

Trabajo Final de Maestría

Una investigación científica acerca del progreso de métodos de
ensamble basados en inteligencia computacional para predicción
de series de tiempo económicas y financieras

Luisa Fernanda Rodríguez Muñoz, I.I.



Director de Tesis

Prof. Juan David Velásquez Henao, Ph.D.

Universidad Nacional de Colombia
Facultad Nacional de Minas
Escuela de Sistemas
Programa de Maestría en Ingeniería – Ingeniería de Sistemas

Noviembre de 2011

A Ana Teresa por todo su apoyo
y su infinita paciencia

Agradecimientos

A todas las personas y entidades que han sido un apoyo durante el desarrollo de esta investigación.

RESUMEN

Contexto: La combinación de pronósticos es un importante método que se usa cuando están disponibles los pronósticos de varios modelos alternativos, con el objetivo de mejorar la precisión del pronóstico. El ensamble mediante el promedio es un método de inteligencia computacional bien conocido para agregar los pronósticos de varios expertos. Ambas metodologías comparten principios y técnicas similares en un grado importante; sin embargo, no hay estudios orientados a determinar el estado actual de las investigaciones y las prácticas en dicha área común.

Objetivo: Este trabajo pretende establecer el estado actual de las investigaciones sobre las metodologías de combinación de pronósticos cuando, al menos, uno de los expertos o modelos de pronóstico o cuando el método de combinación, son herramientas de inteligencia artificial como redes neuronales artificiales, sistemas difusos o neuro-difusos.

Método: Se realizó el mapeo sistemático de 70 estudios usando ocho preguntas de investigación (los artículos seleccionados en el mapeo sistemática, se pueden encontrar comentados en el Apéndice A).

Resultados: La mayoría de los trabajos están orientados a proponer nuevas metodologías para desarrollar ensambles o combinación de pronósticos, y la diversificación es usualmente alcanzada variando un aspecto en los datos o en los expertos; sin embargo, hay preguntas muy importantes que no han sido contestadas: ¿Cuál método es mejor?, ¿Para cuáles casos?, ¿Cómo se deben seleccionar los expertos?

Conclusión: Es necesario realizar estudios posteriores para responder las preguntas anteriores y para ganar una mejor comprensión de estas metodologías.

ABSTRACT

Context: Forecast combination is an important and well-known method used when the forecasts of different alternative models are available, with the aim of improving the forecast accuracy. Ensemble-averaging is a well-known computational intelligence method for aggregate the forecasts of several experts. Both methodologies share similar principles and techniques in an important degree; however, there are not studies focused on determinate the current state of the research and practice in the common area.

Objective: This work is focused on determining the state-of-the-art about the methodologies of forecast combination when, at least, one of the experts or forecasting models or when the combiner are computational intelligence tools as artificial neural networks, fuzzy or neuro-fuzzy systems.

Method: We realize a systematic mapping of 70 studies using eight research questions.

Results: Most of the work is devoted to propose new methodologies for developing ensembles or combining forecasts and diversification is usually achieved by varying one aspect in the data or in the experts; however, important questions are not answered: which method is better? For which cases? How the experts must be selected?.

Conclusion: It is necessary to conduct studies for solving the previous questions and for gaining a better comprehension of these methodologies.

CONTENIDO

1	INTRODUCCIÓN	1
2	CONCEPTOS BÁSICOS EN LA COMBINACIÓN DE PRONÓSTICOS	5
2.1	Métodos tradicionales para realizar la combinación lineal de pronósticos	5
2.2	Principales resultados matemáticos	7
3	METODOLOGÍA	11
4	RESULTADOS.....	15
5	DISCUSIÓN	17
5.1	RQ1: ¿Qué tan fácil es encontrar artículos relevantes?	17
5.2	RQ2: ¿Cuánta actividad investigativa se ha realizado?	17
5.3	RQ3: ¿Cuáles revistas o conferencias han publicado más artículos?.....	18
5.4	RQ4: ¿Cuáles son los artículos más citados?.....	18
5.5	RQ5: ¿Cuáles son los principales modelos en esta aproximación?	19
5.5.1	RQ5.1: ¿Cuáles son los modelos usados como expertos en el sistema?.....	19
5.5.2	RQ5.2: ¿Cómo se logra la diversificación de los expertos en el ensamble?.....	22
5.5.3	RQ5.3: ¿Cuáles son los métodos usados para obtener un pronóstico combinado?	24
5.5.4	RQ5.4: ¿Qué tipos de series de tiempo han sido pronosticadas?.....	26
6	CONCLUSIONES	29
7	ANEXO 1.....	31
	BIBLIOGRAFÍA.....	45

LISTADO DE TABLAS

Tabla 1. Principales Revisas y Congresos	15
Tabla 2. Artículos más citados.....	17

LISTADO DE FIGURAS

Figura 1. Diagrama Esquemático de un Sistema General de Combinación de Pronóstico	2
Figura 2. Actividad Investigativa. (a) Número de Artículos por Año. (b) Número Total de Citaciones para el año de Publicaciones del Artículo Citado.	14

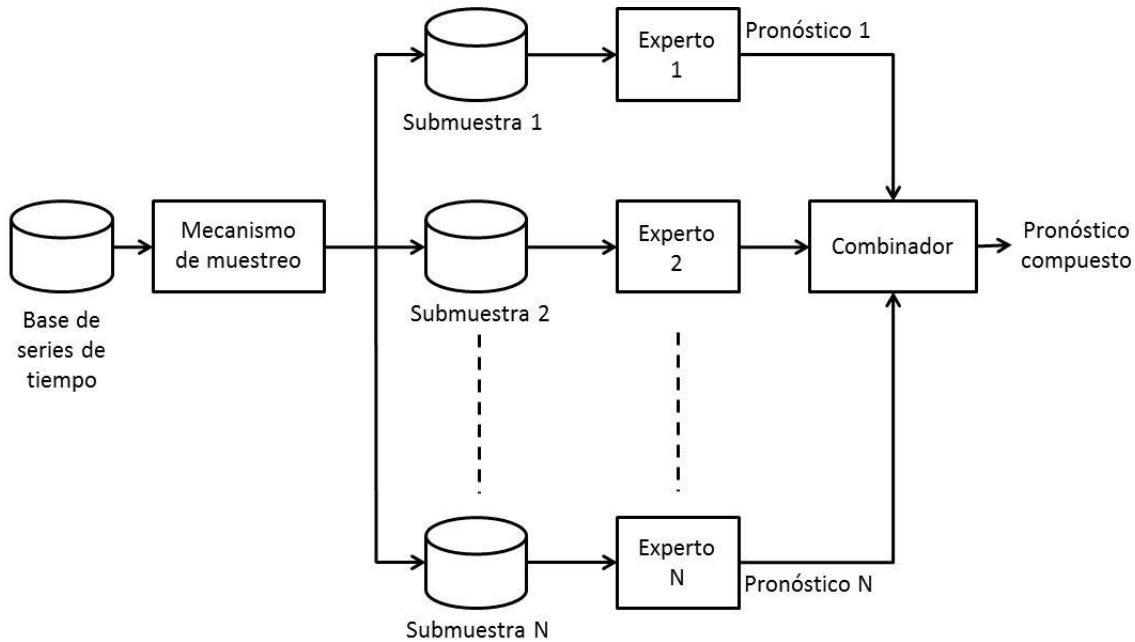
1 INTRODUCCIÓN

En este documento se consideran métodos para obtener pronósticos puntuales combinados de series de tiempo financieras y económicas como una función de los pronósticos puntuales de N modelos numéricos diferentes (o expertos en la jerga de inteligencia computacional). La idea de combinar pronósticos puntuales es utilizada en gran medida en las áreas de inteligencia computacional, estadística y econometría. En la Figura 1 se presenta un diagrama esquemático de un sistema general para la combinación de pronósticos basado en la aproximación de Haykin (1996); el sistema está compuesto de las siguientes partes:

- Una serie de tiempo que debe ser pronosticada.
- Un mecanismo para obtener submuestras de la serie de tiempo original. En los métodos estadísticos las submuestras son, usualmente, idénticas a la serie original.
- Los expertos o modelos de pronóstico. Existe un experto por cada una de las submuestras obtenidas; cada experto es ajustado a una de las submuestras.
- Un mecanismo para combinar de predicción puntual obtenida por cada uno de los expertos.
- El pronóstico compuesto que se obtiene al combinar los pronósticos de los expertos.

En inteligencia computacional, estos métodos son llamados metodologías de ensamble (Haykin, 1996). En los primeros trabajos originales, los expertos son redes neuronales del tipo perceptrón multicapa, de arquitectura idéntica, pero con diferentes valores asociados a los pesos de las conexiones internas entre las neuronas. La diversificación de los expertos es lograda considerando diferentes muestras de entrenamiento para cada experto (Perone, 1993; Breiman, 1996) o diferentes condiciones iniciales para el algoritmo de entrenamiento (Naftaly, Intrator y Horn, 1997). Los pronósticos compuestos son calculados como el promedio aritmético simple de los pronósticos individuales de cada experto o red neuronal artificial (Perone, 1993) o como una combinación lineal óptima (Breiman, 1996; Hashem y Schmeiser, 1993; Hashem & Schmeiser, 1995; Hashem, 1997) de dichos pronósticos individuales.

Figura 1. Diagrama esquemático de un sistema general de combinación de pronósticos.



En los trabajos provenientes de la estadística y la econometría, la diversidad es alcanzada usando diferentes modelos lineales bien conocidos; así, el problema principal es cómo obtener el pronóstico puntual compuesto. Las primeras contribuciones fueron focalizadas en la combinación lineal de los pronósticos puntuales (Clemen, 1989; Winkler y Makridakis, 1983; de Menezes, Bunn y Taylor, 2000; Granger, 1989; Newbold y Granger, 1974) cuando las predicciones individuales son obtenidas a partir de diferentes modelos lineales ajustados usando una función de pérdida cuadrática. Las primeras evidencias a favor de la combinación de pronósticos son presentadas por Reid (1968) y Barnard (1963), pero Bates y Granger (1969) son los primeros en proponer métodos analíticos para calcular promedios ponderados de los pronósticos individuales.

El uso de los métodos de ensamble y de combinación de pronósticos es justificado por muchas experiencias empíricas que indican que el pronóstico combinado es más preciso que el pronóstico individual para cada experto (véase las referencias citadas previamente). Hay muchas razones aceptadas para explicar por qué la combinación de pronósticos es efectiva: primero, cada predicción puede ser obtenida usando diferentes fuentes de información o cada modelo se basa en diferentes supuestos sobre el tipo de relación en los datos. Segundo, cada modelo no está en la capacidad de identificar el proceso real

de generación de los datos, pero captura información complementaria en relación con otros modelos alternativos (Trui y van Dijk, 2002). Y tercero, hay cambios fuertes en la estructura o en los valores de los parámetros del modelo cuando hay nueva información disponible; esto se debe a una especificación pobre del modelo, a la presencia de cambios estructurales (Timmerman, 2006; Hendry y Clements, 2004).

Makridakis (1989) por su parte, explica los factores que contribuyen a realizar una combinación de pronósticos. El primer factor, es que los pronósticos individuales no entregan buenos resultados, esto se ve complementado con los valores de los errores. El segundo son los patrones inestables, o los cambios constantes que por lo general se presenta en una serie de tiempo, las cuales no pueden ser tenidas en cuenta en un pronóstico individual. Tercero, en la combinación de pronóstico se utilizan modelos que minimizan el error.

Si bien es indudable la importancia y el gran interés para investigadores y profesionales en la predicción de series usando metodologías de combinación de pronósticos, no hay una visión clara sobre el progreso en este campo. En esta investigación, se aplica la metodología de mapeo sistemático, la cual fue adaptada de otras disciplinas como medicina, economía y ciencias sociales, por Kitchenham y otros (Kitchenham, 2004; Kitchenham y Charters, 2007; Kitchenham, Dybå y Jørgensen, 2004; Kitchenham et al, 2002). Las intenciones de este trabajo son: obtener un panorama sobre la investigación en métodos de combinación de pronósticos para la predicción de series de tiempo económicas y financieras; y determinar la cantidad e importancia de los artículos, realizados por pares, que han sido publicados en conferencias y revistas. Adicionalmente, se clasificó, ordenó y sintetizó las aproximaciones reportadas en los artículos, con el fin de resumir y entender las variaciones en la metodología e identificar artículos o nuevas áreas para realizar investigaciones posteriores. En la revisión sistemática de la literatura, no se han encontrado estudios previos con los mismos objetivos de esta investigación.

El interés principal de esta investigación es la intercepción de metodologías de ensamble, comúnmente utilizadas en inteligencia computacional, y las técnicas estadísticas para la combinación de pronósticos. Así, este trabajo está focalizado en los métodos donde alguno de los expertos en el modelo o el método de combinación son técnicas de inteligencia computacional tales como redes neuronales artificiales o modelos de lógica difusa.

El objetivo de este trabajo es responder las siguientes preguntas de investigación:

- RQ1: ¿Qué tan fácil es encontrar documentos relevantes?
- RQ2: ¿Qué tanta actividad investigativa a existido?
- RQ3: ¿Cuáles revistas o conferencias publican más artículos?
- RQ4: ¿Cuáles son los artículos más citados?

- RQ5: ¿Cuáles son los principales modelos en esta aproximación? Para responder a esta pregunta de investigación se consideraron las siguientes preguntas:
 - RQ5.1: ¿Cuáles son los modelos usados como expertos en el sistema?
 - RQ5.2: ¿Cómo se alcanza la diversificación para los expertos en el ensamble?
 - RQ5.3: ¿Cuáles son los métodos de combinación usados para tener el pronóstico combinado?
 - RQ5.4: ¿Cuáles son los tipos de series de tiempo que han sido pronosticados?

Este trabajo tiene valor para: (1) investigadores del área de inteligencia computacional interesados en la predicción de series de tiempo y en incorporar nuevas metodologías estadísticas en sus estudios; (2) investigadores del área de predicción estadística interesados en aproximaciones novedosas relacionadas con la combinación de pronósticos; y (3) profesionales de las áreas de inteligencia computacional y de estadística interesados en conocer cuáles son los nuevos modelos prometedores para resolver problemas prácticos de predicción.

Este documento está organizado de la siguiente forma: En el Capítulo 2 se presentan los conceptos matemáticos básicos sobre la combinación de pronósticos; en el Capítulo 3 se discute la metodología de investigación utilizada; en el Capítulo 4 se presentan los resultados obtenidos; en el Capítulo 5 se responden las preguntas de investigación; y finalmente, en el Capítulo 6 se presentan conclusiones.

2 CONCEPTOS BÁSICOS EN LA COMBINACIÓN DE PRONÓSTICOS

A continuación se revisan algunos conceptos matemáticos básicos sobre las metodologías de combinación de pronósticos, las cuales son utilizadas en el combinador presentado en la Figura 1. Tal como ya se indicó en el capítulo introductorio, los primeros trabajos relacionados con la combinación de pronósticos son debidos a Reid (1968) y Barnard (1963); en el caso más general, se considera el problema de cómo combinar los pronósticos individuales, $f_i(t)$, para $i = 1, \dots, N$, de N modelos diferentes, para obtener un único pronóstico $f^*(t)$ para una serie de tiempo $y(t)$ en el instante t .

2.1 Métodos tradicionales para realizar la combinación lineal de pronósticos

Uno de los métodos más comunes para obtener el pronóstico combinado $f^*(t)$ es el promedio aritmético simple, el cual fue utilizado originalmente en los trabajos de Reid (1968) y Barnard (1963):

$$f^*(t) = \frac{1}{N} \sum_{i=1}^N f_i(t) \quad (1)$$

Una variación a este método es el promedio aritmético con recorte (Jose y Winkler, 2008); en este caso, se desecha un porcentaje de los modelos con peor desempeño antes de calcular el promedio.

El modelo de combinación de pronósticos descrito en la ec. (1) puede ser generalizado al considerar que la combinación puede ser realizada como un promedio ponderado:

$$f^*(t) = \sum_{i=1}^N w_i f_i(t) \quad (13)$$

Surge, entonces, el problema de cómo calcular los pesos w_i de forma óptima. Bunn (1975) propone calcular los pesos w_i como una función de que tan bueno fue cada modelo en el pasado.

Bates y Granger (1969) consideran este problema cuando se tienen únicamente dos modelos alternativos y las varianzas de los errores de los dos pronósticos, σ_1^2 y σ_2^2 , son constantes en el tiempo. Sea

k es el peso de los pronósticos del primer modelo y $(1 - k)$ es el peso de los pronósticos del segundo modelo. La varianza del error de la combinación de los pronósticos está dada por:

$$\sigma_c^2 = k^2\sigma_1^2 + (1 - k)^2\sigma_2^2 + 2\rho k\sigma_1(1 - k)\sigma_2 \quad (14)$$

donde ρ es la correlación entre los errores de los pronósticos de los modelos. El valor óptimo de k que minimiza σ_c^2 está dado por:

$$k = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 - \sigma_2^2 - 2\rho\sigma_1\sigma_2} \quad (15)$$

En el caso en el que $\rho = 0$, la expresión anterior se reduce a:

$$k = \frac{\sigma_2^2}{(\sigma_1^2 + \sigma_2^2)} \quad (16)$$

Bates y Granger (1969) demuestran que el valor óptimo de σ_c^2 no podrá ser mayor que la menor de las varianzas individuales. Una generalización de este caso consiste en considerar que el valor de k puede variar a través del tiempo, así, los pesos son función del tiempo:

$$f^*(t) = k_t f_1(t) + (1 - k_t) f_2(t) \quad (17)$$

Otras aproximaciones parten del supuesto de que el pronóstico combinado puede calcularse como una combinación lineal óptima de los pronósticos individuales que minimiza una medida de error entre el pronóstico combinado y el valor real observado $y(t)$. Perrone y Cooper (1992) proponen el método de ensamble generalizado, en el cual, el pronóstico combinado es calculado usando la ec. (8), y los pesos son estimados como:

$$w_i = \frac{\sum_j C_{ij}^{-1}}{\sum_k \sum_j C_{kj}^{-1}} \quad (18)$$

donde C_{ij} son los elementos de la matriz de covarianzas de los errores de los pronósticos individuales $f_i(t)$ y $f_j(t)$; en la solución propuesta por Perrone y Cooper (1992) se cumple que $\sum_{i=1}^N w_i = 1$.

Hashem (1997) considera la determinación de los pesos en la ec. (13) como un problema de regresión en el que se desea minimizar el error cuadrático medio entre el pronóstico combinado y el valor

real $y(t)$; Hanshem (1997) considera los siguientes problemas de regresión para la determinación de los pesos:

- Combinación lineal sin restricciones con un término constante.
- Combinación lineal restringida a que los pesos deben sumar la unidad y con un término constante.
- Combinación lineal sin restricciones sin término constante.
- Combinación lineal restringida a que los pesos deben sumar la unidad y sin término constante.

Adicionalmente a las restricciones anteriores, se suele incorporar la restricción de que los pesos sean iguales o mayores que cero.

Ayda et al. (2001) proponen que los pesos deben ser calculados de forma dinámica de acuerdo con el desempeño de cada modelo considerado. Para ello, desarrollan un sistema experto basado en reglas, el cual calcula la ponderación de cada pronóstico individual en función del desempeño de cada modelo.

En el caso más general, se considera que el pronóstico combinado se obtiene al aplicar una función no lineal $G(\cdot, \dots, \cdot)$ a los pronósticos individuales, tal que:

$$f^*(t) = G(f_1(t), f_2(t), \dots, f_N(t)) \quad (19)$$

En este caso, $G()$ puede ser definida, por ejemplo, como una red neuronal artificial (Lubecke et al, 1998) o un sistema de inferencia difusa (Fiordaliso, 1998).

2.2 Principales resultados matemáticos

De acuerdo con el trabajo de Perone (1993), es posible demostrar matemáticamente la razón por la cual, el pronóstico combinado puede ser más preciso que cada uno de los pronósticos individuales. El error cuadrático medio de cada modelo i , denotado J_i , es descrito como:

$$J_i = E \left[(y(t) - f_i(t))^2 \right] = E[e_i^2(t)] \quad (2)$$

El valor medio de los errores de la predicción denotados como J se calculan como:

$$J = \frac{1}{N} \sum_{i=1}^N J_i \quad (3)$$

Ahora, el error cuadrático para el pronóstico combinado es definido como:

$$\begin{aligned} J^* &= E \left[(y(t) - f^*(t))^2 \right] = E \left[\left(y(t) - \frac{1}{N} \sum_{i=1}^N f_i(t) \right)^2 \right] = E \left[\left(\frac{1}{N} \sum_{i=1}^N y(t) - f_i(t) \right)^2 \right] \\ &= E \left[\frac{1}{N^2} \left(\sum_{i=1}^N y(t) - f_i(t) \right)^2 \right] \end{aligned} \quad (4)$$

Pero la cantidad $y(t) - f_i(t)$ es el error instantáneo del modelo i ; reescribiendo la ecuación previa, se obtiene que:

$$J^* = E \left[\left(\frac{1}{N^2} \sum_{i=1}^N e_i(t) \right)^2 \right] = \frac{1}{N^2} E \left[\left(\sum_{i=1}^N e_i(t) \right)^2 \right] \quad (5)$$

Expandiendo el último término, se puede reescribir la ecuación anterior como:

$$\begin{aligned} J^* &= \frac{1}{N^2} E \left[\sum_{i=1}^N e_i^2(t) + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N e_i(t) e_j(t) \right] \\ &= \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^N E[e_i^2(t)] \right) + \frac{1}{N^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N E[e_i(t)] \cdot E[e_j(t)] \end{aligned} \quad (6)$$

Ahora, cuando los errores son independientes e igualmente distribuidos, entonces $E[e_i(t)] \cdot E[e_j(t)] = 0$, por lo que:

$$J^* = \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^N E[e_i^2(t)] \right) = \frac{1}{N} \left(\frac{1}{N} \sum_{i=1}^N J_i \right) = \frac{1}{N} J \quad (7)$$

La última ecuación implica que el error cuadrático medio del pronóstico combinado, es la N -ésima parte del valor medio del error de predicción de cada modelo.

Naftaly, Intrator y Horn (1997) analizan el comportamiento de la varianza del error para los ensambles concluyendo que esta (la varianza) disminuye considerablemente al considerar los ensambles de pronósticos. En este caso, la varianza del pronóstico combinado será

$$\begin{aligned}
V[f^*(t)] &= E \{ [f^*(t) - E(f^*(t))]^2 \} = E \left\{ \left[\frac{1}{N} \sum_{i=1}^N f_i(t) - E \left(\frac{1}{N} \sum_{i=1}^N f_i(t) \right) \right]^2 \right\} \\
&= E \left[\left(\frac{1}{N} \sum_{i=1}^N f_i(t) \right)^2 \right] - \left(E \left[\frac{1}{N} \sum_{i=1}^N f_i(t) \right] \right)^2
\end{aligned} \tag{8}$$

El primer término de la ecuación anterior puede escribirse como:

$$E \left[\left(\frac{1}{N} \sum_{i=1}^N f_i(t) \right)^2 \right] = \frac{1}{N^2} \sum_{i=1}^N E[f_i^2(t)] + \frac{2}{N^2} \sum_{i < j} E[f_i(t) f_j(t)] \tag{9}$$

Mientras que el segundo término es:

$$\left(E \left[\frac{1}{N} \sum_{i=1}^N f_i(t) \right] \right)^2 = \frac{1}{N^2} \sum_{i=1}^N E[f_i^2(t)] + \frac{2}{N^2} \sum_{i < j} E[f_i(t)] E[f_j(t)] \tag{10}$$

Reemplazando las dos ecuaciones anteriores en la ec. (8) se obtiene que:

$$V[f^*(t)] = \frac{1}{N^2} \sum_{i=1}^N \{ E[f_i^2(t)] - (E[f_i(t)])^2 \} + \frac{2}{N^2} \sum_{i < j} \{ E[f_i(t) f_j(t)] - E[f_i(t)] E[f_j(t)] \} \tag{11}$$

De tal manera que:

$$\begin{aligned}
\frac{1}{N} V[f_i(t)] &\leq V[f^*(t)] \leq \frac{V[f_i(t)] + \max_{i,j} \{ E[f_i(t) f_j(t)] - E[f_i(t)] E[f_j(t)] \}}{Q} \\
&\leq \max_i V[f_i(t)]
\end{aligned} \tag{12}$$

Esto es, cuando se reemplaza $f_i(t)$ por $f^*(t)$, la reducción de la varianza del error es proporcional al grado de independencia entre los predictores en el ensamble.

3 METODOLOGÍA

En esta sección se describe el protocolo de investigación o los pasos metodológicos para la ejecución del alcance de este estudio.

El estudio de mapeo sistemático: En este estudio, el método de investigación utilizado es el mapeo sistemático de literatura (Kitchenham, 2004; Kitchenham et al, 2007; Kitchenham et al, 2004; Kitchenham et al, 2002), con el ánimo de responder las preguntas de investigación presentadas anteriormente. Esta metodología está basada en evidencias, es planeada formalmente, es repetible y ejecutada metodológicamente con el ánimo de responder varias preguntas de investigación, las cuales son diseñadas con el fin de proveer una visión amplia del área de investigación y una identificación y clasificación de los estudios disponibles en el área, para una investigación posterior (Kitchenham et al, 2007). Con el mapeo obtenido es posible detectar vacíos en la investigación actual y proveer las bases para nuevas investigaciones (Kitchenham, 2004; Kitchenham et al, 2007).

La investigación es ejecutada sobre estudios primarios existentes y publicados, cuyos hallazgos son considerados como hechos. Un mapeo sistemático no es una revisión ordinaria de la literatura; es una metodología para el descubrimiento de conocimiento, donde el análisis de datos es usado en la síntesis de los datos requeridos para responder las preguntas de investigación (Kitchenham, 2004; Kitchenham et al, 2007; Kitchenham et al, 2004; Kitchenham et al, 2002). Un mapeo sistemático permite al investigador obviar los siguientes problemas, los cuales están asociados a revisión de literatura tradicional, informal o narrativa: primero, la falta de una búsqueda sistemática y de métodos de análisis de la evidencia; segundo, el sesgo del investigador hacia hechos positivos; tercero, la ausencia de un proceso de investigación ordenado, reproducible y auditable; y cuarto, las deficiencias en la formulación de preguntas de investigación, e inclusive, la ausencia de preguntas (Kitchenham, 2004; Kitchenham et al, 2007; Kitchenham et al, 2004; Kitchenham et al, 2002).

Proceso de búsqueda. Se usó el sistema de indexación Scopus para la ejecución del proceso de búsqueda. La búsqueda tuvo lugar en abril de 2011 y estuvo restringida a las siguientes áreas: Ciencias

computacionales, Inteligencia artificial, Bioinformática, y Matemática. La consulta fue especificada para recuperar artículos, publicados en revistas o memorias de conferencias y congresos, hasta marzo del 2011.

Los criterios de búsqueda usados fueron:

- “ensemble” OR ("linear combination" AND "neural network") AND “time series”
- (“forecast combination” OR “forecasts combination” OR “combining forecasts” OR “combination of forecasts” OR “combined forecasts” OR “forecast combining” OR “combine forecasting”) AND “time series”.

Criterios de inclusión o exclusión: Se realizó una revisión manual del título, el resumen y las palabras claves de los artículos que fueron recuperados usando el proceso automático de búsqueda.

Se incluyeron artículos donde los siguientes criterios fueron satisfechos simultáneamente:

- (1) Se pronostica una serie de tiempo económica, financiera o caótica (benchmark).
- (2) Todos los pronósticos puntuales son obtenidos utilizando modelos matemáticos.
- (3) Al menos un experto (modelo de predicción), o el mecanismo de combinación (o combinador), es una metodología de inteligencia computacional, tal como redes neuronales, sistemas de inferencia difusa o sistemas de inferencia neuro-difusa.

Se excluyeron artículos donde alguno de los siguientes criterios es cumplido:

- (1) El trabajo está concentrado en regresión no lineal, clasificación de patrones u otros tipos de series de tiempo que no son considerados en los criterios de inclusión.
- (2) El objetivo del artículo recuperado es la combinación de pronósticos de la probabilidad de un evento (Clements y Harvey, 2011), de intervalos de probabilidad (Hansen, 2006) o de juicios de expertos (Tay y Wallis, 2000; Hall & Mitchell, 2007).
- (3) La metodología está basada en boosting (Shapire, 1990). En una máquina de boosting, los expertos son entrenados secuencialmente sobre diferentes distribuciones de datos, donde cada experto se concentra sobre los ejemplos no aprendidos adecuadamente por los expertos anteriores.
- (4) El modelo usado es una mezcla-de-expertos (Jacobs, Jordan, Nowlan, Hinton, 1991; Jacob y Jordan, 1991; Jacob y Jordan, 1994). En este caso, cada experto es ajustado sobre una región definida del espacio de entrada y una *gating network*, usando información del espacio de entrada, para determinar la salida del sistema.
- (5) El artículo se concentra en el problema de cómo medir la exactitud de la predicción de cada modelo individual y la comparación entre las predicciones individuales (que usualmente tienen una exactitud similar), con el fin de determinar cuándo son

estadísticamente diferentes y cuándo existe información complementaria (Diebold y López, 1996; Newbold y Harvey 2002; Clements y Harvey, 2009).

Selección de estudio y análisis de datos: Los siguientes datos fueron extraídos de cada artículo seleccionado:

- Autores.
- Nombre de la revista o conferencia.
- Año de publicación.
- Citaciones al artículo.
- Tipos de expertos.
- Metodología para la combinación de pronósticos.
- Tipos de casos de estudio.

Desviaciones del protocolo de investigación: No hay desviaciones del protocolo.

4 RESULTADOS

Inicialmente se mostró un esquema de combinación tradicional que vela por alcanzar la mejor precisión del pronóstico, para esto puede usar diferentes muestras de entrenamiento para los expertos o diferentes modelos o “expertos”.

Se demostró la validez matemática de las metodologías de ensamble, en especial de las metodologías tradicionales, como lo son el combinador de promedio simple, promedio ponderado, promedio ponderado con pesos optimizados, el método del ensamble generalizado, regresiones, modelos dinámicos.

Usando las cadenas de búsqueda descritas en la sección anterior, se encontraron más de 1500 artículos potenciales. Después de realizar una revisión manual aplicando criterios de inclusión y exclusión definidos, se seleccionaron 40 artículos publicados en revistas y 30 artículos de conferencias. Así, se usaron 70 manuscritos para responder las preguntas de investigación. El resultado de estas experiencias de investigación, justifican de alguna manera el uso de estas metodologías de ensamble o combinación de pronósticos, ya que se encuentra que esta es mucho más precisa que las metodologías de pronóstico individual de cada experto.

Los resultados de este trabajo, son un aporte al conocimiento encontrado alrededor de este tipo de metodologías ya que entrega un diagnóstico del estado de la investigación y se determina la cantidad e importancia de los artículos publicados en conferencias y revistas, seleccionados para el análisis; Se encontraron vacíos en la investigación y por ende oportunidades de trabajos posteriores.

5 DISCUSIÓN

En esta sección se responden las preguntas de investigación.

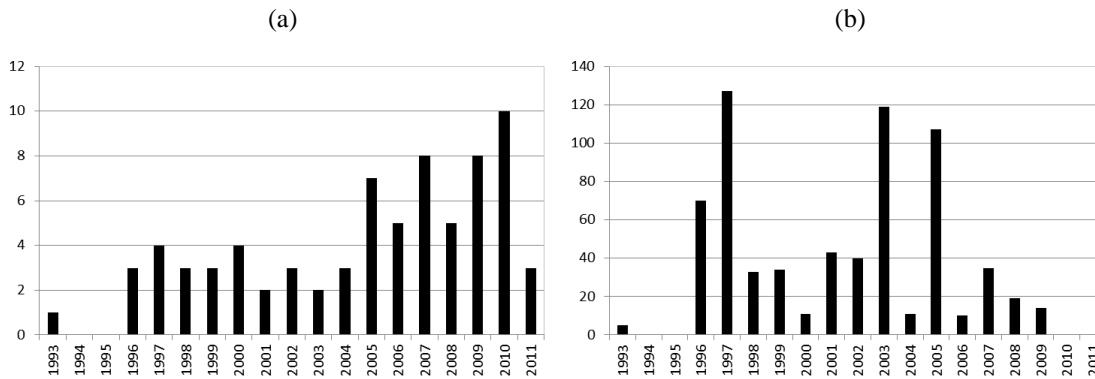
5.1 RQ1: ¿Qué tan fácil es encontrar artículos relevantes?

Realmente es difícil. Hay un importante cuerpo de literatura sobre combinación de pronósticos y métodos de ensamble con más de cuarenta y veinte años de investigación respectivamente. Más aún, el término “combinación de pronósticos” no solamente cubre la predicción puntual de un valor, sino también el pronóstico de probabilidades, la combinación de la densidad de probabilidad de los pronósticos y la predicción de intervalos. En inteligencia computacional, las metodologías de ensamble no son solamente usadas para predicción de series de tiempo sino también en problemas de clasificación y regresión no lineal. En esta investigación únicamente el 5% de los artículos recuperados fueron relevantes.

5.2 RQ2: ¿Cuánta actividad investigativa se ha realizado?

La motivación para esta pregunta es determinar la importancia actual de este tema. La cantidad de estudios publicados por año (Figura 2a) y la cantidad de citaciones de artículos publicados en un año determinado (Figura 2b) son medidas de la cantidad y calidad de la investigación. En la Figura 2a se aprecia que gran parte del trabajo ha sido publicado entre 2005 y 2011; para este período hay un promedio de 7 estudios por año. En la Figura 2b se concluye que gran parte del trabajo más importante, fue publicado en 1993, 1997, 2003 y 2005. Se concluye que el interés en la investigación sobre métodos de combinación de pronósticos está creciendo.

Figura 2. Actividad investigativa. (a) Número de artículos por año. (b) Número total de citas para el año de publicación del artículo citado.



5.3 RQ3: ¿Cuáles revistas o conferencias han publicado más artículos?

En la Tabla 1 se listan las revistas y las conferencias que han publicado al menos dos artículos. Hay 48 publicaciones con un solo artículo relevante para esta investigación. El International Journal of Forecasting publicó tres de los artículos más citados en esta investigación (RQ4) con un total de 96 citas; así, esta revista es la más importante en esta área.

5.4 RQ4: ¿Cuáles son los artículos más citados?

En la Tabla 2 se presentan los artículos con más de 20 citas: dos artículos están relacionados con la combinación no lineal de pronósticos (Fiordaliso, 1998; Donalson y Kamstra, 1996), cuatro artículos describen las metodologías para la construcción de ensambles (Islam, Yao y Murase, 2003; Kim y Kim, 1997; Yu, Wang y Lai, 2005; Zhang, 2007); finalmente, los siguientes artículos presentan casos de aplicación: Teräsvirta, van Dijk y Medeiros (2005), Terui y van Dijk (2002), Zhang y Berardi (2001) y Harrald y Kamstra (1997).

Tabla 1. Principales revistas y congresos.

Nombre	Número de artículos	Refs.
IEEE International Conference on Neural Networks	7	[7][20][29][44][45][48][61]
International Journal of Forecasting	5	[3] [18] [55] [56][60]
Journal of Forecasting	2	[10] [15]
Lecture Notes in Computer Science	5	[5] [30][39][58][41]
Neurocomputing	3	[32][62][66]
International Joint Conference on Neural Networks	3	[12] [25] [50]
Studies in Computational Intelligence	2	[42] [53]

5.5 RQ5: ¿Cuáles son los principales modelos en esta aproximación?

5.5.1 RQ5.1: ¿Cuáles son los modelos usados como expertos en el sistema?

Para responder esta pregunta se consideró que el sistema puede tener expertos del mismo o de diferente tipo. En la literatura sobre predicción, la práctica tradicional consiste en usar diferentes tipos de modelos en el ensamble con el propósito de capturar diversas características de la serie de tiempo. Los modelos más comúnmente usados son: redes neuronales artificiales, ARIMA, suavizado exponencial y regresión polinomial. En inteligencia computacional, lo más común es que el mismo tipo de modelo sea usado para todos los expertos. En este caso, el perceptrón multicapa y las máquinas de vectores de soporte son los tipos de modelos usados con mayor frecuencia.

Tabla 2. Artículos más citados.

Autores	Título (<i>Nombre de la revista</i>)	Año / Número / Referencias de citas
Islam M.M., Yao X., Murase K.	A constructive algorithm for training cooperative neural network ensembles (<i>IEEE Transactions on Neural Networks</i>)	2003 119 [26]
Kim D., Kim C	Forecasting time series with genetic fuzzy predictor ensemble (<i>IEEE Transactions on Fuzzy Systems</i>)	1997 98 [28]
Donaldson, R. G., Kamstra M.	Forecast combining with neural networks (<i>Journal of Forecasting</i>)	1996 67 [15]
Yu, L., Wang, S., Lai, K.K.	A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates (<i>Computers & Operations Research</i>)	2005 63 [67]
Terui N., Van Dijk H.K.	Combined forecasts from linear and nonlinear time series models (<i>International Journal of Forecasting</i>)	2002 36 [56]
Terasvirta T., van Dijk D., Medeiros M.C.	Linear models, smooth transition autoregressions, and neural networks for forecasting macroeconomic time series: A re-examination (<i>International Journal of Forecasting</i>)	2005 35 [55]
Zhang G.P., Berardi V.L	Time series forecasting with neural network ensembles: An application for exchange rate prediction (<i>Journal of the Operational Research Society</i>)	2001 28 [71]
Harrald P.G., Kamstra M.	Evolving artificial neural networks to combine financial forecasts (<i>IEEE Transactions on Evolutionary Computation</i>)	1997 25 [22]
Fiordaliso A.,	A nonlinear forecasts combination method based on Takagi-Sugeno fuzzy systems (<i>International Journal of Forecasting</i>)	1998 25 [18]
Zhang G.P.	A neural network ensemble method with jittered training data for time series forecasting (<i>Information Sciences</i>)	2007 23 [69]

Entre los sistemas que usan el mismo tipo de modelo, se han tenido en cuenta las siguientes metodologías como expertos:

- ANFIS (Chen y Zhang, 2007; Soto, Castillo y Soria, 2010)
- Redes asociativas competitivas (Kurogi, Koyama, Tanaka y Sanuki, 2007)
- Máquinas de aprendizaje extremo (Van Heeswijk et al, 2009)
- Modelos difusos de series de tiempo (Aladag, Egrioglu y Yolcu, 2010; Kim & kim, 1997)

- Perceptrones multicapa (Chiewchanwattana y Lursinsap, 2002; Islam, Yao y Murase, 2003; Palit y Popovic, 2000; Peng y Zhu, 2009; Peralta, Gutierrez y Sanchis, 2010; Ruta y Gabrys, 2007; Schwaerzel y Rosen, 1997; Zhang, 2007; Zhang y Berardi, 2001; Lin & Zhu, 2007; Zhao, Zhang y Liao, 2008).
- Redes neuronales con función de base radial (Wang y Li, 2010; Yu, Lai y Wang, 2008)
- Redes neuronales autogenerativas (Inoue y Narihisa, 2000; Inoue y Narihisa, 2000; Zengli y Qui, 2003)
- Máquina de vectores de soporte (Cai, Hu y Tao, 2007; Deng, Jin y Zhong, 2005; Yang y Ma, 2010; Yang & Shi, 2010).

En el caso de los sistemas que usan diferentes tipos de expertos, las metodologías empleadas son las siguientes:

- ANFIS (Lemke y Gabrys, 2010; Leone, Ohishi y Ballini, 2006)
- Redes neuronales artificiales (Andrawis, Atiya y Shishiny, 2001; Landassuri y Bullinaria, 2009; Leonel, Ohishi y Ballini, 2006; Rutam, Gabrys y Lemke, 2011; Santana, 2006; Shi, Xu y Liu, 1999; Wichard y Ogozalek, 2004; Wichard y Ogozalek, 2004; Wichard y Ogozalek, 2007; Yu, Wang y Lai, 2005)
- Redes neuronales autorregresivas (Teräsvirta, van Dijk y Medeiros, 2005)
- ARIMA (Badri, 1997; Fiordaliso, 1998; Gomes et al, 2006; Palit y Popovic, 2000; Ruta y Gabrys, 2007; Shi y Liu, 1993; Xiao, Gong y Zou, 2009)
- DAN2 (Gomes et al, 2006)
- Redes de Elman (Lemke y Gabrys, 2010)
- Modelos autorregresivos exponenciales (Terui y Dijk, 2002)
- GARCH exponenciales (Hu y Tsoukalas, 1999)
- Suavizado exponencial (Badri, 1997; Badri, Mutawa y Murtagy, 1998; Han, Niu y Yu, 2007; Lemke y Gabrys, 2010; Palit y Popovic, 2000; Ruta y Gabrys, 2007; Santana, 2006; Shi, Xu y Liu, 1999; Shi y Liu, 1993; Xiao, Gong y Zou, 2009)
- GARCH (Donaldson y Kamstra, 1999; Dong, 2002; Harrald y Kamstra, 1997; Tsangari, 2007; Hu y Tsoukalas, 1999)
- Regresión gaussiana (Andrawis, Atiya y El-Shishiny, 2011)
- Autorregresión lineal generalizada (Yu, Wang y Lai, 2005)
- Modelos lineales autoregresivos (Prudencio y Ludermir, 2006; Teräsvirta, van Dijk y Medeiros, 2005; Terui y Dijk, 2002)

- Modelos de medias móviles (Ruta, Gabrys y Lemke, 2011; Xiao, Gong y Zou, 2009)
- Varianza de medias móviles (Donaldson y Kamstra, 1999; Dong, 2002; Harrald y Kamstra, 1997; Hu y Tsoukalas, 1999)
- Perceptrones multicapa (Sun, Xie, Wang y Ahang, 2010)
- Regresión lineal múltiple (Adrawis, Ativa y El-Shishiny, 2011; Fiordaliso, 1998; Han, Niu y Yu, 2007; Leone, Ohishi y Ballini, 2006)
- Modelos del vecino más cercano (Wichard, 2011; Wichard y Ogorzalek, 2004; Wichard y Ogorzalek, 2007)
- Modelos de la trayectoria más cercana (Wichard, 2011; Wichard y Ogorzalek, 2004; Wichard y Ogorzalek, 2007)
- Redes híbridas combinando redes de base radial y perceptrones (Wichard y Ogorzalek, 2004; Wichard & Ogorzalek, 2007)
- Regresión polinomial (Lemke y Gabrys, 2010; Ruta y Gabrys, 2007; Shi, Xu y Liu, 1999; Shi y Liu, 1993; Wichard y Ogorzalek, 2004; Wichard y Ogorzalek, 2007)
- Redes neuronales de base radial (Lemke y Gabrys, 2010)
- Caminata aleatoria (Prudencio y Ludemir, 2006; Ruta y Gabrys, 2007)
- Redes neuronales recurrentes (Lemke y Gabrys, 2010; Leone, Ohishi y Ballini, 2006; Sun, Wang y Zhang, 2010)
- Modelos autorregresivos de transición suave (Teräsvirta, van Dijk y Medeiros, 2005)
- Promedio simple móvil (Andrawis, Atiya y El-Shishiny, 2011)
- Modelos estructurales de series de tiempo (Lemke y Gabrys, 2010; Ruta y Gabrys, 2007)
- Método Theta (Lemke y Gabrys, 2010; Ruta y Gabrys, 2007)
- Modelo autorregresivo con umbral (TAR) (Terui y Dijk, 2002)

5.5.2 RQ5.2: ¿Cómo se logra la diversificación de los expertos en el ensamble?

Varias estrategias han sido usadas para lograr la diversificación entre los expertos del sistema. La más común es la variación del tipo de expertos, seguida de la variación de los datos utilizados para entrenar (ajustar o calibrar) cada experto; dicha variación de los datos es obtenida al aplicar metodologías como validación cruzada, bootstrapping o bagging.

A continuación se listan las metodologías que han sido reportadas en la literatura para realizar la diversificación de los expertos:

(1) Mediante la variación de los datos:

- Incorporación de ruido a los datos (Ruta, Gabrys y Lemke, 2011; Zhang, 2007)
- Bagging (Deng, Jin y Zhong, 2005; Kurogi, Koyama, Tanaka y Sanuki, 2007; Wang y Li, 2010; Lin y Zhu, 2007)
- Boosting (Wang y Li, 2010)
- Bootstrapping (Chen y Zhang, 2005; Deng, Jim y Zhong 2005; Inoue y Narihisa, 2000; Inoue y Narihisa, 2000; Lai, Yu, Wang y Wei, 2006)
- Fuzzy possibility C -means (Yang y Ma; 2010; Yang & Shi, 2010)
- Validación cruzada (Palit y Popovic, 2000; Peng y Zhu, 2009; Peralta, Gutierrez y Sanchis, 2010; Ruta y Gabrys, 2007; Ruta, Gabrys y Lemke, 2011; Wichard y Ogorzalek, 2004; Wichard y Ogorzalek, 2007; Zhang y Berardi, 2001)
- Muestreo al azar sin reemplazo (Chen y Zhang, 2005)
- Selección de diferentes entradas para el modelo (Ruta, Gabrys y Lemke, 2011; Badri, 1997)
- Stacking (Yu, Lai y Wang, 2008)
- Thick modeling (Riedel y Gabrys, 2007).

(2) Mediante la variación de la función de pérdida o error:

- Función de pérdida lineal (Deng, Jin y Zhong, 2005)
- Función de pérdida cuadrática (Deng, Jin y Zhong, 2005).

(3) Mediante la variación del tipo de expertos (Ahmed, Tang y Abdullah, 2009; Aladag, Egrioglu y Yolcu, 2010; Andrawis, Atiya y El-Shishiny, 2011; Donaldson y Kamstra, 1999; Dong, 2002; Fiordaliso, 1998; Gomes et al, 2006; Han, Niu y Yu, 2007; Landassuri y Bullinaria, 2009; Lemke y Gabrys, 2010; Lubeke et al, 1998; Prudencio y Ludermir, 2006; Pulido, Mancilla y Melin, 2009; Riedel y Gabrys, 2007; Ruta, Gabrys y Lemke, 2011; Santana, 2006; Shi, Xu y Liu, 1999; Shi y Liu, 1993; Shi, 2009; Sun, Xie, Wang y Zhang, 2010; Teräsvirta, van Dijk y Medeiros, 2005; Terui y van Dijk, 2002; Tsangari, 2007; Wichard, 2011; Wichard y Ogorzalek, 2004; Xiao, Gong y Zou, 2009; Hu y Tsoukalas, 1999; Yu, Wang y Lai, 2005).

(4) Mediante la variación del tipo de algoritmo de entrenamiento de cada experto (Badri, Al-Mutawa y Murtagy, 1998; Chiewchanwattana y Lursinsap, 2002; Lai, Yu y Wang, 2006; Leone, Ohishi y Ballini, 2006; Ruta, Gabrys y Lemke, 2011; Schwaerzel y Rosen, 1997; Soto, Castillo y Sonia, 2010; Van Heeswijk et al, 2009; Zhang y Berardi, 2001).

- (5) Mediante la variación del método de especificación o de la configuración de los expertos (Pulido, Mancilla y Melin, 2009; Teräsvirta, van Dijk y Medeiros, 2005; Yu, Lai y Wang, 2008; Li y Deng, 2008).
- (6) Mediante el uso de un algoritmo de optimización basado en poblaciones (Harrald y Kamstra, 1997; Islam, Yao y Murase, 2003; Kim y Kim, 1997; Landassuri y Bullinaria, 2009; Zeng y Qui, 2003).
- (7) Mediante la variación de los insumos del modelo (Lin y Zhu, 2007).

5.5.3 RQ5.3: ¿Cuáles son los métodos usados para obtener un pronóstico combinado?

Los métodos más usados para combinar pronósticos son la media aritmética y la combinación lineal óptima. Muchos de los trabajos están basados en la combinación de pronósticos de todos los expertos en el ensamble; sólo en (Landassuri y Bullinaria, 2009) es considerado el problema de la selección apropiada de las variables de entrada de los combinadores. En el caso de la combinación lineal óptima, sólo un estudio utiliza la función de pérdida lineal (Prudencio y Ludermir, 2006). A continuación se enumeran las metodologías reportadas en la literatura:

(1) Para los modelos que usan el mismo tipo de expertos:

- Basados en la varianza (Chen y Zhang, 2005; Inoue y Narihisa, 2000)
- Ensamble de los mejores individuos (Landassuri y Bullinaria, 2009)
- Modelos difusos (Pulido, Mancilla y Melin, 2009)
- Modelo de síntesis difusa (Yang y Ma, 2010; Yang y Shi, 2010)
- Perceptrones multicapa (Aladag, Egrioglu y Yolcu, 2010; Donaldson y Kamstra, 1999; Landassuri y Bullinaria, 2009; Donaldson y Kamstra, 1996)
- Combinación lineal óptima (Kurogi, Koyama, Tanaka y Sanuki, 2007; Palit y Popovic, 2000; Peng y Zhu, 2009; Peralta, Gutierrez y Sanchis, 2010; Ruta y Gabrys, 2007; Soto, Castillo y Sori, 2010; Zeng y Qui, 2003; Yu, Lai y Wang, 2008; Li y Deng, 2008)
- Ordenamiento basado en una combinación lineal (Landassuri y Bullinaria, 2009)
- Funciones de base radial (Yu, Lai y Wang, 2008)
- Conjuntos rugosos (Ahmed, Yang y Abdullah, 2009)
- Media aritmética simple (Deng, Jin y Zhong, 2005; Inoue y Narihisa, 2000; Islam, Yao y Murase, 2003; Landassuri y Bullinaria, 2009; Palit y Popovic, 2000; Peng y

Zhu, 2009; Ruta y Gabrys, 2007; Schwaerzel y Rosen, 1997; Soto, Castillo y Sori, 2010; Van Heeswijk et al, 2009; Ribeiro, Goldschmidt y Choren, 2010)

- Máquinas de vectores de soporte (Lai, Yu y Wang, 2006; Wang y Li, 2010).

(2) Para los modelos que usan diferentes tipos de expertos:

- Combinación dinámica (dynamic pooling) (Riedel y Gabrys, 2007)
- Redes neuronales difusas (Dong, 2000)
- Sistema de inferencia difusa (Palit y Popovic, 2000; Xiao, Gong y Zou, 2009)
- Sistema generalizado difuso de Takagi-Sugeno (Fiordaliso, 1998)
- Métodos de kernel (Chen y Zhang, 2005; Sun, Xiw, Wang y Zhang, 2010)
- Perceptrones multicapa (Shi, Xu y Liu, 1996; Badri, 1997; Badri, Al-Mutawa y Murtagy, 1998; Donaldson y Kamstra, 1999; Harrald y Kamstra, 1997; Shi-Ying, 2009; Shi y Liu, 1993; Hu y Tsoukalas, 1999)
- Redes neuronales neuro-difusas (Palit y Popovic, 2000)
- Combinación óptima de pesos (Han, Niu y Yu, 2007; Harrald y Kamstra, 1997; Lubecke et al, 1998; Tasangari, 2007; Wang y Li, 2010; Yu, Wang y Lai, 2005)
- Por minimización del error medio absoluto sujeto a que los pesos de ponderación sumen la unidad (Prudencio y Ludermir, 2006)
- Método de mejor desempeño (Lemke y Gabrys, 2010; Ruta y Gabrys, 2011)
- Promedio aritmético simple (Andrawis, Atiya y El-Shishiny, 2011; Lemke y Gabrys, 2010; Lubecke et al, 1998; Ruta y Gabrys, 2011; Teräsvirta, van Dijk y Medeiros, 2005)
- Promedio simple con exclusión de datos (Lemke y Gabrys, 2010; Ruta y Gabrys, 2011)
- Sistema difuso de Takagi-Sugeno (Fiordaliso, 1998; Shi-Ying, 2009)
- Combinación lineal cambiante en el tiempo (Terui y Van Dijk, 2002)
- Pesos basados en la varianza (Chiewchanwattana y Lursinsap, 2002; Gomes et al, 2006; Kim y Kim, 1997; Lemke y Gabrys, 2010; Ruta y Gabrys, 2011; Schwaerzel & Rosen, 1997)

5.5.4 RQ5.4: ¿Qué tipos de series de tiempo han sido pronosticadas?

La mayoría de los modelos propuestos son evaluados mediante la comparación de su precisión, cuando se pronostican series de tiempo caóticas. En la revisión sistemática de la literatura, se encontró que el problema de la volatilidad del pronóstico de series de tiempo se ha abordado también por (Harrald y Kamstra, 1997; Hu y Tsoukalas, 1999; Donaldson y Kamstra, 1999; Dong, 2002).

Los tipos de series de tiempo que han sido pronosticadas usando metodologías de combinación de pronósticos son:

- Series de tiempo caóticas comúnmente usadas como benchmark (Chen y Zhang, 2005; Chiewchanwattana y Lursinsap, 2002; Deng et al, 2005; Inoue y Narihisa, 2000; Inoue, 2000; Islam, Yao y Murase, 2003; Landassuri-Moreno y Bullinaria, 2009; Palit y Popovic, 2000; Peng y Zhu, 2009; Peralta, Gutierrez y Sanchis, 2010; Soto, Castillo y Soria, 2010; Yang y Ma, 2010; Yang y Shi, 2010; Zeng y Qiu, 2003; Li y Deng, 2008; Inoue y Narihisa, 2005)
- Series de tiempo no lineales (no caóticas) comúnmente usadas como benchmark (Chiewchanwattana y Lursinsap, 2002; Deng, Jin & Zhong, 2005; Fiordaliso, 1998; Gomez et al, 2006; Landassuri-Moreno y Bullinaria, 2009; Peralta, Gutierrez y Sanchis, 2010; Terui y van Dijk, 2002; Wichard y Ogorzałek, 2004)
- Series de tiempo económicas (Han, Niu y Yu, 2007; Santana, 2006; Teräsvirta, van Dijk y Medeiros, 2005; Terui y van Dijk, 2002; Van Heeswijk et al, 2009; Xiao, Gong y Zou, 2009; Lin y Zhu, 2007)
- Tasas de cambio (Lai, Yu, Wang y Wei, 2006; Lubecke, Nam, Markland y Kwok, 1998; Schwaerzel & Rosen, 1997; Sun, Wang & Zhang, 2010; Tsangari, 2007; Zhang & Berardi, 2001; Yu, Wang & Lai, 2005; Yu, Lai & Wang, 2008)
- Volatilidad de las tasas de cambio (Hu, Tsoukalas, 1999)
- Precios del mercado de valores (Shi, Xu y Liu, 1999; Shi & Liu, 1993)
- La volatilidad de los rendimientos de acciones en mercados de valores (Donaldson y Kamstra, 1999; Dong, 2002; Harrald y Kamstra, 1997)
- Índices de mercados de valores (Aladag, Egrioglu y Yolcu, 2010; Bouchachia, 2009; Lai, Yu, Wang & Wei, 2006; Pulido, Mancilla & Melin, 2009; Wang & Li, 2010)
- Bases de datos de series de tiempo usadas en competencias de pronósticos (Andrawis, Atiya y El-Shishiny, 2011; Kurogi, Koyama, Tanaka & Sanuki, 2007; Lemke & Gabrys,

2010; Prudencio & Ludermir, 2006; Ruta & Gabrys, 2007; Ruta, Gabrys & Lemke, 2011; Wichard, 2011; Wichard & Ogorzałek, 2007).

6 CONCLUSIONES

En este trabajo se presentan los resultados de un estudio de mapeo sistemático realizado con el fin de determinar el estado actual de las investigaciones sobre la combinación de pronósticos y los métodos de ensamble de pronósticos, cuando algunos de los expertos del sistema o la técnica de combinación es una metodología propia de la inteligencia computacional.

La metodología empleada permitió recuperar 70 estudios relevantes respecto a las preguntas de investigación abordadas en este trabajo; se encontró que este tipo de metodologías han sido abordadas en estadística y en inteligencia computacional bajo ópticas diferentes en sus inicios, pero la evolución de los trabajos permite observar la convergencia de las investigaciones. La temática investigada es relevante actualmente de tal forma que se ha dado un incremento en el número de publicaciones en los últimos seis años.

Las principales investigaciones sobre esta temática han sido ampliamente difundidas en la comunidad científica a través de la publicación de artículos en revistas de alto impacto, tal como el *International Journal of Forecasting* o las memorias de la *IEEE International Conference on Neural Networks*.

Los resultados obtenidos muestran una amplia gama de técnicas utilizadas como expertos o para la combinación de los pronósticos; no obstante, los trabajos presentan dispersión respecto a las series de tiempo utilizadas, de tal manera que es imposible que se puedan extraer recomendaciones sobre cuáles podrían ser las metodologías más apropiadas, e inclusive hablar de una metodología para la construcción de este tipo de modelos.

La investigación realizada es limitada en el sentido de que sólo se consideraron aquellos artículos en que uno de los expertos o la metodología para la combinación de pronósticos es una técnica de inteligencia computacional. De esta forma, no se tuvo en cuenta un amplio cuerpo de literatura sobre la combinación de pronósticos, pero, que en un momento determinado debería vincularse para alcanzar una visión más general sobre el área de la combinación de pronósticos.

7 ANEXO 1

A continuación se presenta el listado de los artículos utilizados en la revisión sistemática de la literatura:

1. Ahmed, E. F., Yang, W. J., Abdullah, M. Y. (2009). Novel Method of the Combination of Forecasts Based on Rough Sets. *Journal of Computer Science*, 5 (6), 440-444.

Este estudio utiliza la metodología AHP (Analytic Hierarchy Process) para estimar los pesos de los atributos en un modelo de combinación de pronósticos.

2. Aladag, C. H., Egrioglu, E., Yolcu, U. (2010). Forecast Combination by using Artificial Neural Networks. *Neural Processing Letters*, 32 (3), 269-276.

En este caso, el pronóstico es obtenido mediante diferentes modelos de inferencia difusa, y son combinados utilizando redes neuronales artificiales. Esta metodología es aplicada a una serie de tiempo del índice IMKB, y los resultados son comparados con la literatura. Los autores demuestran las mejoras en cuanto a la precisión del pronóstico que entrega la metodología de combinación propuesta.

3. Andrawis, R. R., Atiya, A. F., El-Shishiny, H. (2011). Forecast Combinations of Computational Intelligence and Linear Models for the NN5 Time Series Forecasting Competition. *International Journal of Forecasting*, 27 (3), 672-688.

El modelo de combinación propuesto utiliza como expertos una red neuronal, un proceso de regresión Gaussiano y un modelo lineal, los cuales son combinados unas un promedio simple. Se hace énfasis en la estacionalidad de las series.

4. Ashour Z.H., Hashem S.R., Fayed H.A. (2008). Combining neural networks during training for real time series modeling and forecasting. *Journal of Engineering and Applied Science*.

Los autores desarrollan una serie de combinación de modelos, con el fin de probar el desempeño entre ellos, y así encontrar la mejor combinación posible. En los resultados experimentales muestran las diferencias entre un modelo simple y un modelo combinado.

5. Azmy, W. M., Atiya, A. F., El-Shishiny, H. (2010). Forecast Combination Strategies for Handling Structural Breaks for Time Series Forecasting. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5997 LNCS, 245-253.

Se analizan estrategias de combinación de pronóstico teniendo en cuenta los puntos de quiebre, con el fin de mitigar la discontinuidad de las series de tiempo.

6. Badri, M. A., Al-Mutawa, A., Murtagy, A. (1998). Neural Networks of Data Inhibiting Long Memory Pattern. *Computers and Industrial Engineering*, 35 (3-4), 551-554.

Los autores presentan tres modelos de redes neuronales para pronóstico de series de tiempo que contienen patrones de memoria larga. El objetivo al analizar tres modelos, era obtener un mejor entendimiento sobre el comportamiento de las redes neuronales para el caso en que los datos cuentan con patrones de memoria larga, y así encontrar el modelo óptimo. Éste trabajo es la continuación del desarrollado por el autor en 1997.

7. Badri, M.A. (1997). Neural Networks of Combination of Forecasts for Data with Long Memory Pattern. *IEEE International Conference on Neural Networks - Conference Proceedings*, 1, 359-364.

El autor busca un mejor entendimiento de las redes neuronales para datos que cuentan con patrones de memoria larga. Para esto usa tres modelos de redes neuronales. El primero es el más simple, incluye sólo el pronóstico de la serie de tiempo. El segundo modelo utiliza los mismos datos de entrada del primero y le agrega una combinación de promedios. El tercer modelo combina por medio de una red neuronal los pronósticos individuales.

8. Bouchachia, A. (2009). Radial Basis Function Nets for Time Series Prediction. *International Journal of Computational Intelligence Systems*, 2 (2), 147-57.

Se presenta una metodología novedosa de combinación de pronósticos basada en Redes Recurrentes de Base Radial (RRBFN). La red adaptativa propuesta está basada en un modelo autoregresivo no lineal con variables de entrada exógenas. Con el fin de evaluar los modelos, los autores concluyen sobre el desempeño de las metodologías individuales y las metodologías de ensamble.

9. Cai, J.W., Hu, S. S., Tao, H.F. (2007). Prediction of Chaotic Time Series Based on Selective Support Vector Machine Ensemble. *Wuli Xuebao/Acta Physica Sinica* 56 (12), 6820-7.

Los autores presentan un estudio de predicción de series de tiempo utilizando metodologías de combinación de pronóstico, basadas en Support Vector Regression (SVR). Este análisis se concentra en series de tiempo caóticas.

10. Chen, A.N.S., Leung, M.T. (2005). Performance Evaluation of Neural Network Architectures: The Case of Predicting Foreign Exchange Correlations. *Journal of Forecasting*, 24 (6), 403-420.

En este artículo, los autores comparan el desempeño de dos metodologías de pronóstico basadas en la arquitectura de redes neuronales: multi-layered feedforward (MLFN) y una red neuronal de regresión general. Los resultados demuestran que la totalidad de las características de los datos no son asimiladas por ambos modelos, lo que resalta la importancia de tener modelos que combinen pronósticos, ya que cada pronóstico individual puede especializarse en características particulares.

11. Chen, D.E.W, Zhang, J.P. (2005). Time Series Prediction Based on Ensemble ANFIS. 2005 *International Conference on Machine Learning and Cybernetics, ICMLC*.

En este estudio, un ensamble de modelos ANFIS es propuesto. Los expertos son generados usando bootstrap. El autor concluye que el ensamble tiene mayor precisión que un modelo ANFIS individual cuando se pronostican series de tiempo caóticas y series de tráfico.

12. Chiewchanwattana, S., Lursinsap, C. (2002). FI-GEM Networks for Incomplete Time-Series Prediction. *Proceedings of the International Joint Conference on Neural Networks 2*, 1757-1762.

Este artículo considera el problema de la predicción de series de tiempo incompletas usando redes FI-GEM (Fill-in Generalized Ensemble Method). El método consta de dos pasos: en el primero se llenan los datos faltantes de la serie. En el segundo, se usa un ensamble de Perceptrones Multicapa para generar un único resultado predictivo.

13. Deng, Y.F, Jin, X., Zhong, Y.X. (2005). Ensemble SVR for Prediction of Time Series. *International Conference on Machine Learning and Cybernetics, ICMLC 2005*, 3528-3534.

En este trabajo, se comparan los pronósticos realizados mediante ensambles de Máquinas de Vectores de Soporte generados usando las metodologías de Bagging y Boosting. Para validar los resultados, se utilizan las series de Sunspots y Mackey-Glass.

14. Donaldson, R.G., Kamstra, M. (1996). Forecast combining with neural networks. *Journal of Forecasting 15*, 49-61.

Con este artículo, los investigadores reportan los resultados de la predicción de los índices financieros de USA, Canadá, Japón e Inglaterra. Demuestran que las combinaciones con redes neuronales artificiales generalmente producen pronósticos con los cuales se disminuye el error de la precisión. Esta superioridad se presenta debido a su flexibilidad para modelar las relaciones no lineales potencialmente complejas y difícilmente capturadas por las metodologías lineales tradicionales.

15. Donaldson, R.G., Kamstra, M. (1999). Neural network forecast combining with interaction effects. *Journal of Forecasting 15*, 49-61.

Se demuestra que las redes neuronales artificiales (ANN) pueden capturar interacciones y efectos entre las series de tiempo, entregando así mejores resultados del pronóstico. En este artículo se utilizan series de volatilidad financiera.

16. Dong, J. (2000). Research on the Method of Nonlinear Combining Forecasts Based on Fuzzy-Neural Systems. *Proceedings of the World Congress on Intelligent Control and Automation (WCICA) 2*, 899-903.

En este trabajo se presenta un método de combinación de pronóstico no lineal basado en un sistema difuso de redes neuronales. Se desarrolla un algoritmo backpropagation para ajustar los parámetros del modelo difuso. Con los resultados se muestra que cada modelo se adapta a las características y propiedades de los datos.

17. Dong, J.R. (2002). A Nonlinear Combining Forecast Method Based on Fuzzy Neural Network. *Proceedings of 2002 International Conference on Machine Learning and Cybernetics 4*, 2160-2164.

Se sugiere una red neuronal difusa (FNN) como experto. El desempeño de esta metodología es evaluado con dos metodologías de pronóstico individual y tres metodologías de combinación lineal. Los resultados muestran que la predicción con la metodología (FNN) es mucho mejor que los pronósticos entregados por las metodologías individuales.

18. Fiordaliso, A. (1998). A nonlinear forecast combination method based on Takagi-Sugeno fuzzy systems. *International Journal of Forecasting* 14, 367-379.

Se utiliza un sistema de inferencia difusa de Takagi-Sugeno para los combinar pronósticos individuales. Para validar los resultados, se compara el modelo con un tradicional modelo lineal de combinación de pronóstico. Con la metodología se muestra una ventaja en los pronósticos de series de tiempo no lineales.

19. Fiordaliso, A. (2001). A Constrained Takagi-Sugeno Fuzzy System that Allows for Better Interpretation and Analysis. *Fuzzy Sets and Systems*, 118 (2), 307-18.

En este trabajo se presenta una modificación al modelo generalizado de inferencia difusa de Takagi-Sugeno El objetivo es combinar los patrones de entrada con reglas de restricción. Estas reglas luego son aplicadas al modelo de combinación de pronósticos.

20. Gomes, G.S.S., Maia, A.L.S., Ludermir, T.B., De Carvalho, F.D.A.T., Araujo, A.F.R. (2006). Hybrid Model with Dynamic Architecture for Forecasting Time Series. *IEEE International Conference on Neural Networks - Conference Proceedings*, 3742-3747.

En este estudio, se utiliza la arquitectura dinámica de una red neuronal, propuesta por Ghiassi and Saidane (2005). Se utilizan DAN2 y los resultados son comparados con un modelo ARIMA. Como resultado de la investigación, los autores proponen un modelo híbrido con esta arquitectura dinámica (HAD) basada en combinaciones de pronósticos individuales de DAN2 y ARIMA.

21. Han, D., Niu, W.Q., Yu, C. (2007). The Comparative Study on Linear and Non-Linear Forecast-Combination Methods Based on Neural Network. *International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2007*, 5627-5630.

Este trabajo compara el desarrollo de un modelo de combinación de pronóstico lineal y no lineal, utilizando una metodología de combinación de red neuronal con pesos óptimos. Finalmente los autores encuentran que la metodología entrega mejores resultados en series de tiempo no lineales.

22. Harrald, P.G., M. Kamstra. (1997). Evolving Artificial Neural Networks to Combine Financial Forecasts. *IEEE Transactions on Evolutionary Computation*, 1(1), 40-52.

En este artículo se utiliza la programación evolutiva con el fin de desarrollar redes neuronales que permitan la combinación de pronósticos. Los autores prueban el modelo con la volatilidad del precio de las acciones, y lo testean con metodologías de mínimos cuadrados y el método de Kernel.

23. Inoue, H., Narihisa, H. (2000). Effect of Parallel Ensembles to Self-Generating Neural Networks for Chaotic Time Series Prediction. *Neural Networks for Signal Processing - Proceedings of the IEEE Workshop 2*, 896-905.

En este trabajo, se investiga la capacidad de mejora de la predicción y el paralelo de eficiencia de un ensamble SGNNs (ESGNNs) para tres series de tiempo caóticas. Se asigna una SGNN a cada procesador, los resultados demuestran que mientras más aumente el número de procesadores, se obtiene más exactitud en la predicción.

24. Inoue, H., Narihisa, H. (2005). Parallel Performance of Ensemble Self-Generating Neural Networks for Chaotic Time Series Prediction Problems. *Systems and Computers in Japan*, 36 (10), 82-92.

Con este artículo los investigadores muestran la capacidad de mejorar la precisión de la predicción y la eficiencia de la metodología de ensamble SGNNs, para series de tiempo caóticas. Con el fin de ver el desempeño de cada modelo, se asigna a cada SGNNs un procesador, los resultados muestran que cuando más aumenta el número de procesadores se mejora la precisión de la predicción.

25. Inoue, H., Narihisa, H. (2000). Predicting chaotic time series by ensemble self-generating neural networks. *Proceedings of the International Joint Conference on Neural Networks*, 2, 231-236.

En este artículo se introduce el ensamble de redes neuronales autogenerativas (ESGNNs) para la predicción de series de tiempo caóticas. Las (ESGNNs) son combinadas por una metodología de ensamble con SGNNs. Los resultados son comparados con una red neuronal backpropagation en varias series de tiempo caóticas. Se muestra que usando varios SGNTs en el ensamble, se mejora la precisión del pronóstico de las series caóticas.

26. Islam, M.M., Yao, X., Murase, K. (2003). A constructive algorithm for training cooperative neural network ensembles. *IEEE Transactions on Neural Networks*, 14 (4), 820-834.

Este artículo presenta un algoritmo constructivo (CNNEs), para el ensamble de redes neuronales individuales (NNs). Las CNEEs han sido evaluadas por los autores en problemas benchmark tanto financieros como en problemas de salud (Cancer, Diabetes, entre otros). Se muestra que las CNNE pueden producir ensambles de redes neuronales más generalizados.

27. Jian L., Bangzhu Z., (2007). Neural network ensemble based on feature selection. *IEEE International Conference on Control and Automation, ICCA*. 1844-1847.

Se presenta una metodología de combinación de pronóstico de series de tiempo NNEIPCABag, la cual combina el análisis de los componentes principales (IPCA) y la metodología Bagging. Los resultados muestran que la capacidad de generalización del modelo propuesto puede ser mucho mayor a la de un modelo de redes neuronales.

28. Kim, D., Kim, C. (1997). Forecasting Time Series with Genetic Fuzzy Predictor Ensemble. *IEEE Transactions on Fuzzy Systems*, 5 (4), 523-535.

Este artículo propone un algoritmo genético difuso (GFPE) con dos estados, para la predicción de series de tiempo caóticas. En el primer estado se generan reglas para las predicciones. En el segundo estado se genera una función que minimiza el error. Estos dos estados se desarrollan independientemente de la combinación de los datos de entrada.

29. Kurogi, S., Koyama, R., Tanaka, S., Sanuki, T. (2007). Forecasting using First-Order Difference of Time Series and Bagging of Competitive Associative Nets. *IEEE International Conference on Neural Networks - Conference Proceedings*, 166-171.

Este artículo describe un método de Bagging usado en la competencia de pronósticos de redes neuronales e inteligencia computacional en 2007. Para hacer frente a la estacionalidad de los datos mensuales de la serie de tiempo, la metodología utiliza la diferencia de primer orden. Se muestra como esta metodología reduce la varianza de la predicción de la serie de tiempo analizada.

30. Lai, K. K., Yu, L., Wang, S., Wei, H. (2006). A novel nonlinear neural network ensemble model for financial time series forecasting. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 3991 LNCS - I*, 790-793.

Los autores proponen un nuevo modelo de ensamble con redes neuronales para problemas no lineales, con el fin de ser utilizado en el pronóstico de series de tiempo financieras. Los expertos son redes neuronales. El objetivo es seleccionar apropiadamente los miembros del ensamble. Finalmente la metodología de Máquinas de Soporte Vectorial es usada para ensamblar los resultados de las redes neuronales.

31. Landassuri-Moreno, V.M., Bullinaria, J.A. (2009). Neural network ensembles for time series forecasting. *Paper presented at the Proceedings of the 11th Annual Genetic and Evolutionary Computation Conference, GECCO*, 1235-1242.

Se analiza el uso del algoritmo genético EPNNet para crear ensambles de redes neuronales. Este algoritmo ha sido estudiado ampliamente en problemas de clasificación donde ha entregado muy buenos resultados. Sin embargo no ha pasado lo mismo con problemas de predicción de series de tiempo. Se compara con metodologías de combinación lineal. Es evaluado con 21 tipos diferentes de series de tiempo. Se muestra que es posible mejorar el desempeño del pronóstico con metodologías de combinación.

32. Lemke, C., Gabrys, B. (2010). Meta-Learning for Time Series Forecasting and Forecast Combination. *Neurocomputing*, 73 (10-12), 2006-2016.

Se investiga la aplicabilidad de diferentes aproximaciones de meta aprendizaje, con el fin de conocer qué modelo es mejor en qué situación, y mejorar el desempeño de los pronósticos. Los resultados muestran la superioridad de los métodos de combinación.

33. Leone-Filho, M.A., Ohishi, T., Ballini, R. (2006). Ensembles of selected and evolved predictors using genetic algorithms for time series prediction. *IEEE Congress on Evolutionary Computation, CEC*, 2872-2879.

Se utilizan cuatro tipos de redes neuronales para estimar con más exactitud los parámetros del pronóstico: Perceptron multicapa, Red neuronal recurrente, Red neuronal de base radial y una Red neuro fuzzy. Finalmente se hace uso de un algoritmo genético para encontrar los candidatos del ensamble que optimicen la predicción. El modelo es validado con series de tiempo del Mercado eléctrico. Los resultados muestran la superioridad del modelo cuando se tiene una metodología de selección de expertos.

34. Li X., Deng Z. (2008). Prediction of turning points for chaotic time series using ensemble ANN model. *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*.
- Se presenta un modelo de aprendizaje de máquinas, para predecir series de tiempo caóticas. El modelo de ensamble está basado en un algoritmo de aprendizaje de parámetros EM-like. El modelo es validado con una serie de tiempo benchmark: Mackey-Gass. Los resultados muestran una importante mejora en los pronósticos realizados con el modelo de combinación.
35. Lubecke, T.H., Nam, K.D., Markland, R.E., Kwok, C.C.Y. (1998). Combining foreign exchange rate forecasts using neural networks. *Global Finance Journal*, 9 (1), 5-27.
- Los autores utilizan cuatro metodologías simples de redes neuronales para el pronóstico de: juicios de valor, series econométricas, tasa de interés técnico. Para combinar las metodologías individuales utilizan técnicas de promedio simple. Los resultados de la metodología compuesta son comparados con modelos individuales de pronóstico. Se demuestran los buenos resultados de las metodologías de combinación.
36. Palit, A.K., Popovic, D. (2000). Nonlinear Combination of Forecasts Using Artificial Neural Network, Fuzzy Logic and Neuro-Fuzzy Approaches. *IEEE International Conference on Fuzzy Systems 2*, 566-571.
- Se proponen combinaciones no lineales de series de tiempo, basadas en el uso de redes neuronales, sistemas fuzzy y neuro fuzzy. Con algunos ejemplos prácticos demuestran que la combinación no lineal de pronósticos, captan las mejores características de las metodologías individuales de cada uno de los pronósticos individuales.
37. Peng, S., Zhu, S. (2009). Application of Neural Network Ensemble in Nonlinear Time-Series Forecasts. *2nd International Conference on Intelligent Computing Technology and Automation, ICICTA 2009, 1, art. 5287713*, 45-47.
- Con este trabajo, los autores proponen la combinación de redes neuronales para la predicción de series de tiempo no lineales. En el modelo se aplican los datos del sistema de Lorénz, para la simulación numérica. Los resultados muestran que el modelo de ensamble entrega buenos resultados.
38. Peralta, J., Gutierrez, G., Sanchis, A. (2010). Time Series Forecasting by Evolving Artificial Neural Networks using "Shuffle", Cross-Validation and Ensembles. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 6352 LNCS (PART 1)*, 50-53.
- Este trabajo evalúa dos metodologías para obtener patrones que pueden ser usados en modelos de redes neuronales artificiales. El primero, es llamado por el autor "Shuffle", el segundo es desarrollado con una metodología de Cross-validation. El estudio evalúa en ambos métodos la exactitud del pronóstico.
39. Prudencio, R.B.C., Ludermir, T.B. (2006). Learning Weights for Linear Combination of Forecasting Methods. *Proceedings of the Ninth Brazilian Symposium on Neural Networks, SBRN'06, art. 4026820*.

Se utilizan las técnicas de aprendizaje de máquinas, para definir los pesos adecuados de la combinación lineal de pronósticos. Las técnicas de aprendizaje de máquinas son usadas para definir las características de las series y para adecuar los pesos predefinidos. En la implementación de la solución se usa una red neuronal para combinar los pronósticos individuales. El desempeño experimental revela significativa precisión del pronóstico.

40. Prudencio, R., Ludermir, T. (2004). Using Machine Learning Techniques to Combine Forecasting Methods. *Lecture Notes in Computer Science, Vol 3339, AI 2004: Advances in Artificial Intelligence*, 277-286.

En este trabajo, el autor presenta técnicas de aprendizaje de máquinas para combinar pronósticos de series de tiempo. El propósito es que la metodología capture las características de los datos, y que defina los pesos adecuados de los pronósticos individuales. El modelo de combinación es un promedio de los pesos.

41. Pulido, M., Mancilla, A., Melin, P. (2009). An Ensemble Neural Network Architecture with Fuzzy Response Integration for Complex Time Series Prediction. *Studies in Computational Intelligence* 257, 85-110.

Este artículo describe la aplicación de la arquitectura de un ensamble de redes neuronales para la predicción de series de tiempo. Las series de tiempo consideradas son la Mackey-glass, Dow Jones y la bolsa de valores mexicana. Los resultados se utilizan en un conjunto de entrenamientos del ensamble de la red neuronal, integrados con promedio, promedio ponderado por pesos e integración Fuzzy. Se muestran los buenos resultados de la metodología.

42. Ribeiro, C., Goldschmidt, R., Choren, R. (2010). An Extensible Ensemble Environment for Time Series Forecasting. *ICEIS 2010 - Proceedings of the 12th International Conference on Enterprise Information Systems, 2 AIDSS*, 404-407.

Con este trabajo, el autor presenta un análisis, que puede ser usado para crear, experimentar, y analizar ensambles para el pronóstico de series de tiempo. Pretende entregar una herramienta flexible que incluya el cálculo de la proyección individual, y posteriormente la combinación.

43. Riedel, S., B. Gabrys. (2007). Dynamic Pooling for the Combination of Forecasts Generated using Multi Level Learning. *IEEE International Conference on Neural Networks - Conference Proceedings. art. 4370999*, 454-459.

Este artículo entrega los resultados experimentales y previos hallazgos teóricos respecto a la combinación de pronósticos, diversificados por tres diferentes métodos: Con parámetros de aprendizaje y diferentes niveles de agregación de datos, mediante un modelado robusto y por el uso de diferentes metodologías de pronóstico. Se demuestran los buenos resultados de la metodología cuando se compara el error con el algoritmo Aiolfi y Timmermann.

44. Ruta, D., Gabrys, B. (2007). Neural network ensembles for time series prediction. *IEEE International Conference on Neural Networks - Conference Proceedings, art. 4371129*, 1204-1209.

En este artículo el autor propone un modelo compuesto, utilizando la fortaleza de las redes neuronales múltiples (NN) y la variación de las características de los datos de entrada. El modelo fue evaluado usando los datos de la competencia NISIS 2006 y la competencia NN3 de 2007.

45. Ruta, D., Gabrys, B., Lemke, C. (2011). A generic multilevel architecture for time series prediction. *IEEE Transactions on Knowledge and Data Engineering*, 23(3), 350-359.

Con este trabajo, los autores proponen una arquitectura generalizada para muchos métodos de combinación de series de tiempo. La clave de esta arquitectura, es la habilidad para construir ensambles diversificados de predicciones individuales, que forman muchas variables de entrada.

46. Santana, J. C. (2006). Forecasting Time Series with Neural Networks: An Application to the Colombian Inflation. *Revista Colombiana de Estadística*, 29 (1), 77-92.

Con este análisis los autores evalúan la utilidad de las metodologías de redes neuronales para predecir la inflación en Colombia. Los resultados muestran que las predicciones pueden ser mucho más precisas usando redes neuronales que metodologías como Suavización exponencial y métodos SARIMA. Los resultados muestran que la combinación de pronóstico entrega resultados más precisos que las metodologías de predicción individual.

47. Schwaerzel, R., Rosen, B. (1997). Improving the Accuracy of Financial Time Series Prediction using Ensemble Networks and High Order Statistics. *IEEE International Conference on Neural Networks - Conference Proceedings 4*, 2045-2050.

Se aplica un ensamble de red neuronal para pronosticar series de tiempo financieras, teniendo en cuenta información estadística como parte de los datos de entrada. Los autores muestran cómo el uso de este tipo de metodologías mejora notablemente el resultado de la precisión del pronóstico. Presentan una métrica basada en las simulaciones del precio de los índices financieros que mide la precisión del pronóstico entregado por la metodología planteada.

48. Shi, S. M., Xu, L. D., Liu, B. (1999). Improving the Accuracy of Nonlinear Combined Forecasting using Neural Networks. *Expert Systems with Applications*, 16 (1), 49-54.

Con este trabajo los autores proponen una metodología de combinación de redes neuronales para la predicción de series de tiempo no lineales. Esta metodología es comparada y evaluada con tres metodologías de predicciones individuales y con dos metodologías de combinación lineal. Los resultados revelan que los pronósticos realizados con la metodología de combinación de redes neuronales son más ajustados que los realizados con las metodologías convencionales.

49. Shi, S., Liu, L. (1993). Nonlinear Combination of Forecasts with Neural Networks. *Proceedings of the International Joint Conference on Neural Networks*, 1, 959-962.

Contiene una aproximación a la combinación no lineal de pronóstico basado en redes neuronales. Los resultados sugieren que las combinaciones de pronósticos no lineales con redes neuronales son bastante efectivas para la combinación de pronósticos.

50. Shi, S., Xu, L.D., Liu, B. (1996). Applications of Artificial Neural Networks to the Nonlinear Combination of Forecasts. *Expert Systems*, 13 (3), 195-201.

Con este trabajo, los autores proponen una red neuronal artificial como herramienta para la combinación de pronósticos individuales. Tres modelos de pronósticos individuales son usados para compararlos con dos modelos de combinación de pronóstico. Este artículo sugiere que las metodologías de redes neuronales pueden ser usadas como una alternativa para entregar de una manera más precisa, pronósticos de series de tiempo.

51. Shi-Ying, S. (2009). Study on the Technique and Error Analysis of Nonlinear Combining Forecasts Based Fuzzy System. *Proceedings-2009 International Conference on Electronic Computer Technology, ICECT 2009*, art. 4795989, 385-387.

En este trabajo se presenta una metodología de combinación de pronóstico basada en un sistema de lógica difusa. En general, la idea es que el método de pronóstico para series de tiempo no lineales sea una función $\Phi(x)$ que pueda ser simulada por un modelo Fuzzy Takagi Sugeno. Con este trabajo se encuentra que no solo se reduce la dificultad de construir la función $\Phi(x)$, sino que también se mejora mucho la precisión del pronóstico.

52. Soto, J., Castillo, O., Soria, J. (2010). Chaotic time series prediction using ensembles of ANFIS. *Studies in Computational Intelligence* 318, 287-301.

Este artículo propone una arquitectura ANFIS (Adaptative Network Based Fuzzy Inference System) para ensambles, con el fin de predecir series de tiempo caóticas como Mackey Glass. Los métodos usados para los ensambles de los promedios son: promedio simple, y promedio ponderado con pesos.

53. Sun B., Xie C., Wang G., Zhang J. (2010). The combining prediction of the RMB exchange rate series based on diverse architectural artificial neural network ensemble methodology. *Proceedings 2010 IEEE 5th International Conference on Bio-Inspired Computing: Theories and Applications, BIC-TA 2010*.

Este trabajo entrega diversas arquitecturas de redes neuronales artificiales para optimizar los pesos en metodologías de combinación de pronósticos. Por otro lado, arquitecturas como: Perceptron Multicapa (MLP), Redes Neuronales recurrentes (RNNs) son adoptadas con el fin de diversificar el mecanismo de aprendizaje. La técnica no paramétrica de Suavización de kernel es aplicada para hacer la combinación de pronóstico. Los resultados muestran que la metodología mejora significativamente el pronóstico.

54. Teräsvirta, T., van Dijk, D., Medeiros, M.C. (2005). Linear Models, Smooth Transition Autoregressions, and Neural Networks for Forecasting Macroeconomic Time Series: A Re-Examination. *International Journal of Forecasting*, 21 (4), 755-74.

Se examina la precisión en el pronóstico, de modelos autoregresivos lineales, suavizamiento de transición autorregresivo (STAR), y redes neuronales (NN); con el fin de predecir series de tiempo financieras y económicas. Los resultados de la red neuronal son mezclados con la percepción del horizonte de tiempo del pronóstico.

55. Terui, N., van Dijk, H. K. (2002). Combined Forecasts from Linear and Nonlinear Time Series Models. *International Journal of Forecasting*, 18 (3), 421-438.

Este artículo examina la combinación de pronósticos para modelos lineales y no lineales. Los pronósticos son combinados por un modelo de regresión con coeficientes constantes, así como un método de variables en el tiempo. La metodología de variables en el tiempo focaliza la modelación no lineal. Estas metodologías son aplicadas a tres subconjuntos de series financieras y económicas.

56. Tsangari, H. (2007). An Alternative Methodology for Combining Different Forecasting Models. *Journal of Applied Statistics*, 34 (4), 403-21.

Este trabajo comienza con un análisis comparativo entre un modelo Generalized Autoregressive Conditional Heteroskedasticity (GARCH) y una red neuronal. Los autores continúan con un estudio sobre los principales modelos de combinación de pronóstico. Finalmente proponen una metodología para combinar pronósticos efectivamente.

57. Van Heeswijk, M., Miche, Y., Lindh-Knuutila, T., Hilbers, P.A.J., Honkela, T., Oja, E., Lendasse, A. (2009). Adaptive Ensemble Models of Extreme Learning Machines for Time Series Prediction. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5769 LNCS (PART 2), 305-314.

En este trabajo, los autores investigaron la aplicación de modelos de ensamble adaptativos para Máquinas de Aprendizaje Extremo (ELMs), para el problema de predicción en series de tiempo no estacionarias. Los autores demuestran que el error del modelo adaptado es semejante a los mejores métodos de pronóstico, además demuestran el bajo costo computacional de la metodología.

58. Wang, D., Y. Li. (2010). A Novel Nonlinear RBF Neural Network Ensemble Model for Financial Time Series Forecasting. *3rd International Workshop on Advanced Computational Intelligence, IWACI 2010*, art. 5585218, 86-90.

En este artículo, los autores estudian un modelo de red neuronal no lineal con base radial (RBF-NN), modelo basado en un Support Vector Machine (SVM). En el proceso del ensamble los datos son divididos en dos subconjuntos de entrenamiento, uno usado en una metodología de bagging y otro usado en boosting. Luego los subconjuntos sirven como input de en los diferentes modelos de predicción (RBF-NN). A continuación la técnica (PLS) Partial Least Square es usada para seleccionar los miembros adecuados para el ensamble. Finalmente un (SVM) es usado para ensamblar todos los pronósticos calculados con las metodologías (RBF-NN). Para la validación del modelo, presentan predicciones de series de tiempo financieras.

59. Wichard, J.D. (2011). Forecasting the NN5 time series with hybrid models. *International Journal of Forecasting*, 27 (3), 700-707.

En este trabajo se propone una metodología para predecir series de tiempo con períodos de recurrencia estacional. Los valores faltantes de la serie son estimados e interpolados en un paso de procesamiento dentro del análisis. El modelo es una combinación entre redes neuronales y el modelo de trayectoria más corta. Se evalúa en la competencia NN5.

60. Wichard, J.D., Ogorzałek, M. (2004). Iterated time series prediction with ensemble models. *Proceedings of the IASTED International Conference on Modelling, Identification and Control* 23, art. 412185, 512-517.

Los autores muestran las clásicas metodologías de ensamble por medio de redes neuronales, utilizando diferentes tipos de arquitectura. Se sugiere un procedimiento de predicción iterativo para seleccionar los miembros finales del ensamble. El modelo de validación de este análisis es una extensión de la técnica Cross-Validation.

61. Wichard, J.D., Ogorzałek, M. (2007). Time series prediction with ensemble models applied to the CATS benchmark. *Neurocomputing*, 70 (13-15), 2371-2378.

Los autores muestran las clásicas metodologías de ensamble por medio de redes neuronales, utilizando diferentes tipos de arquitectura. Se sugiere un procedimiento de predicción iterativo para seleccionar los miembros finales del ensamble.

62. Xiao, Z., Gong, K., Zou, Y. (2009). A Combined Forecasting Approach Based on Fuzzy Soft Sets. *Journal of Computational and Applied Mathematics* 228 (1), 326-333.

Con este trabajo los autores proponen un modelo de predicción de series de tiempo basado en una metodología fuzzy. Los resultados son evaluados con datos del comercio Chino desde 1993 hasta 2006. Los resultados muestran que los pronósticos encontrados con metodologías de combinación son más precisos que los pronósticos encontrados con metodologías individuales.

63. Yang, H., J. Shi. (2010). A Three-Stage SVM Ensemble Algorithm for Chaotic Time Series Prediction. *2nd International Workshop on Education Technology and Computer Science, ETCS 2010*.

En este trabajo, los autores entregan un modelo compuesto por tres partes. En la primera parte un algoritmo de cluster parte los datos en subconjuntos. En la segunda parte se perfecciona el algoritmo de predicción con ayuda de un (PSO) Particle Swarm Optimization. Finalmente se utiliza un algoritmo fuzzy para combinar las predicciones.

64. Yang, H., Ma, H. (2010). Chaotic Time Series Prediction Based on Fuzzy Possibility C-Mean and Composite Kernel Support Vector Regression. *ICIME 2010-2010 2nd IEEE International Conference on Information Management and Engineering 2*, art. 5477442, 236-239.

En este trabajo, el autor presenta un modelo basado en Support Vector Machine (SVM) con el fin de obtener un pronóstico combinado de series de tiempo caóticas. Inicialmente se utiliza el algoritmo de cluster (FPCM) con el fin de encontrar los subconjuntos de los datos de entrada. Luego se adaptan los parámetros del modelo por medio de un SVM (Máquinas de Soporte Vectorial). Finalmente un algoritmo Fuzzy es empleado para combinar los pronósticos de las submuestras.

65. Yu, L., Lai, K.K., Wang, S. (2008). Multistage RBF Neural Network Ensemble Learning for Exchange Rates Forecasting. *Neurocomputing*, 71 (16-18), 3295-3302.

En este estudio se utiliza una función radial no lineal (RBF) como base para el modelo de ensamble de predicción de series de tiempo con redes neuronales. El primer paso de este análisis consiste en generar muchos pronósticos con (RBF). El segundo paso consiste en seleccionar los miembros del ensamble por medio de una varianza generalizada condicional (CGV). Para testear el modelo, se compara con algunos ensambles de series de tiempo existentes.

66. Yu, L., Wang, S., Lai, K.K. (2005). A Novel Nonlinear Ensemble Forecasting Model Incorporating GLAR and ANN for Foreign Exchange Rates. *Computers and Operations Research*, 32 (10), 2523-2541.

En este trabajo, los autores proponen un modelo de ensamble de pronóstico no lineal, integrando una metodología (GLAR) con redes neuronales artificiales (ANN). Los resultados son comparados con dos metodologías de pronósticos individuales (GLAR) y (ANN). Se encuentra que la combinación de pronósticos, entrega mejores resultados.

67. Zeng, F., Zhengding, Q. (2003). Predicting Chaotic Time Series by Ensemble Self Generating Neural Networks Merged with Genetic Algorithm. *Proceedings of 2003 International Conference on Neural Networks and Signal Processing, ICNNSP'03*.

En este artículo los autores analizan la correlación que existe entre los componentes del ensamble. En este sentido, el autor propone una metodología basada en algoritmos genéticos para optimizar los componentes del ensamble. Los resultados mostrados en el artículo sobre dos series de tiempo demuestran que la metodología propuesta mejora notablemente la precisión de la predicción de la series de tiempo.

68. Zhang, G.P. (2007). A neural network ensemble method with jittered training data for time series forecasting. *Information Sciences*, 177 (23), 5329-5346.

Los autores presentan un modelo de combinación de pronóstico basado en la validación de los datos de entrada, con el fin de imprimir ruido a los datos y que el modelo pueda tener diferentes pero relacionados conjuntos de entrada.

69. Zhang, G.P., Berardi, V.L. (2001). Time series forecasting with neural network ensembles: An application for exchange rate prediction. *Journal of the Operational Research Society*, 52(6), 652-664.

Los autores presentan una investigación del uso de metodologías de combinación de redes neuronales para pronósticos de series de tiempo, desarrollando el modelo Keep-the-best (KTB). Se proponen dos aproximaciones de redes neuronales con el fin de predecir series de tiempo financieras. Los resultados muestran que los ensambles de redes basados en diferentes subconjuntos de los datos de entrada son más efectivos que los modelos con los cuales se utiliza todo el conjunto de datos.

70. Zhao, Z., Zhang, Y., Liao, H. (2008). Design of Ensemble Neural Network using the Akaike Information Criterion. *Engineering Applications of Artificial Intelligence*, 21 (8), 1182-8.

En este artículo el autor propone un algoritmo de ensamble de redes neuronales basado en Akaike Information Criterion (AIC). El ensamble de redes neuronales basado en AIC investiga la mejor configuración de los pesos de las redes. Dos funciones analíticas: la función "Peak" y la "Friedman" son usadas para encontrar la precisión propuesta del ensamble.

BIBLIOGRAFÍA

- Adya, M., Collopy, F., Armstrong, J., Kennedy, M. (2001). Automatic identification of time series features for rule based forecasting. *International Journal of Forecasting*, 17(2), 143-157.
- Assimakopoulos, V; Nikolopoulos, K. (2000). The Theta Model: A Decomposition Approach to Forecasting. *Journal Forecasting*. 16. (4). 521 – 530.
- Barnard G.A. (1963). New Methods of Quality Control. *Journal of the Royal Statistical Society A*, 126, 255-259.
- Bates, J. M., Granger, C. W. J. (1969). The Combination of Forecasts. *Operational Research Society*. 20 (4). 451-468.
- Bates, J.M., Granger, C.W.J. (1969). The combination of forecasts. *Operations Research Quarterly*, 20, 451-468.
- Bordley, R.F. (1982). The Combination of Forecasts: A Bayesian Approach. *The Journal of the Operational Research Society*. 33 (2). 171- 174.
- Breiman, L. (1996). Stacked regressions. *Machine Learning*, 24, 49-64.
- Bunn, D. (1975). A Bayesian approach to the linear combination of forecasts. *Operational Research Quarterly* 26(2), 325-329.
- Clemen, R.T. (1989). Combining forecasts: a review and annotated bibliography. *International Journal of Forecasting*, 5, 559-583.
- Clements, M.P., Harvey, D.I. (2009). Forecast combination and encompassing. In T. C. Mills, & K. Patterson (Eds.), *Palgrave handbook of econometrics, volume 2: Applied econometrics*. 169–198. Basingstoke: Palgrave MacMillan.
- Clements, M.P., Harvey, D.I. (2011). Combining probability forecasts. *International Journal of Forecasting*, 27 (2), 208-223.
- De Menezes, L., Bunn, D., Taylor, J. (2000). Review of guidelines for the use of combined forecasts. *European Journal of Operational Research*, 120, 190-204.
- Diebold, F. X., López, J.A. (1996). Forecast evaluation and combination. In *Handbook of statistics: Vol. 14 (1996)* 241–268. Amsterdam: North-Holland.
- Granger, C.W.J. (1989). Combining forecasts—twenty years later. *Journal of Forecasting*, 8, 167-173.

- Hall, S.G., Mitchell, J. (2007). Combining density forecasts. *International Journal of Forecasting*, 23 (1), 1-13.
- Hansen, B.E. (2006). Interval forecasts and parameter uncertainty. *Journal of Econometrics*, 135 (1-2), 377-398.
- Hansen, K.-L., Salamon, P. (1990). Neural Network Ensembles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol 12 N°10.
- Hansen, S.; Schmeiser, B. (1995). Improving Model Accuracy Using Optimal Lineal Combinations of Trained Neural Networks. *IEEE Transactions on Neural Networks*. Vol 6. N°3.
- Hashem, S. (1997). Optimal linear combinations of neural networks. *Neural Networks*, 10, 599-614.
- Hashem, S., Schmeiser, B. (1993). Approximating a function and its derivatives using MSE-optimal linear combinations of trained feedforward neural networks. In *Proceedings of the 1993 World Congress on Neural Networks*, 1, 617-620. New Jersey: Lawrence Erlbaum Associates.
- Hashem, S., Schmeiser, B. (1995). Improving model accuracy using optimal linear combinations of trained neural networks. *IEEE Transactions on Neural Networks*, 6 (3), 792-794.
- Hashem, Sherif. (1997). Optimal Lineal Combinations of Neural Networks. *Neural Networks*. Vol 10 N°4. 599-614.
- Haykin, S. (1996). *Neural Networks: A Comprehensive Study*. Prentice Hall International.
- Hendry, D., Clements, M. (2004). Pooling of forecasts. *Econometrics Journal*, 7, 1-31.
- Hu, M. Y., Tsoukalas, C. (1999). Combining Conditional Volatility Forecasts using Neural Networks: An Application to the EMS Exchange Rates. *Journal of International Financial Markets, Institutions and Money*, 9 (4), 407-422.
- Jacobs, R.A., Jordan, M. (1991). A competitive modular connectionist architecture. In R. Lippman, J. Moody, & D. Touretzky (Eds.), *Advances in Neural Information Processing System*, 3 767-773. California: Morgan Kaufman.
- Jacobs, R.A., Jordan, M. (1994). Hierarchical mixtures of experts and the EM algorithm. *Neural Computation*, 6, 181-214.
- Jacobs, R.A., Jordan, M.I., Nowlan, S.J., Hinton, G.E. (1991). Adaptive mixtures of local experts. *Neural Computation*, 3, 79-87.
- Jose, V.R.R., Winkler, R.L. (2008). Simple robust averages of forecasts: some empirical results. *International Journal of Forecasting* 24 (1), 163-169.
- Kitchenham, B.A. (2004). Procedures for Undertaking Systematic Reviews. *Joint Technical Report, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd. (0400011T.1)*.
- Kitchenham, B.A., Charters, S. (2007). Guidelines for Performing Systematic Literature Reviews in Software Engineering. *Technical Report EBSE-2007-01*.
- Kitchenham, B.A., Dybå, T., Jørgensen, M. (2004). Evidence-based software engineering. *Proceedings of the 26th International Conference on Software Engineering, (ICSE'04), IEEE Computer Society, Washington DC, USA*, 273-281.

- Kitchenham, B.A., Pfleeger, S.L., Pickard, L.M., Jones, P.W., Hoaglin, D.C., El Emam, K., Rosenberg, J. (2002) Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering*, 28 (8), 721-734.
- Li, X., Z. Deng. (2008). Prediction of Turning Points for Chaotic Time Series using Ensemble ANN Model. *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*, art. 4593474, 3459-3464.
- Makridakis, S. (1989). Why combining works?. *International Journal of Forecasting* 5, 601 - 603.
- Makridakis, S., Winkler, R.L. (1983) Averages of Forecasts: Some Empirical Results. *Inform Management Science*, 29 (9). 987-996.
- Naftaly, U., Intrator, N., Horn, D. (1997). Optimal ensemble averaging of neural networks. *Network: Computation in Neural Systems*, 8 (3), 283-296
- Newbold, P., Granger, C.W.J. (1974). Experience with forecasting univariate time series and the combination of forecasts (with discussion). *Journal of the Royal Statistical Society, Series A*, 137, 131–164.
- Newbold, P., Harvey, D. I. (2002). Forecasting combination and encompassing. In M. Clements, & D. Hendry (Eds.), *A companion to economic forecasting*, 268–283. Oxford: Blackwells.
- Opitz, D., Maclin, R. (1999). Popular Ensemble Methods: an Empirical Study. *Journal of Artificial Intelligence Research* 11. 169 – 198.
- Perrone, M.P. (1993). Improving regression estimation: Averaging methods for variance reduction with extensions, to general convex measure optimization. PhD. Thesis. Brown University, Rhode Island.
- Perrone M.P. y Cooper, N. (1992). When networks disagree: Ensemble methods for hybrid neural networks. In *Neural Networks for Speech and Image processing*. Chapman & Hall
- Reid, D.J. (1968). Combining three estimates of gross domestic products. *Economica*, 35, 431–444.
- Schapire, R. (1990). The strength of weak learnability. *Machine Learning*, 5, 197–227.
- Sharkey, A.J.C. (2010). On Combining Artificial Neural Nets. Department of Computer Science, University of Sheffield, Sheffield, UK. Online publication. <http://www.informaworld.com/smpp/title~content=t713411269>.
- Sun, B., Xie, C., Wang, G., Zhang, J. (2010). The Combining Prediction of the RMB Exchange Rate Series Based on Diverse Architectural Artificial Neural Network Ensemble Methodology. *Proceedings 2010 IEEE 5th International Conference on Bio-Inspired Computing: Theories and Applications, BIC-TA 2010*, art. 5645218, 743-749.
- Tay, A. S.; Wallis, K. F. (2000). Density forecasting: A survey. *Journal of Forecasting*, 19, 235–254.
- Taylor, J.W. (2003). Exponential Smoothing with a Damped Multiplicative Trend. *Journal Forecasting* 19 (4), 715 – 725.
- Timmerman, A. (2006). Forecast combinations. In G. Elliot, C. W. J. Granger, & A. Timmerman (Eds.), *Handbook of economic forecasting* 135–196. Amsterdam: North-Holland.
- Winkler, R.L., Makridakis, S. (1983). The combination of forecasts. *Journal of the Royal Statistical Society, Series A*, 146 (2), 150-157.

Wolpert, D.H. (1992). Stacked Generalization. Complex System Group, Theoretical Division, and Center for Non-Linear Studies, MS B213.

Zhang, G; Patuwo, B; Hu, M. (1998). Forecasting with Artificial Neural Networks: The State of the Art. *Journal Forecasting*, 14 (1). 35-62.