).

**Figura 10.1** Número máximo de Neuronas en la simulación de la Producción de café en Brasil



La metodología que minimiza los costos computacionales empleando redes neuronales estáticas para simular la producción de café en Brasil y determinar el número máximo de neuronas en la capa oculta fue la propuesta por Masters (1993) ( $h = \sqrt{mn}$ ).

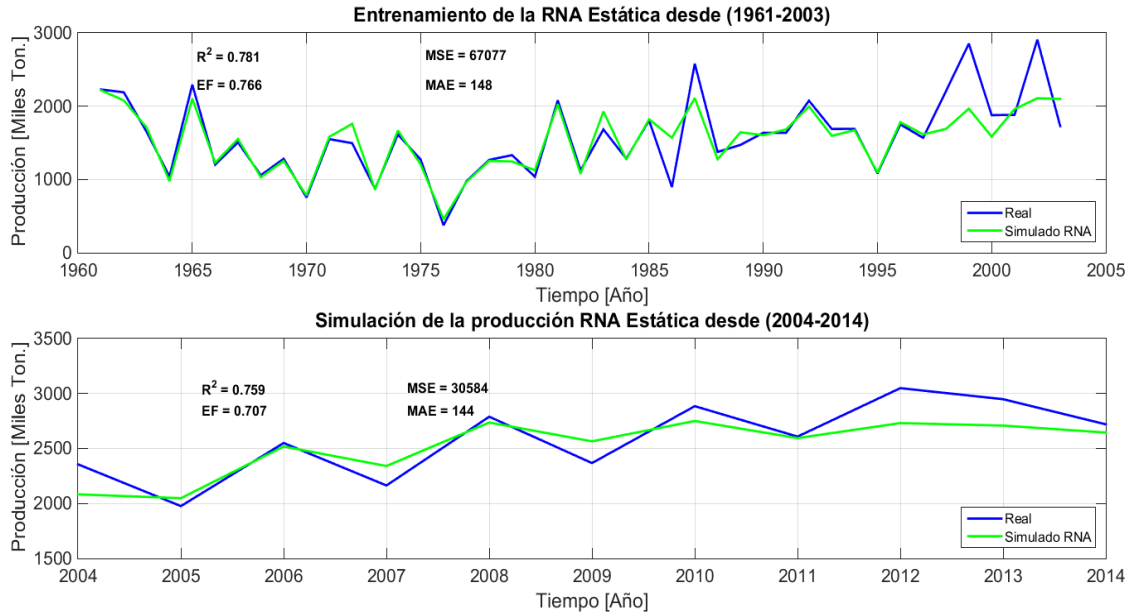
El mejor resultado para simular la producción de café en Brasil empleando redes neuronales estáticas se alcanzó con 3 neuronas en la capa oculta y distribución de datos 70-15-15 para el proceso de entrenamiento validación y prueba. Alcanzando un error medio absoluto de 144 mil toneladas (Tabla 10). Asimismo, la red estática describe el comportamiento de la oferta del grano y (Figura 10.2).

**Tabla 10** Mejor resultado de las redes neuronales estáticas.

Proceso	ítem	50-25-25	60-20-20	70-15-15
Validación cruzada	n	24	3	3
	R <sup>2</sup>	0,51	0,77	0,78
	MSE	233792	67911	67077
	EF	0,18	0,76	0,77
	MAE	407	175	148
Simulación	R <sup>2</sup>	0,76	0,72	0,76

<b>MSE</b>	363119	30283	30584
<b>EF</b>	0,65	0,71	0,71
<b>MAE</b>	162	145	144

**Figura 10.2** Resultado del mejor modelo neuronal estático

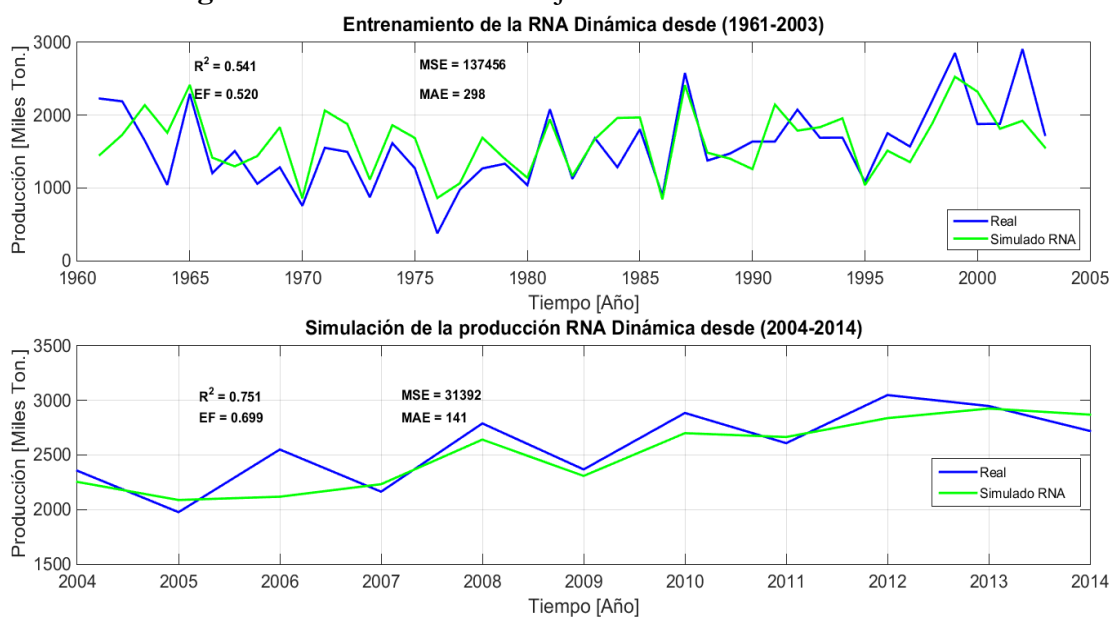


Con redes neuronales dinámicas, la metodología en la cual se encontró el número máximo de neuronas en la capa oculta para disminuir los costos computacionales fue  $h = n$  (Tabla 10.1). El empleo de un mayor número de neuronas en la capa oculta desarrollando la metodología propuesta por Hecht-Nielsen (1987) no mejoró el desempeño de las redes neuronales utilizadas en este trabajo. Asimismo, el costo computacional de la red dinámica es superior al de la red estática, debido a que los retrasos en este tipo de redes duplica el número de parámetros (pesos).

**Tabla 10.1** Resultados de las redes neuronales dinámicas.

Proceso		ítem	50-25-25	60-20-20	70-15-15
Validación cruzada		<b>n</b>	30	18	18
		<b>R<sup>2</sup></b>	0,72	0,54	0,88
		<b>MSE</b>	85748	137456	47078
		<b>EF</b>	0,7	0,52	0,84
		<b>MAE</b>	146	298	102
Simulación		<b>R<sup>2</sup></b>	0,61	0,75	0,68
		<b>MSE</b>	43674	31392	40382
		<b>EF</b>	0,58	0,7	0,61
		<b>MAE</b>	173	141	165

El mejor desempeño estadístico utilizando redes neuronales dinámicas para simular la producción cafetalera de Brasil, se logró con 18 neuronas en la capa oculta ( $h = \frac{2n}{3}$ ) y distribución de datos para el proceso de validación cruzada de, 60% para entrenamiento, 20% para la validación, y 20% para la prueba del modelo (Tabla 10.1). Logrando un error medio absoluto (MAE) fue de 141 mil toneladas de café por año (Figura 10.3).

**Figura 10.3** Resultado del mejor modelo neuronal Dinámico

El coeficiente de determinación ( $R^2$ ) como medidas de desempeño para la evaluación de un modelo, algunas veces no es una buena medida de ajuste, ya que cambia drásticamente de acuerdo a la variabilidad de los errores (Tabla 10 y 10.1).

La red neuronal dinámica logró un mayor desempeño estadístico para simular la producción de café en Brasil, en comparación, con la red neuronal estática, alcanzando una diferencia de 3 toneladas de café.

Dumedah *et al.*, (2014), probando 9 diferentes tipos de redes neuronales y 5 métodos estadísticos para el llenado de datos perdidos de humedad del suelo a diferentes profundidades en 13 estaciones de monitoreo, encontraron resultados similares, donde las redes neuronales dinámicas presentaron un mejor desempeño que las redes estáticas y otros métodos empleados.

Resultados análogos fueron encontrados por Sundermeyer *et al.*, (2013), quienes analizaron el buen ajuste en una tarea de reconocimiento de voz francesa empleando redes estáticas y dinámicas, hallando que las redes neuronales dinámicas superaron a las redes estáticas.

Asimismo, Gençay *et al.*, (1997), en su trabajo de modelación no lineal y predicción con redes estáticas y dinámicas con datos ruidosos, encontraron que la red dinámica presenta mejor desempeño que la red estática con muestras de datos pequeñas y grandes.

Sin embargo, Mańdziuk *et al.*, (2002), encontraron que las redes estáticas presentaron un mejor desempeño que las redes dinámicas. Los mismos autores encontraron que una red neuronal con dos capas es superior a la que tiene una sola capa oculta para la predicción de corto plazo en series de tiempo caótica.



### 10.3 Conclusión

Las metodologías utilizadas en este trabajo no optimizan el tiempo de simulación y no pueden encontrar el número máximo de neuronas en una capa oculta empleando redes neuronales estáticas y dinámicas. Ya que, la red neuronal estática logra su mejor desempeño estadístico con la metodología propuesta por Masters, (1993) y la red neuronal dinámica lo logró con la metodología intermedia.

La mejor metodología para distribuir los datos en el proceso de validación cruzada (Entrenamiento, Validación y Prueba) empleando redes neuronales estáticas fue 70-15-15 con un error medio absoluto (MAE) de 144 mil toneladas por año, sin embargo, la distribución 60-20-20 alcanzó un MAE de 145 mil toneladas de café. Asimismo, las redes neuronales dinámicas lograron su mejor desempeño estadístico con distribución 60-20-20 alcanzando un error medio absoluto de 141 mil toneladas de café por año.

Los modelos neuronales dinámicos presentaron un mejor desempeño en la simulación de la producción cafetalera en Brasil, pues la diferencia entre el error medio absoluto (MAE) con las redes neuronales estáticas fue de 3 mil toneladas de café por año.

### 10.4 Agradecimientos

A la Organización Internacional del café (ICO), por facilitar información relevante para la culminación del presente trabajo.

### 10.5 Referencias

- Bahrammirzaee, A. (2010). A comparative survey of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, 19(8), 1165-1195.
- Dumedah, G., Walker, J. P., & Chik, L. (2014). Assessing artificial neural networks and statistical methods for infilling missing soil moisture records. *Journal of Hydrology*, 515, 330-344.
- FAO (2014). Organización de las naciones unidas para la alimentación y la agricultura. FAOSTAT. [Fecha de consulta: 30 de agosto de 2014] Disponible en: <http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/S>
- Gençay, R., & Liu, T. (1997). Nonlinear modelling and prediction with feedforward and recurrent networks. *Physica D: Nonlinear Phenomena*, 108(1), 119-134.
- Govindaraju, R. S. (2000). Artificial neural networks in hydrology. I: Preliminary concepts. *Journal of Hydrologic Engineering*, 5(2), 115-123.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Printice-Hall. Inc., New Jersey.
- Kůrková, V. (1992). Kolmogorov's theorem and multilayer neural networks. *Neural networks*, 5(3), 501-506.

ICO (2015). International Coffee Organization. Historical Data on the Global Coffee Trade. [Fecha de consulta: 08 de marzo de 2015] Disponible en:  
<http://www.ico.org/historical/1990%20onwards/PDF/1a-total-production.pdf>

INMET (2015). Instituto Nacional de meteorología. Datos Históricos. Banco de datos históricos para enseñanza y meteorología. [Fecha de consulta: 09 de mayo de 2015]. Disponible en:  
<http://www.inmet.gov.br/portal/index.php?r=bdmep/bdmep>

Mańdziuk, J., & Mikołajczak, R. (2002). Chaotic time series prediction with feed-forward and recurrent neural nets. *Control and Cybernetics*, 31, 383-406.

Masters, T. (1993). *Practical neural network recipes in C++*. Morgan Kaufmann. 490p.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2011). *Introduction to time series analysis and forecasting* (Vol. 526). John Wiley & Sons. 441 p.

Pelupessy, W. (2007). The World behind the World Coffee Market. *Études rurales*, 189-211.

Santos, J. D. F. (2011). Conjuntura E Perspectivas Da Economia Cafeeira—Uma Breve Análise. VII Simpósio de Pesquisa dos Cafés do Brasil. Araxá-Brasil. 3 p.

Sugai, Y., Teixeira, F. A. R., Contini, E., (2004). Impacto de Exportação do Café na Economia do Brasil. EMBRAPA. 42 p.

Sundermeyer, M., Oparin, I., Gauvain, J. L., Freiberg, B., Schluter, R., & Ney, H. (2013, May). Comparison of feedforward and recurrent neural network language models. In *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on (pp. 8430-8434). IEEE.

Valdes, C. (2006). Brazil's booming agriculture faces obstacles. *Amber Waves*, 4(5), 28-35.

Veselý, A. (2011). Economic classification and regression problems and neural networks. *Agricultural Economics (Zemědělská Ekonomika)*, 57(3), 150-157.

Wallach, D., Makowski, D., Jones, J. W., & Brun, F. (2013). *Working with Dynamic Crop Models: Methods, Tools and Examples for Agriculture and Environment*. Academic Press. 487 p.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press. 676 p.

Yildirim, I., Ozsahin, S., & Akyuz, K. C. (2011). Prediction of the Financial Return of the Paper Sector with Artificial Neural Networks. *BioResources*, 6(4).