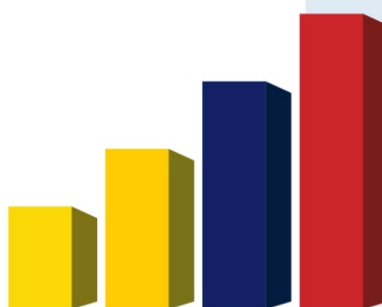


Analítica

Modelación de series económicas mediante métodos
automáticos de regresión difusa

Modeling of economic time series using automated fuzzy
regression methods

Rodrigo Cajamarca - Hermann Mena



www.inec.gob.ec | www.ecuadorencifras.com

Modelación de series económicas mediante métodos automáticos de regresión difusa

Modeling of economic time series using automated fuzzy regression methods

Rodrigo Cajamarca[†] y Hermann Mena[‡]

[†]Ministerio de Educación, Quito-Ecuador

[‡]Departamento de Matemática, Escuela Politécnica Nacional, Quito - Ecuador

[†]rpaulcajamarca@hotmail.com, [‡]hermann.mena@epn.edu.ec

Recibido: 20 de octubre de 2011

Aceptado: 8 de diciembre de 2011

Resumen

Este trabajo se desarrolla en un contexto donde Estadística e Inteligencia Computacional convergen: la Regresión Difusa. La idea fundamental de esta técnica es generalizar conceptos de regresión tradicional a conjuntos difusos. Concretamente, se investigará el potencial de aplicar los métodos automáticos de regresión difusa a cierto tipo de series económicas. En particular, se estudian los siguientes métodos: mínimos cuadrados por lotes, mínimos cuadrados recursivo, aprendizaje desde el ejemplo modificado y agrupamiento difuso combinado. Adicionalmente, se propone el método de mínimos cuadrados recursivo combinado, el cual es una de las principales contribuciones de este trabajo. Cada uno de estos métodos ha sido descrito e implementado en R para el caso unidimensional y se generaliza para el caso de entradas y salidas múltiples. Finalmente, se muestran resultados numéricos de la Reserva Internacional de Libre Disponibilidad y un índice de liquidez financiera, en las cuales se visualiza el comportamiento y desempeño de los métodos comparándolos con modelos SARIMA.

Palabras clave: Conjuntos difusos, funciones de pertenencia, lógica difusa, métodos automáticos de regresión difusa.

Abstract

This document works in a context where Statistics and Computational Intelligence converge: Fuzzy Regression. The idea of this technique is to generalize concepts from statistic regression into fuzzy sets; more precisely, to research the possibility of applying automated fuzzy regression methods for modeling economic data. It discusses the Batch Least Squares, Recursive Least Squares, Learning by Example and Combined Clustering methods. The main theoretical contribution of this research is the proposal of a novel method: the Combined Recursive Least Squares. All the methods are described and implemented in R for the one-dimensional case and generalized for the multiple input/output case. To conclude, numerical tests from the Free Liquidity International Reserve (RILD) and a financial index show the good performance of these methods compared to SARIMA models.

Keywords: Fuzzy sets, Membership Functions, Fuzzy Logic, Automated Fuzzy Regression.

Código JEL: C02, C22, C32.

1 Introducción

Actualmente, existen muchos métodos y técnicas diseñadas para modelar fenómenos económicos. Lamentablemente, muchos de estos métodos no han dado buenos resultados debido, sobre todo, a la presencia de incertidumbre en los datos; este es el caso de varias técnicas tradicionales, *e.g.*, regresión clásica. En general, la falta de conocimiento del sistema, confianza en la información, o la presencia de datos incompletos/imputados constituyen los obstáculos más grandes que los métodos de modelización deben enfrentar. Afortunadamente, en situaciones como esta la modelación difusa constituye una opción real y competitiva. [9, 15].

La idea fundamental de la regresión difusa es generalizar conceptos de la regresión tradicional a datos que se pueden modelar con conjuntos difusos. Estos conjuntos, en general, describen de mejor manera los fenómenos inmersos en el sistema. La regresión probabilística solamente puede modelar el fenómeno de la incertidumbre mediante la inclusión de intervalos de confianza. La aleatoriedad puede considerarse como uno de los componentes de la vaguedad, junto a la imprecisión. Otras formas de incertidumbre se pueden encontrar en la ambigüedad, la incongruencia, problemas de especificación y factores no considerados por la alta complejidad del sistema. En algunos casos, es posible incorporar estas restricciones a ciertos modelos de regresión probabilística. Sin embargo, la inclusión de estas restricciones implica que algunos supuestos requeridos por la teoría ya no se satisfagan, *e.g.*, que el valor esperado de los residuos sea cero [6]. Esto no ocurre en la regresión difusa, método en el que la presencia de un número mayor de restricciones no constituye un impedimento para su aplicación, puesto que las desviaciones entre los valores de pertenencia observados y los valores de pertenencia estimados se asumen dependientes de la incertidumbre.

Este trabajo está orientado a revisar y extender algunas propuestas que se han formulado en la literatura sobre la regresión difusa, con el objetivo de proporcionar un instrumento metodológico a problemas reales en los cuales los datos disponibles se ven afectados por factores como la imprecisión e incertidumbre. En particular, nos enfocamos en los métodos automáticos de regresión difusa. En la Sección 2 se revisan algunas definiciones y los métodos de mínimos cuadrados por lotes, mínimos cuadrados recursivo, aprendizaje desde el ejemplo modificado, agrupamiento difuso combinado; además, se propone el método de mínimos cuadrados recursivo combinado. Cada uno de estos métodos ha sido descrito e implementado para el caso unidimensional, y generalizado para el caso de entradas y salidas múltiples. En la Sección 3, se muestran resultados numéricos de la Reserva Internacional de Libre Disponibilidad (RILD) y un índice de liquidez financiera, en los cuales se visualiza el comportamiento y desempeño de los métodos. Finalmente, en la Sección 4 se presentan algunas

consideraciones sobre lo analizado y planteamientos para trabajos futuros.

2 Metodología

2.1 Principales definiciones

La teoría de conjuntos difusos fue propuesta originalmente por Lotfi Zadeh [18] para describir matemáticamente la imprecisión y la vaguedad que se presentan en el lenguaje diario. Formalmente, un conjunto difuso \tilde{A} en el conjunto universo X , está caracterizado por una *función de pertenencia* $\mu_{\tilde{A}}$ que toma valores en el intervalo $[0, 1]$, [18]. En este sentido, un conjunto difuso es una generalización de un conjunto clásico pues la función de pertenencia permite que tome valores en el intervalo $[0, 1]$ en lugar de solo $\{0, 1\}$ (no pertenece o pertenece, respectivamente). Concretamente, el conjunto \tilde{A} se representa como el conjunto de pares,

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\}$$

donde $\mu_{\tilde{A}} : X \rightarrow [0, 1]$ es la función de pertenencia para \tilde{A} ; para más detalles ver [5].

Sea $g : \tilde{X} \rightarrow \tilde{Y}$; $\tilde{X} \subset \mathbb{R}^n$ y $\tilde{Y} \subset \mathbb{R}$, la función que describe el sistema en estudio; se quiere construir un sistema difuso representado por la función $f : X \rightarrow Y$; $X \subset \tilde{X}$ y $Y \subset \tilde{Y}$, mediante la elección de un vector de parámetros θ . Se espera aproximar la función g de la forma:

$$g(x) = f(x \mid \theta) + e(x), \quad (1)$$

para todo $x = [x_1, x_2, \dots, x_n]^T \in X$; donde $e(x)$ representa el error aproximación. La i -ésima entrada-salida de la función g se nota como (x^i, y^i) donde $x^i \in X$, $y^i \in Y$ con $y^i = g(x^i)$, para $i = 1, \dots, M$. Luego $x^i = [x_1^i, x_2^i, \dots, x_n^i]^T$ corresponde el vector de entradas para el i -ésimo par de datos. El conjunto de parejas ordenadas de entrada-salida del sistema se conoce como *conjunto de datos de entrenamiento* y se nota

$$G = \{(x^1, y^1), \dots, (x^M, y^M)\} \subset X \times Y \quad (2)$$

donde M es la cardinalidad de G .

Los métodos automáticos presentados en este trabajo generan reglas base, *i.e.*, reglas que describen la relación que existe entre las variables de entrada y salida, o en su defecto, usan una regla base predeterminada para modelar el sistema. En cualquier caso, las reglas están compuestas por antecedentes y consecuentes, *i.e.*,

SI *antecedente* ENTONCES *consecuente*,

las reglas permiten predecir y/o gobernar la salida del sistema con conocimiento a priori de sus entradas [12].

En los métodos que se revisan a continuación, tanto las variables de entrada como las de salida se deben describir con funciones de pertenencia específicas. En la mayoría de métodos se incorporan las funciones de pertenencia Gaussianas para las variables de entrada, pues describen una amplia gama de fenómenos y la función de pertenencia Delta para las variables de salida. Cabe señalar que los métodos pueden ser modificados para cualquier tipo de función de pertenencia de entrada o salida.

2.2 Mínimos cuadrados por lotes (MCL)

Este método aplica el estimador de mínimos cuadrados ordinarios a conjuntos difusos. Para ello, es necesario caracterizar el comportamiento del sistema mediante una regla-base (centros y dispersiones de las funciones de pertenencia para las variables de entrada y salida). En los casos en los cuales se desconoce el comportamiento del sistema, es posible utilizar otros algoritmos con la capacidad de formar reglas, como los métodos de aprendizaje desde el ejemplo; ver [15].

Para aplicar los mínimos cuadrados, se utiliza el sistema

$$f(x^i | \theta) = \frac{\sum_{l=1}^R b_l \mu_{il}}{\sum_{l=1}^R \mu_{il}} \quad i = 1, \dots, M \quad (3)$$

donde,

$$\mu_{il} = \prod_{j=1}^n \exp \left[-\frac{1}{2} \left(\frac{x_j^i - c_j^l}{\sigma_j^l} \right)^2 \right] \quad l = 1, \dots, R \quad (4)$$

R representa el número de reglas en la regla-base y n es la dimensión del vector x . Es importante señalar que el parámetro R se desconoce para algunos métodos. En ciertos casos, este parámetro es determinado por el algoritmo directamente *e.g.*, aprendizaje desde el ejemplo modificado. Los b_l son los centros de las funciones de pertenencia de salida y los c_j^l, σ_j^l son los centros y dispersiones de las funciones de pertenencia de entrada, respectivamente. Hay que señalar que la dispersión relativa $\sigma_j^l > 0$ para todo l y j . Los centros y dispersiones de las funciones de pertenencia de entrada se pueden inicializar utilizando un proceso heurístico, tal que,

$$\sigma_j^l = \sigma > 0 \quad l = 1, \dots, R, j = 1, \dots, n \quad (5)$$

Luego por (3)

$$f(x^i | \theta) = \frac{b_1 \mu_{i1}}{\sum_{l=1}^R \mu_{il}} + \frac{b_2 \mu_{i2}}{\sum_{l=1}^R \mu_{il}} + \dots + \frac{b_R \mu_{iR}}{\sum_{l=1}^R \mu_{il}} \quad i = 1, \dots, M$$

si se nota:

$$\tilde{\zeta}_{il} = \frac{\mu_{il}}{\sum_{l=1}^R \mu_{il}} \quad l = 1, \dots, R,$$

entonces,

$$f(x^i | \theta) = b_1 \tilde{\zeta}_{i1} + b_2 \tilde{\zeta}_{i2} + \dots + b_R \tilde{\zeta}_{iR}$$

es decir,

$$y^i = f(x^i | \theta) = \theta^\top \tilde{\zeta}^i \quad (6)$$

donde:

$$\theta = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_R \end{bmatrix} \quad \text{y} \quad \tilde{\zeta}^i = \begin{bmatrix} \tilde{\zeta}_{i1} \\ \tilde{\zeta}_{i2} \\ \vdots \\ \tilde{\zeta}_{iR} \end{bmatrix} \quad i = 1, \dots, M.$$

Se puede apreciar que (6) es similar al modelo de regresión lineal [6]. La diferencia radica en que, en lugar de la matriz X , se consideran los números difusos de la matriz $\tilde{\zeta}$, de dimensión $R \times M$, es decir, son los lotes o particiones formadas por las funciones de pertenencia de las variables de entrada [13].

Finalmente, si se nota $\Phi = \tilde{\zeta}$ y se aplica el método de mínimos cuadrados el estimador está dado por

$$\hat{\theta} = (\Phi^\top \Phi)^{-1} \Phi^\top y \quad (7)$$

El método de mínimos cuadrados por lotes ha probado ser muy eficiente para una gran variedad de aplicaciones. Sin embargo, la matriz $\Phi^\top \Phi$ podría ser mal condicionada. Una generalización de este método para el caso de entrada-salida múltiple se presenta en el Algoritmo 1.

Algoritmo 1 Mínimos cuadrados por lotes entrada-salida múltiple

Entrada: $X_{M \times n}$ matriz de las variables de entrada
 $Y_{M \times H}$ matriz de las variables de salida
 $\sigma > 0$

- 1: $R \leftarrow M - 1$
 - 2: $c_j^l \leftarrow x_j^l + \frac{1}{2}(x_j^{l+1} - x_j^l) \quad l = 1, \dots, R, j = 1, \dots, n$
 - 3: $\sigma_j^l \leftarrow \sigma \quad l = 1, \dots, R, j = 1, \dots, n$
 - 4: **para** $i \leftarrow 1, M$ **hacer**
 - 5: **para** $l \leftarrow 1, R$ **hacer**
 - 6: $\mu_{il} \leftarrow \prod_{j=1}^n \exp \left[-\frac{1}{2} \left(\frac{x_j^i - c_j^l}{\sigma_j^l} \right)^2 \right]$
 - 7: **fin para**
 - 8: $\tilde{\zeta}_{il} \leftarrow \frac{\mu_{il}}{\sum_{l=1}^R \mu_{il}}$
 - 9: **fin para**
 - 10: $\Phi \leftarrow \tilde{\zeta}(x)$
 - 11: **para** $h \leftarrow 1, H$ **hacer**
 - 12: $\hat{\theta}_h \leftarrow (\Phi^\top \Phi)^{-1} \Phi^\top y_h$
 - 13: **devolver** $f_h(x | \hat{\theta}_h) \leftarrow \hat{\theta}_h^\top \tilde{\zeta}(x)$
 - 14: **fin para**
-

2.3 Mínimos cuadrados recursivo (MCR)

El método de mínimos cuadrados recursivo es, como su nombre lo indica, una versión recursiva del método de mínimos cuadrados por lotes. Este método utiliza en cada iteración una parte del conjunto de entrenamiento y no requiere el cálculo de la inversa de $\Phi^T \Phi$; de esta manera su desempeño es más eficiente que el MCL. El algoritmo calcula $\hat{\theta}(k)$ mediante $\hat{\theta}(k-1)$, x^k , y y^k ; donde k representa un ciclo sobre el conjunto de entrenamiento G , la iteración se repite K veces; para más detalles ver [11].

Al igual que para el MCL, se utiliza la función de pertenencia Gaussiana para las entradas y una función Delta para las salidas. Los parámetros b_l , c_j^l y $\sigma_j^l > 0$ se definen como en el método anterior.

Sea $P(k)$ la matriz $n \times n$ definida como

$$P(k) = (\Phi^T \Phi)^{-1} = \left(\sum_{i=1}^k x^i (x^i)^T \right)^{-1}, \quad (8)$$

se obtiene

$$P(k) = P(k-1) - P(k-1)x^k (I + (x^k)^T P(k-1)x^k)^{-1} (x^k)^T P(k-1), \quad (9)$$

y

$$\hat{\theta}(k) = \hat{\theta}(k-1) + P(k)x^k (y^k - (x^k)^T \hat{\theta}(k-1)). \quad (10)$$

Si se aplica el lema de la matriz inversa, se evita la inversión de una matriz; y, en su lugar, únicamente se calcula el inverso del escalar $(I + (x^k)^T P(k-1)x^k)^{-1}$. Las ecuaciones (9) y (10) caracterizan $P(k)$ y $\hat{\theta}(k)$ y permiten describir el método de mínimos cuadrados recursivo.

Para inicializar el algoritmo, se emplean generalmente $\hat{\theta}(0) = 0$ y $P(0) = P_0 = \alpha I$ para algún $\alpha > 0$. Otra opción es utilizar una aproximación heurística.

Además se puede usar el método de mínimos cuadrados recursivo ponderado. En este caso se utiliza un factor $0 < \lambda \leq 1$ que se conoce como un "factor de pérdida de memoria" y sirve para dar mayor peso a la información más reciente. Si $\lambda = 1$ se obtiene el método MCR original. En el método de mínimos cuadrados recursivo $P(k)$ se define como:

$$P(k) = \frac{1}{\lambda} (I - P(k-1)x^k (\lambda I + (x^k)^T P(k-1)x^k)^{-1} (x^k)^T) P(k-1) \quad (11)$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + P(k)x^k (y^k - (x^k)^T \hat{\theta}(k-1))$$

Si en el sistema difuso (3), se reemplaza x^k con $\xi(x^k)$ en (11) se obtiene

$$P(k) = \frac{1}{\lambda} \left\{ I - P(k-1)\xi(x^k) \left[\lambda I + (\xi(x^k))^T P(k-1)\xi(x^k) \right]^{-1} (\xi(x^k))^T \right\} P(k-1)$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + P(k)\xi(x^k) \left[y^k - (\xi(x^k))^T \hat{\theta}(k-1) \right]$$

Al igual que en la sección anterior, se generalizó este método para el caso de entrada-salida múltiple; ver Algoritmo 2.

2.4 Aprendizaje desde el ejemplo modificado (AEM)

Este método tiene un enfoque muy intuitivo para la construcción de un sistema difuso el cual se basa en técnicas de optimización para determinar los parámetros del sistema. En el AEM se generan automáticamente reglas y funciones de pertenencia; por esta razón, este método es una alternativa viable a problemas en los cuales se carece de información *a priori* del sistema. De manera análoga a la sección anterior, se consideran funciones de pertenencia Gaussiana y Delta; el método se describe de la siguiente manera:

1. Construcción de un sistema difuso inicial: dados el par de datos de entrada y salida $(x^i, y^i) \in G, i = 1, \dots, M$, se forma un sistema difuso inicial con (x^1, y^1) y se inicializa el número de reglas $R = 1$

$$f(x^i | \theta) = b_1 \quad (12)$$

para $c_j^1 = x_j^1, \sigma_j^1 = \sigma_0$ (parámetros de las funciones de pertenencia de las variables de entrada) y $b_1 = y^1$ (parámetro de la función de pertenencia de la variable de salida), para todo $j = 1, \dots, n$.

2. Evaluación de los datos de entrenamiento: se utiliza un factor de tolerancia ε_f que determina el error. Para cada $(x^i, y^i) \in G$ se evalúa:

- si $|f(x^i | \theta) - y^i| \leq \varepsilon_f$, entonces el sistema difuso f representa correctamente (x^i, y^i) , y por lo tanto, no es necesario añadir ninguna regla. Se repite este paso, con el siguiente par de datos de entrenamiento.
- si $|f(x^i | \theta) - y^i| > \varepsilon_f$, entonces se añade una regla para representar (x^i, y^i) modificando los parámetros de las funciones de pertenencia; continuar al Paso 3

3. Codificación de parámetros del sistema difuso: si no se cumple la condición del Paso 2, se añade una nueva regla $R = R + 1, b_R = y^i$ y $c_j^i = x_j^i$, para todo $j = 1, \dots, n$. La modificación de los σ_j^i para $i = R$ se realiza utilizando los centros de las funciones de pertenencia de la forma;

$$n_j^* = \arg \min \{ |c_j^{i'} - c_j^i| : i' = 1, 2, \dots, R, i' \neq i \},$$

Algoritmo 2 Mínimos cuadrados recursivo entrada-salida múltiple

Entrada: $X_{M \times n}$ matriz de las variables de entrada

$Y_{M \times H}$ matriz de las variables de salida

K número de iteraciones del algoritmo

$\sigma > 0, \alpha > 0, 0 < \lambda \leq 1$

- 1: $\hat{\theta}(0) \leftarrow 0$
 - 2: $P(0) \leftarrow \alpha I$
 - 3: $c_j^l \leftarrow x_j^l + \frac{1}{2}(x_j^{l+1} - x_j^l) \quad l = 1, \dots, R, j = 1, \dots, n$
 - 4: $\sigma_j^l \leftarrow \sigma \quad l = 1, \dots, R, j = 1, \dots, n$
 - 5: **para** $i \leftarrow 1, M$ **hacer**
 - 6: **para** $l \leftarrow 1, R$ **hacer**
 - 7: $\mu_{il} \leftarrow \prod_{j=1}^n \exp \left[-\frac{1}{2} \left(\frac{x_j^i - c_j^l}{\sigma_j^l} \right)^2 \right]$
 - 8: **fin para**
 - 9: $\xi_{il} \leftarrow \frac{\mu_{il}}{\sum_{l=1}^R \mu_{il}}$
 - 10: **fin para**
 - 11: **para** $h \leftarrow 1, H$ **hacer**
 - 12: $s \leftarrow 0$ contador auxiliar
 - 13: **para** $k \leftarrow 1, K \times M$ **hacer**
 - 14: **si** $s \leq M$ **entonces**
 - 15: $s \leftarrow s + 1$
 - 16: **si no**
 - 17: $s \leftarrow 1$
 - 18: **fin si**
 - 19: $P_h(k) \leftarrow \frac{1}{\lambda} \left\{ I - P_h(k-1) \xi(x^s) \left[\lambda I + (\xi(x^s))^\top P_h(k-1) \xi(x^s) \right]^{-1} (\xi(x^s))^\top \right\} P_h(k-1)$
 - 20: $\hat{\theta}_h(k) \leftarrow \hat{\theta}_h(k-1) + P_h(k) \xi(x^s) \left[y_h^s - (\xi(x^s))^\top \hat{\theta}_h(k-1) \right]$
 - 21: **fin para**
 - 22: **devolver** $f_h(x | \hat{\theta}_h) \leftarrow \hat{\theta}_h^\top \xi(x)$
 - 23: **fin para**
-

donde $j = 1, 2, \dots, n$ y c_j^i es fijo. Finalmente se actualiza los σ_j^i para $i = R$, i.e.,

$$\sigma_j^i = \frac{1}{\omega} |c_j^i - c_j^{n^*}| \quad j = 1, 2, \dots, n, \quad (13)$$

donde ω es un factor que determina la sobreposición de las funciones de pertenencia. Se puede observar en (13) que el factor de ponderación ω y la dispersión σ_j^i tienen una relación inversa, i.e., para un ω más grande hay una menor sobreposición. Una vez que el Paso 3 se ha completado, se repite el Paso 2 hasta que el número de datos de entrenamiento $(x^i, y^i) \in G$ se haya terminado.

Dado que el parámetro ε_f caracteriza la exactitud con la que opera el sistema difuso, la elección de ε_f determina el número de reglas. Una de las partes más importantes en la construcción de sistemas difusos es la elección del número de reglas difusas R , para lo cual es necesario el criterio de expertos [8]. Algunas aplicaciones particulares limitan el número de reglas del sistema, ya sea por las características específicas del fenómeno o por su costo computacional.

En general, existe una compensación entre la complejidad computacional y la capacidad funcional basada en el número de reglas difusas R . La generalización para el caso de entrada-salida múltiple de este método se presenta en el Algoritmo 3.

2.5 Agrupamiento difuso combinado (ADC)

Este método se compone de dos etapas. En la primera, se identifican las reglas del sistema mediante el algoritmo c-medias difuso, el cual particiona los datos de entrada en conjuntos difusos datos similares [4]. En la segunda etapa, se aplica mínimos cuadrados para caracterizar el conjunto de salida.

El agrupamiento difuso c-medias permite que los elementos de los grupos tengan diferentes grados de pertenencia. Estos grados de pertenencia μ_{il} y los centros de clase ϑ^l se determinan de manera iterativa minimizando la siguiente función objetivo,

$$J = \sum_{i=1}^M \sum_{l=1}^R (\mu_{il})^m |x^i - \vartheta^l|^2, \quad (14)$$

Algoritmo 3 Aprendizaje desde el ejemplo modificado entrada-salida múltiple

Entrada: $X_{M \times n}$ matriz de las variables de entrada
 $Y_{M \times H}$ matriz de las variables de salida
 $\epsilon_f > 0, \omega > 0$

```

1:  $R \leftarrow 1$ 
2:  $b_1 \leftarrow y^1$ 
3:  $\hat{\theta} \leftarrow b$ 
4:  $c_j^1 \leftarrow x_j^1 \quad j = 1, \dots, n$ 
5: para  $h \leftarrow 1, H$  hacer
6:   para  $i \leftarrow 1, M$  hacer
7:     para  $l \leftarrow 1, R$  hacer
8:        $\mu_{il} \leftarrow \prod_{j=1}^n \exp \left[ -\frac{1}{2} \left( \frac{x_j^i - c_j^l}{\sigma_j^l} \right)^2 \right]$ 
9:     fin para
10:     $\zeta_{il} \leftarrow \frac{\mu_{il}}{\sum_{l=1}^R \mu_{il}}$ 
11:     $f^l(x) \leftarrow \hat{\theta}^\top \zeta(x)$ 
12:    si  $|f_h^l(x) - y_h^l| > \epsilon_f$  entonces
13:       $R \leftarrow R + 1$ 
14:       $b_R \leftarrow y_h^l$ 
15:       $\hat{\theta}_h \leftarrow b$ 
16:       $c_j^R \leftarrow x_j^l \quad j = 1, \dots, n$ 
17:       $n_j^* \leftarrow \arg \min \{ |c_j^{l'} - c_j^l| : l' = 1, 2, \dots, R, l' \neq l \}$ 
18:       $\sigma_j^l \leftarrow \frac{1}{\omega} |c_j^l - c_j^{n_j^*}|$ 
19:    fin si
20:  fin para
21:  devolver  $f_h(x | \hat{\theta}_h) \leftarrow \hat{\theta}_h^\top \zeta_h(x)$ 
22: fin para

```

donde $m > 1$ es un parámetro que sirve para determinar la sobreposición entre las clases, M es el número de pares de entrada-salida (x^i, y^i) en el conjunto de entrenamiento G , R es el número de reglas que se quiere calcular (clases o grupos), $\underline{v}^l = [v_1^l, v_2^l, \dots, v_n^l]^\top$ para $l = 1, \dots, R$ es el vector de centros de clases y μ_{il} para $i = 1, \dots, M$ y $l = 1, \dots, R$ es el grado de pertenencia de x^i en el l -ésimo grupo, $|x| = \sqrt{x^\top x}$, [17].

El sistema difuso considera que la salida es una función lineal de las entradas, *i.e.*,

$$\text{Si } H^l \text{ entonces } g_l(x) = a_{l,0} + a_{l,1}x_1 + \dots + a_{l,n}x_n, \quad (15)$$

donde n es el número de entradas y H^l es un conjunto difuso de entrada dado por

$$H^l = \{(x, \mu_{H^l}(x)) : x \in \mathcal{X}_1 \times \dots \times \mathcal{X}_n\},$$

\mathcal{X}_i es el i -ésimo universo y $\mu_{H^l}(x)$ es la función de pertenencia asociada a H^l , la cual representa la premisa para la regla l y $g_l(x) = \underline{a}_l^\top \hat{x}$; $\underline{a}_l^\top = [a_{l,0}, a_{l,1}x_1, \dots, a_{l,n}x_n]$ y $\hat{x} = [1, x^\top]^\top$ para $l = 1, \dots, R$. El sistema difuso resultante es el promedio ponderado de las salidas $g_l(x)$ y esta defi-

nido por

$$f(x | \theta) = \frac{\sum_{l=1}^R g_l(x) \mu_{H^l}(x)}{\sum_{l=1}^R \mu_{H^l}(x)}, \quad (16)$$

donde R es el número de reglas en la regla-base. El algoritmo funciona de la siguiente manera:

1) inicialización de parámetros: se especifica el "factor de solapamiento" m . Si $m > 1$ los puntos con menor grado de pertenencia tienen menor influencia en el cálculo de un nuevo centro. Luego, se especifica el número de grupos (clases) R que se desea calcular; el cual es igual al número de reglas en la regla base y debe ser menor o igual al número de datos en el conjunto de entrenamiento G , *i.e.*, $R \leq M$. Entonces, se especifica la tolerancia $\epsilon_c > 0$ que determina el error permitido en el cálculo para los centros de los grupos. Finalmente, se procede a inicializar los centros de los grupos \underline{v}_0^l .

2) cálculo de los nuevos centros: se utilizan los centros iniciales \underline{v}_0^l de manera que la función objetivo (14) sea minimizada. Las condiciones para la minimización J están dadas por:

$$\underline{v}_{new}^l = \frac{\sum_{i=1}^M x^i (\mu_{il}^{new})^m}{\sum_{i=1}^M (\mu_{il}^{new})^m}, \quad (17)$$

donde,

$$\mu_{il}^{new} = \left[\sum_{k=1}^R \left(\frac{|x^i - \underline{v}_{old}^l|^2}{|x^i - \underline{v}_{old}^k|^2} \right)^{\frac{1}{m-1}} \right]^{-1}, \quad (18)$$

para $i = 1, 2, \dots, M$ y $l = 1, 2, \dots, R$ y $\sum_{l=1}^R \mu_{il}^{new} = 1$

3) evaluación: se compara la distancia entre los centros nuevos de los grupos \underline{v}_{new}^l y los centros previos \underline{v}_{old}^l mediante,

$$|\underline{v}_{new}^l - \underline{v}_{old}^l| < \epsilon_c \quad l = 1, 2, \dots, R \quad (19)$$

si se cumple (19) significa que los centros \underline{v}_{new}^l representan correctamente a los datos, el algoritmo de agrupamiento termina y se procede con la decodificación de la salida (Paso 4). Caso contrario, se continua iterando mediante (17) y (18) hasta que los nuevos centros satisfagan (19).

4) decodificación de la salida: se calcula la función $g_l(x) = \underline{a}_l^\top(\hat{x})$, $l = 1, 2, \dots, R$ para cada regla, es decir para cada centro de grupo \underline{v}_l , se minimiza la función J_l

$$J_l = \sum_{i=1}^M (\mu_{il})^2 (y^i - (\hat{x}^i)^\top \underline{a}_l)^2 \quad (20)$$

donde, $l = 1, 2, \dots, R$, y^i es la parte de la salida del i -ésimo par de datos $d^{(i)} = (x^i, y^i)$ y el producto entre $(\hat{x}^i)^\top$ y \underline{a}_l define la salida asociada con la l -ésima regla para el i -ésimo punto de datos de entrenamiento, [4]. Al observar (20), se puede notar que la minimización de J_l respecto a \underline{a}_l corresponde a un problema de mínimos cuadrados ponderados, cuya solución es:

$$\underline{a}_l = (\hat{X}^\top D_l^2 \hat{X})^{-1} \hat{X}^\top D_l^2 Y,$$

donde,

$$\hat{X} = \begin{bmatrix} 1 & \dots & 1 \\ x^1 & \dots & x^M \end{bmatrix},$$

$$Y = [y^1, \dots, y^m]^\top,$$

$$D_l^2 = (\text{diag}[\mu_{1l}, \dots, \mu_{Ml}])^2.$$

Una descripción detallada de este método generalizado al caso entrada-salida múltiple, se encuentra en el Algoritmo 4.

Algoritmo 4 Agrupamiento difuso combinado entrada-salida múltiple

Entrada: $X_{M \times n}$ matriz de las variables de entrada

$Y_{M \times H}$ matriz de las variables de salida

R número de reglas (clases o grupos)

$\epsilon_c > 0, m > 1$

1: $X_{\min} \leftarrow \min\{x_j, j = 1, \dots, n\}$

2: $X_{\max} \leftarrow \max\{x_j, j = 1, \dots, n\}$

3: $\underline{v}_{0j}^l \leftarrow X_{\min,j} + l \left(\frac{X_{\max,j} - X_{\min,j}}{R + 1} \right)$ $j = 1, \dots, n$,
 $l = 1, \dots, R$

4: $\underline{v}_{new}^l \leftarrow \underline{v}_0^l$ $l = 1, \dots, R$

5: **mientras** $|\underline{v}_{new}^l - \underline{v}_{old}^l| > \epsilon_c$ $i \leftarrow 1, \dots, M, l \leftarrow 1, \dots, R$
hacer

6: $\underline{\mu}_{il}^{new} \leftarrow \left[\sum_{k=1}^R \left(\frac{|x^i - \underline{v}_{old}^k|^2}{|x^i - \underline{v}_{old}^k|^2} \right)^{\frac{1}{m-1}} \right]^{-1}$

7: $\underline{v}_{new}^l \leftarrow \frac{\sum_{i=1}^M x^i (\underline{\mu}_{il}^{new})^m}{\sum_{i=1}^M (\underline{\mu}_{il}^{new})^m}$

8: **fin mientras**

9: **para** $i \leftarrow 1, M$ **hacer**

10: $\hat{x}^i \leftarrow [1, x^i]$

11: **fin para**

12: $\hat{X} \leftarrow \hat{x}^\top$

13: $Y \leftarrow [y^1, \dots, y^m]^\top$

14: $D_l^2 \leftarrow (\text{diag}[\mu_{1l}, \dots, \mu_{Ml}])^2$

15: **para** $h \leftarrow 1, H$ **hacer**

16: **para** $l \leftarrow 1, R$ **hacer**

17: $\underline{a}_{lh} \leftarrow (\hat{X}^\top D_l^2 \hat{X})^{-1} \hat{X}^\top D_l^2 Y_h$

18: **devolver** $g_{lh}(x) \leftarrow \underline{a}_{lh}^\top \hat{x}$

19: **fin para**

20: **fin para**

2.6 Mínimos cuadrados recursivo combinado (MCRC)

Hasta ahora se han descrito y generalizado los métodos automáticos de regresión difusa que se encuentran en la literatura, [13, 15]. En esta sección, se propone el método de mínimos cuadrados recursivo combinado, el cual constituye el principal aporte teórico de este trabajo. Este método se compone de dos etapas; en primer lugar, se utiliza el algoritmo c-medias difuso para la identificación de las reglas del sistema; luego, se aplica el método de mínimos cuadrados recursivo para caracterizar el conjunto de salida [2]. Los resultados numéricos demuestran el buen desempeño de este método, como se verá en la Sección 3.

El método de mínimos cuadrados recursivo, requiere que la regla-base del sistema esté completamente especificada, *i.e.*, número de reglas, centros de funciones de pertenencia, etc. Puesto que las reglas del sistema en los métodos automáticos se encargan de predecir y/o gobernar las salidas de los sistemas difusos, es necesario un algoritmo que permita una adecuada especificación de reglas. El algoritmo c-medias difuso especifica eficientemente las reglas del sistema; y luego, se procede a la aplicación del método de mínimos cuadrados recursivo. Este método posee características muy útiles para ser implementado:

- no requiere la inversión de una matriz, como en el caso de los mínimos cuadrados por lotes y agrupamiento difuso combinado. En su lugar, se calcula únicamente el inverso de un escalar,
- los mínimos cuadrados recursivos son más eficientes para incluir las reglas del sistema [13],
- la convergencia del método es competitiva respecto a otras técnicas [7].

Estas razones justifican la implementación del método de mínimos cuadrados recursivo.

El método MCRC se describe a continuación:

1. **Determinar el número de centros:** se inicializa R ,

$$R = c, \\ \underline{v}_{new}^l = \underline{v}_0^l \quad l = 1, \dots, R$$

para lo cual se aplica algún método heurístico.

2. **Calcular una partición de los datos:** se asigna cada punto al grupo más cercano utilizando la función de pertenencia

$$\underline{\mu}_{il}^{new} = \left[\sum_{k=1}^R \left(\frac{|x^i - \underline{v}_{old}^k|^2}{|x^i - \underline{v}_{old}^k|^2} \right)^{\frac{1}{m-1}} \right]^{-1},$$

3. **Calcular los nuevos centros:** mediante

$$\underline{v}_{new}^l = \frac{\sum_{i=1}^M x^i (\underline{\mu}_{il}^{new})^m}{\sum_{i=1}^M (\underline{\mu}_{il}^{new})^m},$$

4. **Evaluar:** si $|\underline{v}_{new}^l - \underline{v}_{old}^l| \leq \epsilon_c$ para $i = 1, \dots, M, l = 1, \dots, R$ la partición es estable, el proceso se detiene. Caso contrario, se regresa al Paso 2.
5. **Inicializar los parámetros del MCR:** $\hat{\theta}(0) = 0, P(0) = \alpha I$
6. **Calcular los grados de pertenencia:** utilizando los centros, $i = 1, \dots, M$ y $l = 1, \dots, R$

$$\mu_{il} = \prod_{j=1}^n \exp \left[-\frac{1}{2} \left(\frac{x_j^i - \underline{v}_j^l}{\sigma_j^l} \right)^2 \right]$$

$$\zeta_{il} = \frac{\mu_{il}}{\sum_{l=1}^R \mu_{il}}$$

7. Aplicar el MCR

$$P(k) = \frac{1}{\lambda} \left\{ I - P(k-1) \zeta(x^k) \left[\lambda I + (\zeta(x^k))^T P(k-1) \zeta(x^k) \right]^{-1} (\zeta(x^k))^T \right\} P(k-1)$$

$$\hat{\theta}(k) = \hat{\theta}(k-1) + P(k) \zeta(x^k) \left[y^k - (\zeta(x^k))^T \hat{\theta}(k-1) \right]$$

Al aplicar el algoritmo c-medias difuso, se debe inicializar el número de reglas R (grupos o clases) y los valores de los centros de las clases cuidadosamente, ya que los resultados son sensibles a la elección de estos valores iniciales. La generalización al caso de entrada-salida múltiple de este método se encuentra en el Algoritmo 5.

3 Aplicación

En esta sección, se presentan resultados de las pruebas numéricas realizadas con información de la Reserva Internacional de libre Disponibilidad (RILD) de Ecuador¹ y un índice de liquidez financiera². El objetivo de validar los métodos y analizar su comportamiento.

Los resultados para cada una de las aplicaciones numéricas fueron obtenidos mediante una aplicación desarrollada en R [3]. Para ello se implementará los métodos:

- mínimos cuadrados por lotes (MCL),
- mínimos cuadrados recursivo (MCR),
- aprendizaje desde el ejemplo modificado (AEM),
- agrupamiento difuso combinado (ADC),
- mínimos cuadrados recursivo combinado (MCRC).

Estos métodos fueron implementados utilizando funciones de pertenencia Gaussianas y Triangulares. Únicamente en el caso del agrupamiento difuso combinado no se implementó estas funciones, debido a que este método define específicamente su función de pertenencia.

Para visualizar de mejor manera el desempeño de cada método, se estandarizaron los datos, *i.e.*, se utilizó una escala porcentual o, en su defecto, se normalizaron y estandarizaron las variables, [10]. Se dividió al conjunto de datos en un conjunto de entrenamiento y un conjunto de pruebas (*back testing*); además, para contrastar los resultados se calcularon modelos de series de tiempo (SARIMA) para las aplicaciones de estudio, utilizando el paquete estadístico *forecast* y la función *auto.arima* del software R [14].

3.1 Reserva Internacional de Libre Disponibilidad (RILD) de Ecuador

La administración de las reservas internacionales requiere de objetivos claros, sistemas de control fuertes y una apreciación realista de las limitaciones que se tienen. Un manejo eficiente de estas reservas implica una administración macroeconómica exitosa para un país. La mayor parte de los bancos centrales tienden a mantener reservas en forma de bonos, los cuales devengan intereses de bajo riesgo, o de otros activos de riesgo similar en algún tipo de divisa. Lo anterior constituye un intento de asegurar los valores de mercado de sus activos externos, los cuales son susceptibles a sufrir variaciones, de manera que sean siempre mayores que el valor de los pasivos monetarios.

Es de gran importancia para un país asegurar y fortalecer su capacidad para afrontar acontecimientos inesperados, como son los choques financieros exógenos. Estrategias sanas de manejo de reservas conllevan una amplia gama de objetivos de política, tales como:

- asegurar un alto nivel de confianza en las políticas monetaria y cambiaria de la economía, particularmente en regímenes de tipo de cambio fijo;
- mantener divisas líquidas durante los ataques de los choques externos;
- darle confianza a la comunidad internacional de que la economía es capaz de cumplir con sus obligaciones externas.

Particularmente, los depósitos realizados en el Banco Central del Ecuador por instituciones del sector público y del sistema financiero privado, por ser exigibles a la vista, están respaldados por la Reserva Internacional de Libre Disponibilidad (RILD). Ésta, por su carácter de activo de reserva del país y por su tamaño, es el componente de mayor importancia entre los activos del Banco Central del Ecuador.

¹Información proveniente de la Dirección de Estudios del Banco Central de Ecuador.

²Información correspondiente a los Boletines Financieros publicados mensualmente por la Superintendencia de Bancos y Seguros.

Los recursos de la RILD son invertidos, por mandato legal, en los mercados financieros internacionales, bajo los principios de seguridad, liquidez y rentabilidad. Una parte de los rendimientos obtenidos se revierte al Presupuesto General del Estado y otra parte es transferida mensualmen-

te al Instituto Ecuatoriano de Seguridad Social (IESS), para mejorar las prestaciones a favor de sus afiliados. Básicamente, la RILD corresponde al monto de activos externos de alta liquidez que tienen una contrapartida en los pasivos del balance del Banco Central del Ecuador³.

Algoritmo 5 Mínimos cuadrados recursivo combinado entrada-salida múltiple

Entrada: $X_{M \times n}$ matriz las variables de entrada

$Y_{M \times H}$ matriz de las variables de salida

K número de iteraciones del algoritmo

R número de reglas (clases o grupos)

$\epsilon_c > 0, m > 1, 0 < \lambda \leq 1$

1: $X_{\min} \leftarrow \min\{x_j, j = 1, \dots, n\}$

2: $X_{\max} \leftarrow \max\{x_j, j = 1, \dots, n\}$

3: $\vartheta_{0j}^l \leftarrow X_{\min,j} + l \left(\frac{X_{\max,j} - X_{\min,j}}{R + 1} \right) \quad j = 1, \dots, n, l = 1, \dots, R$

4: $\vartheta_{new}^l \leftarrow \vartheta_0^l \quad l \leftarrow 1, \dots, R$

5: **mientras** $|\vartheta_{new}^l - \vartheta_{old}^l| > \epsilon_c \quad i \leftarrow 1, \dots, M, l \leftarrow 1, \dots, R$ **hacer**

6: $\mu_{il}^{new} \leftarrow \left[\sum_{k=1}^R \left(\frac{|x^i - \vartheta_{old}^l|^2}{|x^i - \vartheta_{old}^k|^2} \right)^{\frac{1}{m-1}} \right]^{-1}$

7: $\vartheta_{new}^l \leftarrow \frac{\sum_{i=1}^M x^i (\mu_{il}^{new})^m}{\sum_{i=1}^M (\mu_{il}^{new})^m}$

8: **fin mientras**

9: $\hat{\theta}(0) \leftarrow 0$

10: $P(0) \leftarrow \alpha I$

11: $\sigma_j^l \leftarrow \sigma \quad j = 1, \dots, n$

12: **para** $i \leftarrow 1, M$ **hacer**

13: **para** $l \leftarrow 1, R$ **hacer**

14: $\mu_{il} \leftarrow \prod_{j=1}^n \exp \left[-\frac{1}{2} \left(\frac{x_j^i - \vartheta_j^l}{\sigma_j^l} \right)^2 \right]$

15: **fin para**

16: $\zeta_{il} \leftarrow \frac{\mu_{il}}{\sum_{l=1}^R \mu_{il}}$

17: **fin para**

18: **para** $h \leftarrow 1, H$ **hacer**

19: $s \leftarrow 0$ contador auxiliar

20: **para** $k \leftarrow 1, K \times M$ **hacer**

21: **si** $s \leq M$ **entonces**

22: $s \leftarrow s + 1$

23: **si no**

24: $s \leftarrow 1$

25: **fin si**

26: $P_h(k) \leftarrow \frac{1}{\lambda} \left\{ I - P_h(k-1)\zeta(x^s) \left[\lambda I + (\zeta(x^s))^\top P_h(k-1)\zeta(x^s) \right]^{-1} (\zeta(x^s))^\top \right\} P_h(k-1)$

27: $\hat{\theta}_h(k) \leftarrow \hat{\theta}_h(k-1) + P_h(k)\zeta(x^s) \left[y_h^s - (\zeta(x^s))^\top \hat{\theta}_h(k-1) \right]$

28: **fin para**

29: **devolver** $f_h(x | \hat{\theta}_h) \leftarrow \hat{\theta}_h^\top \zeta(x)$

30: **fin para**

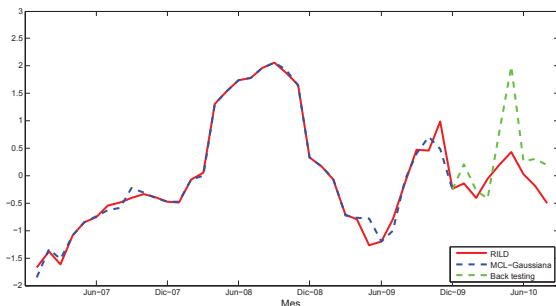
³Antes de adoptar la dolarización en la economía ecuatoriana, el concepto que se manejaba era el de Reserva Monetaria Internacional, el mismo que representaba el saldo entre activos y pasivos externos, y está sujeto al control de las autoridades económicas del país; sus componentes eran: el oro monetario, los derechos especiales de giro, la posición y uso de la reserva en el Fondo Monetario Internacional (FMI), los activos en divisas y otros títulos de crédito.

El Banco Central del Ecuador calcula la reserva monetaria de libre disponibilidad que sirve de base para mantener los sistemas de canje, reserva financiera y otras operaciones del Banco Central del Ecuador, [1].

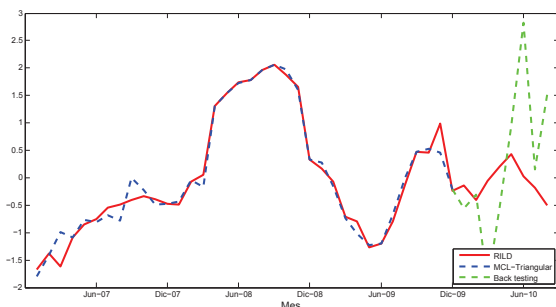
Se propone modelar la RILD, considerando las variables:

Variable	Descripción
Y	Reserva Internacional de Libre Disponibilidad (RILD) de Ecuador
X ₁	Ingresos al sistema financiero
X ₂	Egresos del sistema financiero
X ₃	Exportaciones de petróleo
X ₄	Importaciones de derivados de petróleo

La información corresponde a series mensuales del período enero 2007 - agosto 2010. Para este caso, se considera como conjunto de entrenamiento al período comprendido enero 2007 - diciembre 2009 y, como conjunto de pruebas, (*back testing*) al período enero 2010 - agosto 2010.



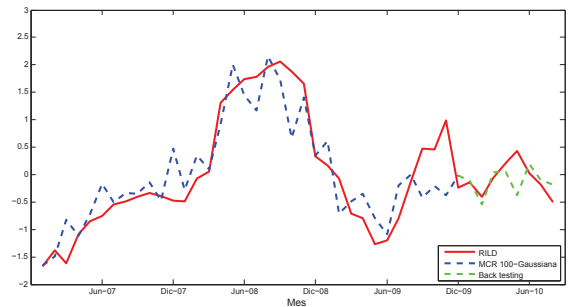
(a) Función gaussiana



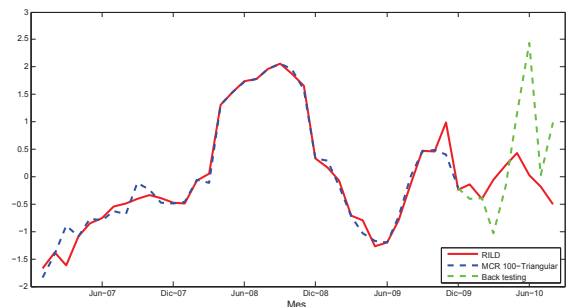
(b) Función triangular

Figura 1. RILD, mínimos cuadrados por lotes (MCL)

La Figura 1 muestra el ajuste para el método de mínimos cuadrados por lotes, con funciones de pertenencia Gaussiana y Triangular, se observa que el ajuste en el conjunto de entrenamiento es bueno para ambas series; sin embargo, en el *back testing*; los pronósticos para ambas series se tornan volátiles.

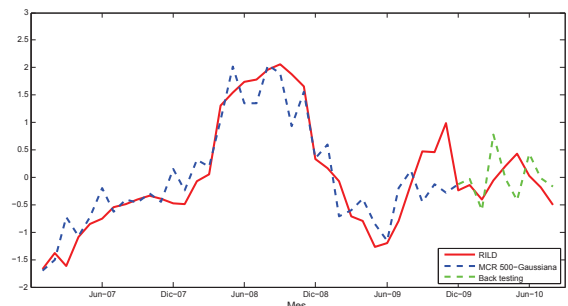


(a) Función gaussiana

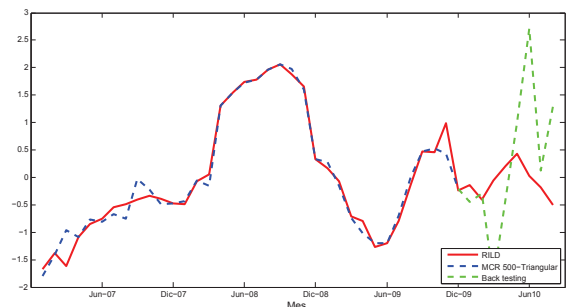


(b) Función triangular

Figura 2. RILD, mínimos cuadrados recursivo 100 iteraciones (MCR)



(a) Función gaussiana

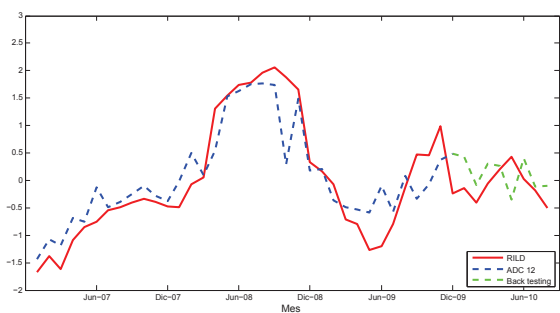


(b) Función triangular

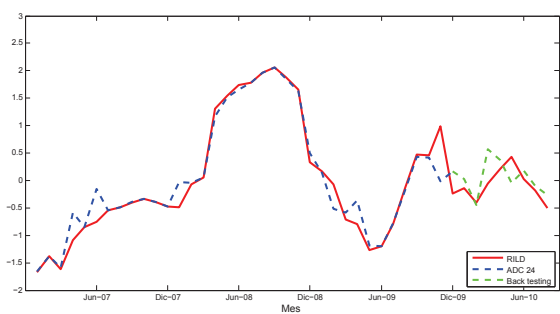
Figura 3. RILD, mínimos cuadrados recursivo 500 iteraciones (MCR)

En el caso del método de mínimos cuadrados recursivo con 100 y 500 iteraciones (Figura 2 y Figura 3 respectivamente), para la función Gaussiana (Figura 2a y Figura 3a), el ajuste en el entrenamiento es un tanto volátil, aún cuando existe una similitud en el comportamiento entre la serie entrenada y la serie original. Al analizar el *back testing*, se nota cierta semejanza con los valores reales de la RILD. En el caso de la función de pertenencia Triangular (Figura 2b y Figura 3b), se observa que el entrenamiento es superior, pero los pronósticos no se asemejan a los de la serie original.

Para el agrupamiento difuso combinado, se utilizaron 12 y 24 centros, los resultados se pueden observar en la Figura 4. En el caso en el cual se utiliza 12 centros (Figura 4a) en el conjunto de entrenamiento, se aprecia una semejanza con la serie original; algo similar ocurre en el *back testing*. Cuando se aumenta a 24 centros (Figura 4b), se observa una mejora muy significativa, tanto en el conjunto de entrenamiento como en el *back testing*, capturándose de mejor manera los cambios de tendencia.

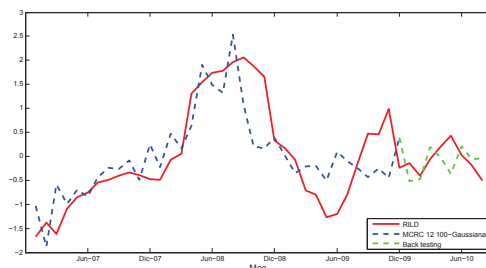


(a) Centros $c = 12$

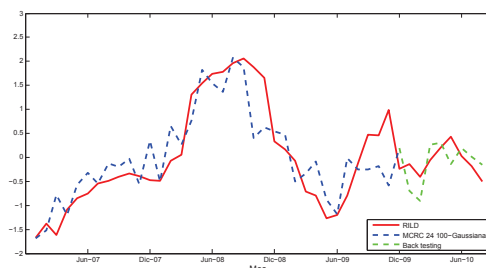


(b) Centros $c = 24$

Figura 4. RILD, agrupamiento difuso combinado (ADC)

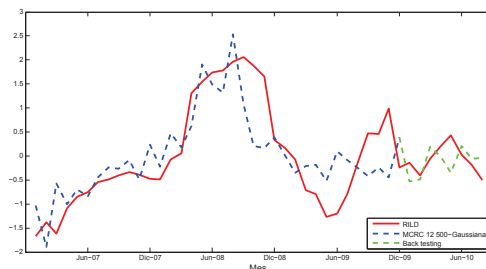


(a) Centros $c = 12$

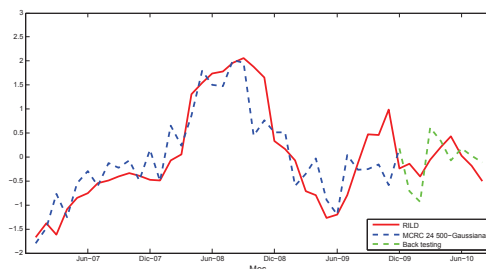


(b) Centros $c = 24$

Figura 5. RILD, mínimos cuadrados recursivo combinado 100 iteraciones (MCRC) Gaussiana



(a) Centros $c = 12$



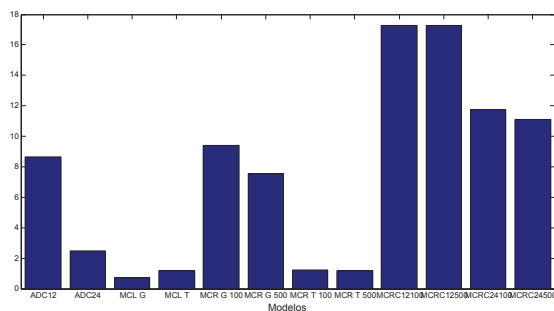
(b) Centros $c = 24$

Figura 6. RILD, mínimos cuadrados recursivo combinado 500 iteraciones (MCRC) Gaussiana

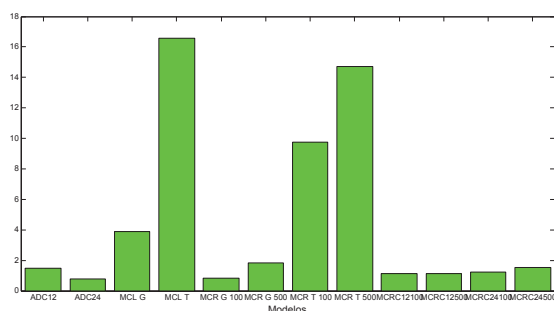
En el método de mínimos cuadrados recursivo combinado con 100 y 500 iteraciones (Figura 5 y Figura 6), se utilizaron 12 y 24 centros. Con 100 iteraciones, el ajuste del conjunto de entrenamiento es similar para 12 y 24 centros; pero cuando se calculan los pronósticos, hay mejores resultados con 12 centros; ver Figura 5a. Lo mismo sucede cuando se incrementa a 500 iteraciones, *i.e.*, con 12 centros los pronósticos se asemejan mucho más a la serie real (Figura 6a).

La Figura 7 presenta, para cada uno de los métodos, la suma de errores del modelo y *back testing*. Los mejores métodos, según estos criterios, son mínimos cuadrados por lotes, agrupamiento difuso combinado 24 (24 grupos) y mínimos cuadrados recursivo combinado 12-100 (12 centros, 100 iteraciones).

Con el objetivo de comparar los resultados de los métodos automáticos, se calculó un modelo SARIMA para la RILD, utilizando el mismo conjunto de entrenamiento. Los resultados del modelo SARIMA muestran que el modelo más apropiado para su estimación y pronóstico es un SARIMA(3, 0, 0)(1, 0, 0). Estos resultados son comparados con los mejores métodos automáticos en la Figura 8. Se puede apreciar que los métodos de mínimos cuadrados por lotes, agrupamiento difuso combinado 24 y el modelo SARIMA son los mejores cuando se ejecuta el entrenamiento (residuos de los modelos); pero, cuando se calculan los pronósticos, los métodos de agrupamiento difuso combinado 24 y mínimos cuadrados recursivo combinado 12-100 son más competitivos.



(a) Residuos de los modelos



(b) Residuos del *back testing*

Figura 7. RILD, residuos de los métodos automáticos y *back testing*

3.2 Índice de Liquidez de Instituciones Financieras

En este apartado se aplican los métodos descritos para un índice de liquidez de los bancos: Pichincha, Internacional y Amazonas, en ejemplos del caso de entrada y salida múltiple.

La liquidez de una institución financiera se refiere a su capacidad de contar con los fondos necesarios para:

- cumplir oportunamente con los compromisos financieros adquiridos, cubrir retiros de depósitos a la vista, pagar una deuda, o pagar a proveedores;
- aprovechar oportunidades de negocio, *i.e.*, compra de cartera, inversión a tasa atractiva, etc.

Cuando una institución enfrenta una escasez de fondos para cumplir con sus obligaciones, tiene la necesidad de conseguir recursos alternativos o de vender activos en condiciones desfavorables, con un alto costo financiero o una elevada tasa de descuento incurriendo en pérdidas de valoración; por esta razón, una institución debe establecer e implementar políticas y procedimientos idóneos que le permitan una adecuada administración de su liquidez, considerando la complejidad y el volumen de las operaciones que realiza. Dichas políticas y procedimientos deben tener en cuenta los posibles escenarios y la forma en la que la institución respondería, en el caso de que tales eventos sean reales [16]. Sin embargo, la preferencia por la liquidez puede ser contraproducente para la economía general; *e.g.*, si hay incertidumbre o los precios decaen, la población en general puede preferir mantener sus activos, tendencia que, si se generaliza, puede llevar a una gran disminución de la demanda.

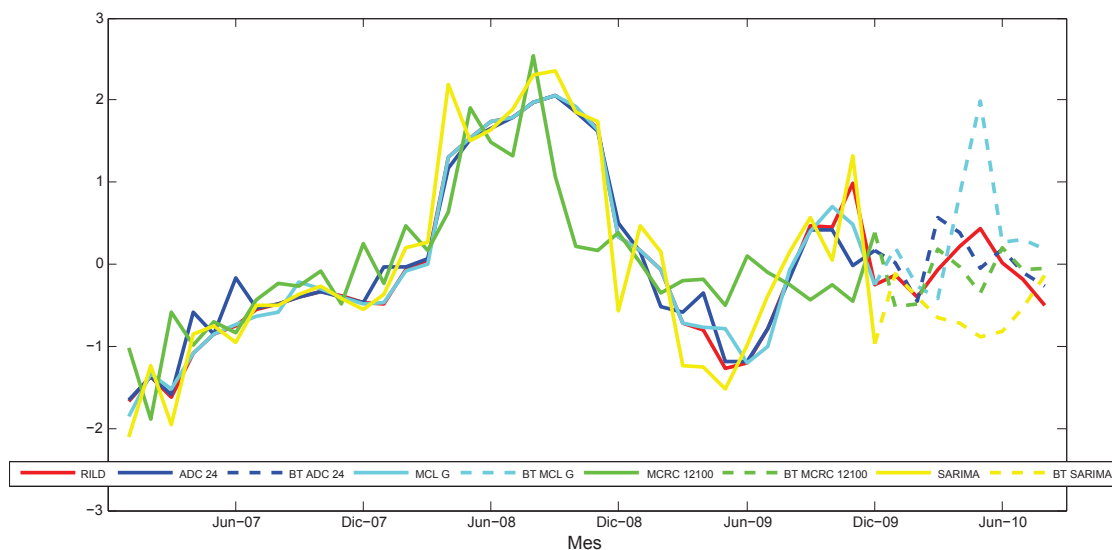
El siguiente modelo considera el índice definido como:

$$\text{Liquidez} = \frac{\text{Fondos disponibles}}{\text{Total de depósitos a corto plazo}}$$

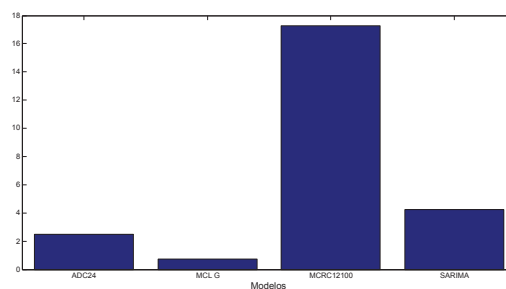
por lo que se consideraron las siguientes variables:

Variable	Descripción
Y_1	Índice de liquidez Banco Pichincha
Y_2	Índice de liquidez Banco Internacional
Y_3	Índice de liquidez Banco Amazonas
X_1	Índice de precios al consumidor
X_2	Inflación

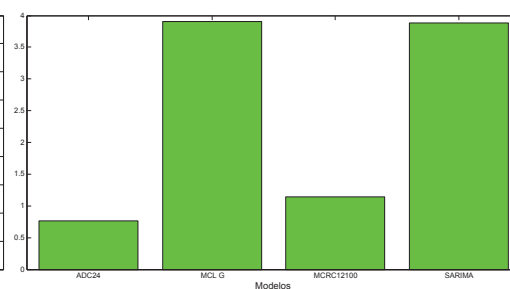
La información corresponde a las series mensuales del período enero 2006 - octubre 2010. Como conjunto de entrenamiento se consideró al período enero 2006 - diciembre 2009 y, como conjunto de *back testing*, al período enero 2010 - octubre 2010.



(a) Contraste

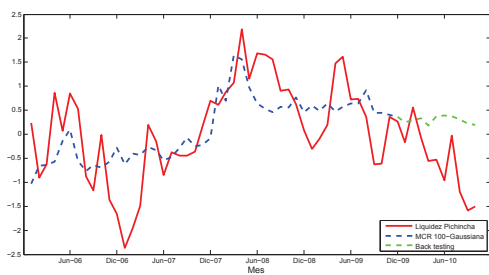


(b) Residuos de los modelos

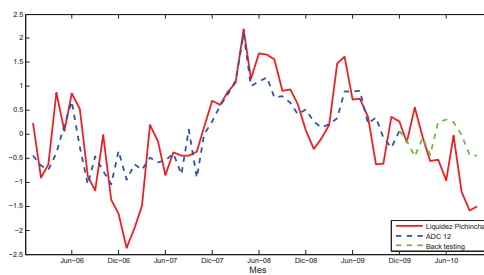


(c) Residuos del back testing

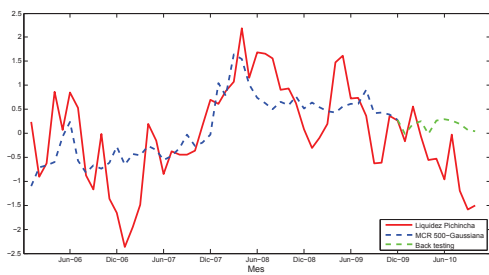
Figura 8. RILD, contraste



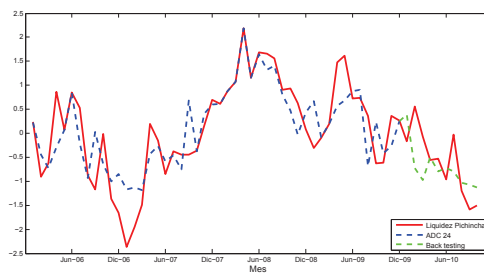
(a) 100 iteraciones



(a) Centros $c = 12$



(b) 500 iteraciones



(b) Centros $c = 24$

Figura 9. Banco Pichincha, mínimos cuadrados recursivo (MCR)

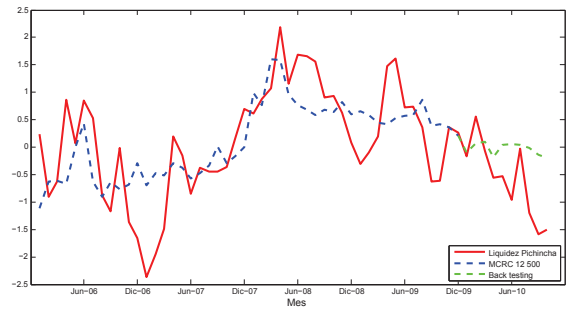
Figura 10. Banco Pichincha, agrupamiento difuso combinado (ADC)

A continuación, se presentan los resultados de los métodos de mínimos cuadrados recursivo, agrupamiento difuso combinado y mínimos cuadrados recursivo combinado para el Banco Pichincha.

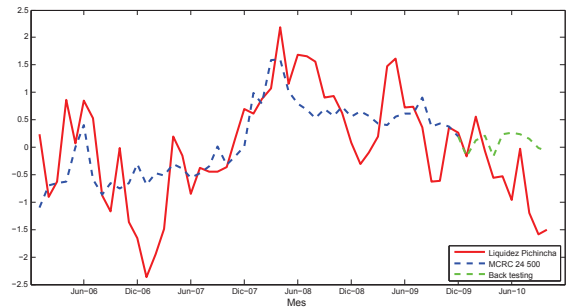
La Figura 9 muestra el ajuste para el método de mínimos cuadrados recursivo con 100 y 500 iteraciones y con función de pertenencia Gaussiana. Se observa que el ajuste en el conjunto de entrenamiento recoge la tendencia de la serie original. Sin embargo, el *back testing* no se asemeja a los valores observados.

Para el agrupamiento difuso combinado, se aplicó 12 y 24 centros; los resultados se observan en la Figura 10. Con 12 centros (Figura 10a) en el conjunto de entrenamiento, se aprecia una semejanza con la serie original; algo similar sucede con los pronósticos en el *back testing*. Pero cuando se incrementa a 24 centros (Figura 10b), se aprecia una mejora significativa tanto en el conjunto de entrenamiento como en el *back testing*.

En el método de mínimos cuadrados recursivo combinado con 100 y 500 iteraciones (Figura 11 y Figura 12), se utilizaron 12 y 24 centros. Con 100 iteraciones el ajuste del conjunto de entrenamiento es muy parecido para 12 y 24 centros; lo mismo ocurre cuando se calcula los pronósticos; en estos casos no se asemejan a la serie del índice de liquidez; ver Figura 11a. Un fenómeno similar ocurre cuando se aumenta a 500 iteraciones (Figura 12a). Esto podría deberse a que el comportamiento de este indicador es muy volátil.

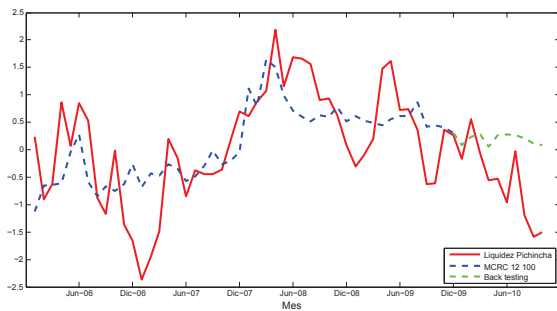


(a) Centros $c = 12$

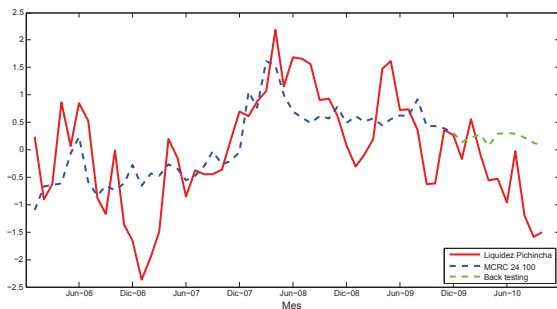


(b) Centros $c = 24$

Figura 12. Banco Pichincha, mínimos cuadrados recursivo combinado 500 iteraciones (MCRC) Gaussiana

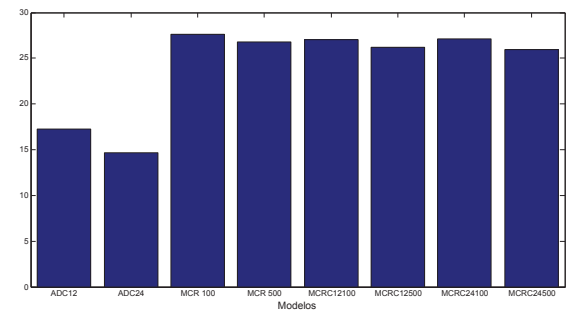


(a) Centros $c = 12$

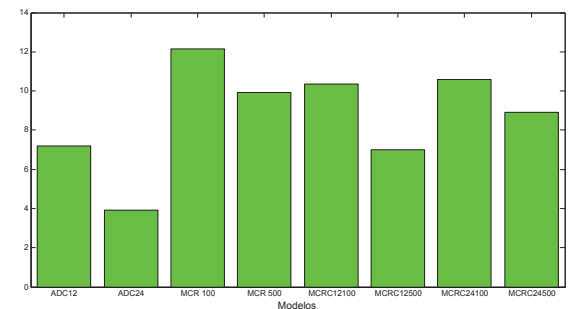


(b) Centros $c = 24$

Figura 11. Banco Pichincha, mínimos cuadrados recursivo combinado 100 iteraciones (MCRC) Gaussiana

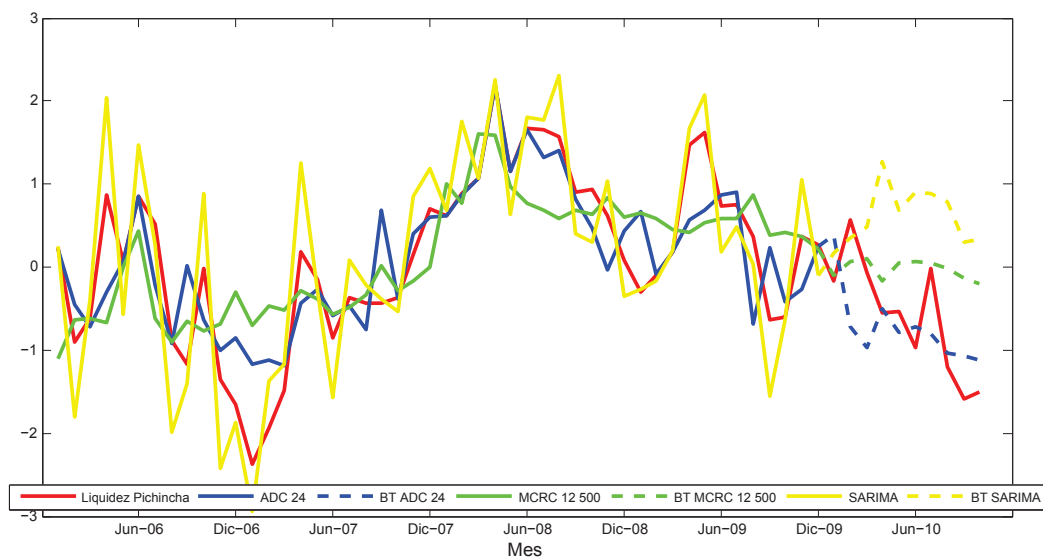


(a) Residuos de los modelos

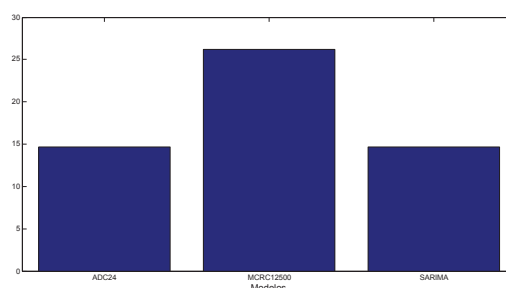


(b) Residuos del *back testing*

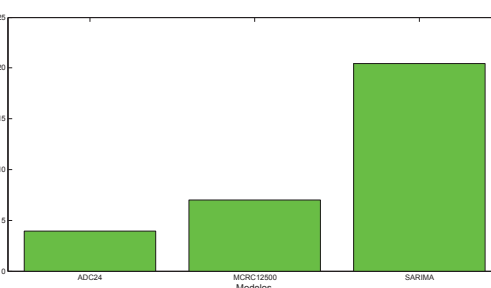
Figura 13. Banco Pichincha, residuos de los métodos automáticos y *back testing*



(a) Contraste

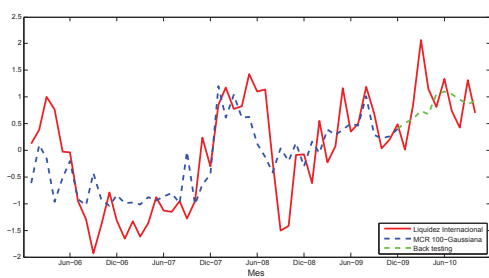


(b) Residuos de los modelos

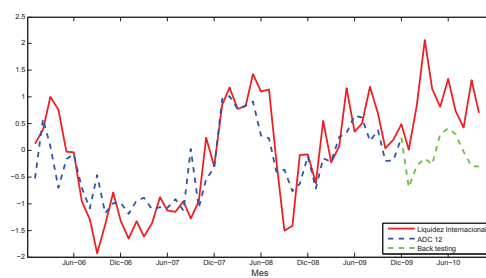


(c) Residuos del back testing

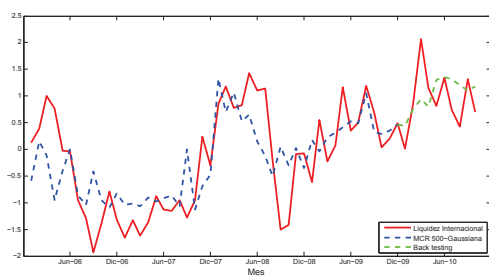
Figura 14. Banco Pichincha contraste



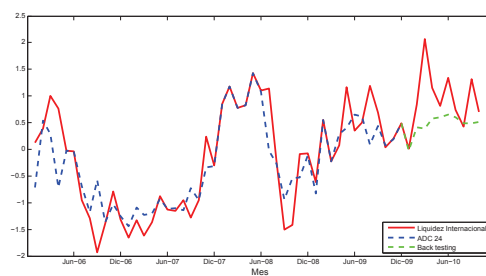
(a) 100 iteraciones



(a) Centros $c = 12$



(b) 500 iteraciones



(b) Centros $c = 24$

Figura 15. Banco Internacional, mínimos cuadrados recursivo (MCR)

Figura 16. Banco Internacional, agrupamiento difuso combinado (ADC)

La Figura 13 presenta para cada uno de los métodos la suma de errores del modelo y *back testing*. Los mejores métodos, según estos criterios, son agrupamiento difuso combinado 24 (24 grupos) y mínimos cuadrados recursivo combinado 12-500 (12 centros, 500 iteraciones).

Finalmente, se calculó el modelo SARIMA para el índice de liquidez del Banco Pichincha utilizando el mismo conjunto de entrenamiento; estos resultados muestran que el modelo más apropiado para su estimación y pronóstico es un SARIMA(0, 1, 0)(2, 0, 0). Este modelo es comparado con los mejores métodos automáticos en la Figura 14. El método de agrupamiento difuso combinado 24 y el modelo SARIMA son los mejores al momento de la estimación (residuos de los modelos); si se evalúan los pronósticos, los métodos de agrupamiento difuso combinado 24 y mínimos cuadrados recursivo combinado 12-500 son más competitivos.

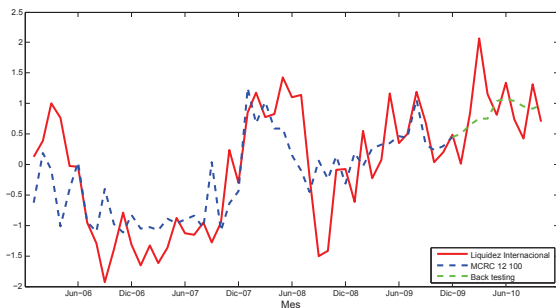
A continuación, se presentan los resultados de los métodos de mínimos cuadrados recursivo, agrupamiento difuso combinado y mínimos cuadrados recursivo combinado para el Banco Internacional.

La Figura 15 muestra el método de mínimos cuadrados recursivo con 100 y 500 iteraciones con función de perte-

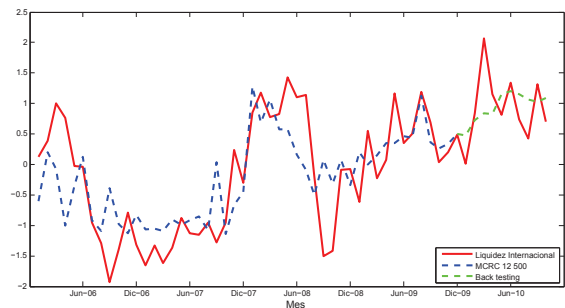
nencia Gaussiana. Se observa que el ajuste en el conjunto de entrenamiento recoge la tendencia de la serie original, y lo mismo ocurre cuando se ejecuta el *back testing*. En este caso no se aprecia ninguna diferencia significativa cuando se varía el número de iteraciones.

Para el método de agrupamiento difuso combinado, se aplicó 12 y 24 centros; los resultados se observan en la Figura 16. Con 12 centros (Figura 16a) en el conjunto de entrenamiento, se aprecia una semejanza con la serie original; lo mismo sucede con el *back testing*. Al incrementar a 24 centros (Figura 16b), se observa una mejora muy significativa en el *back testing*, capturando de manera más adecuada la tendencia de la serie original.

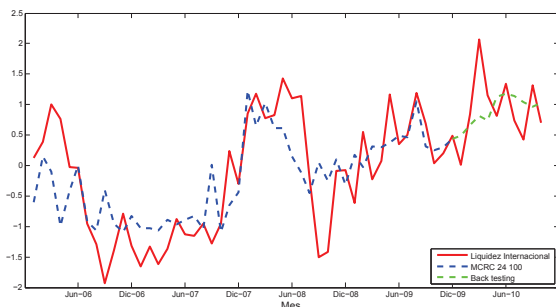
En el método de mínimos cuadrados recursivo combinado con 100 y 500 iteraciones (Figura 17 y Figura 18) se utilizaron 12 y 24 centros. Con 100 iteraciones, el ajuste del conjunto de entrenamiento es muy parecido para 12 y 24 centros; lo mismo ocurre cuando se calcula los pronósticos, ver Figura 17a. De la misma forma, cuando se aumenta a 500 iteraciones, no se observa una mejora significativa (Figura 18a). Aun si el ajuste no se asemeja a la serie, este método modela la tendencia del fenómeno.



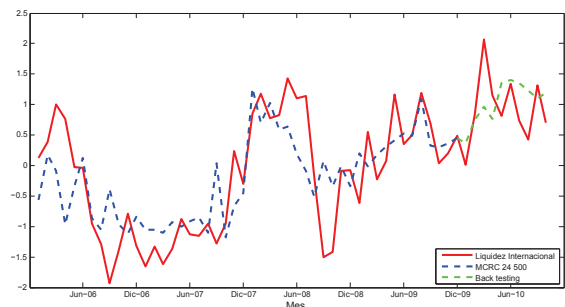
(a) Centros $c = 12$



(a) Centros $c = 12$



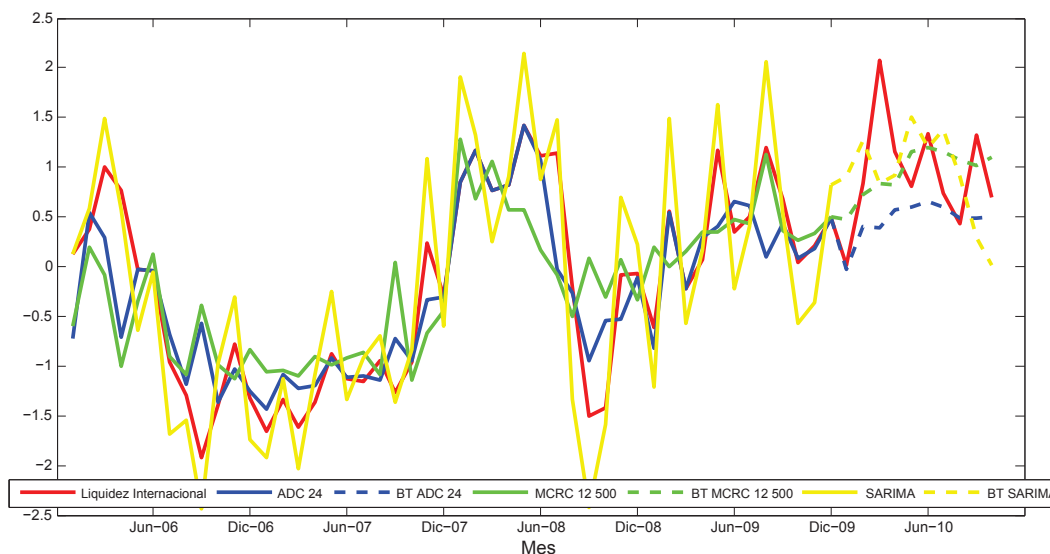
(b) Centros $c = 24$



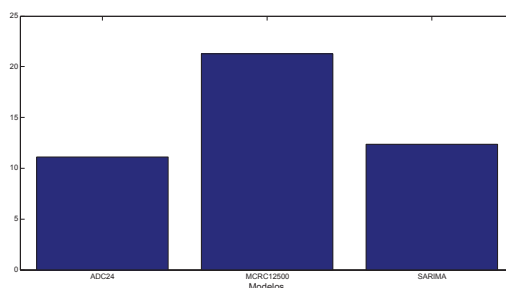
(b) Centros $c = 24$

Figura 17. Banco Internacional, mínimos cuadrados recursivo combinado 100 iteraciones (MCRC) Gaussiana

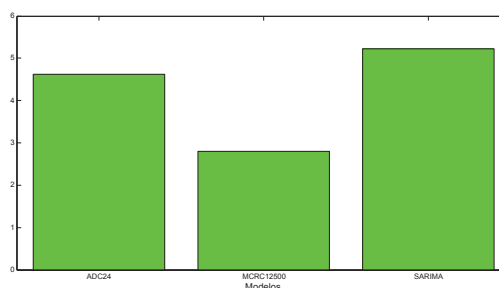
Figura 18. Banco Internacional, mínimos cuadrados recursivo combinado 500 iteraciones (MCRC) Gaussiana



(a) Contraste

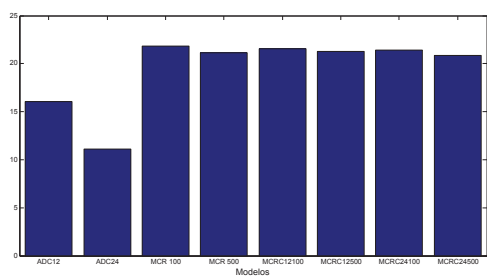


(b) Residuos de los modelos

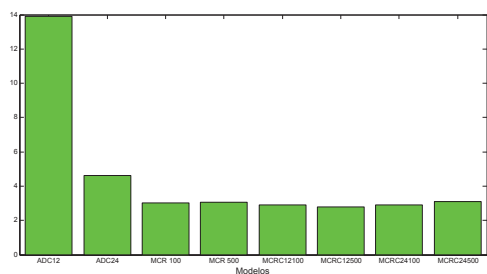


(c) Residuos del back testing

Figura 19. Banco Internacional, contraste



(a) Residuos de los modelos



(b) Residuos del back testing

Figura 20. Banco Internacional, residuos de los métodos automáticos y back testing

La Figura 20 presenta para cada uno de los métodos la suma de errores del modelo y el *back testing*. Los mejores métodos según estos criterios son agrupamiento difuso combinado 24 (24 grupos) y mínimos cuadrados recursivo combinado 12-500 (12 centros, 500 iteraciones).

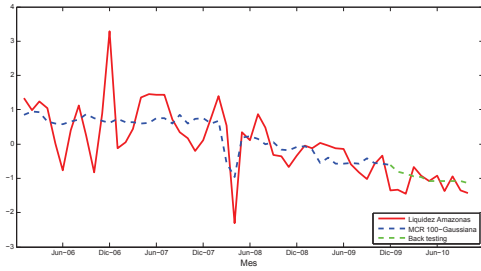
Nuevamente, se calculó un modelo SARIMA para el índice de liquidez del Banco Internacional, utilizando el mismo conjunto de entrenamiento, estos resultados muestran que el modelo más apropiado para su estimación y pronóstico es un SARIMA(0, 1, 0)(0, 0, 2). Este modelo es comparado con los mejores métodos automáticos en la Figura 19. Al igual que al calcular el índice de liquidez de Banco Pichincha, se aprecia que el método de agrupamiento difuso combinado 24 y el modelo SARIMA son los mejores al momento de la estimación (residuos de los modelos); para calcular los pronósticos, el método mínimos cuadrados recursivo combinado 12-500 es más competitivo.

A continuación, se presentan los resultados de los métodos de mínimos cuadrados recursivo, agrupamiento difuso combinado y mínimos cuadrados recursivo combinado para el Banco Amazonas.

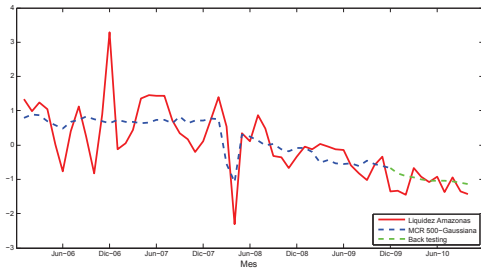
En la Figura 21 se observa el método de mínimos cuadrados recursivo con 100 y 500 iteraciones con función de

pertenencia Gaussiana. El ajuste en el conjunto de entrenamiento recoge la tendencia de la serie original; lo mismo ocurre cuando se ejecuta el *back testing*; nuevamente no se aprecia ninguna diferencia significativa cuando se varía el número de iteraciones.

Para el agrupamiento difuso combinado, se utilizó 12 y 24 centros (Figura 22). En este caso se observa que, al incrementar el número de centros a 24, se logra un mejor ajuste en el entrenamiento; ello se refleja en los pronósticos, cuando se aplica el *back testing*.

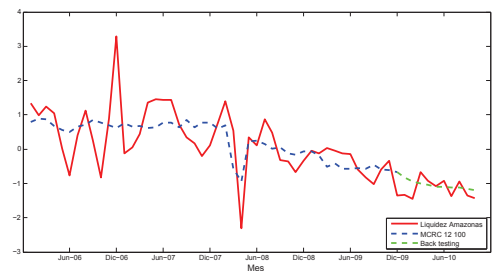


(a) 100 iteraciones

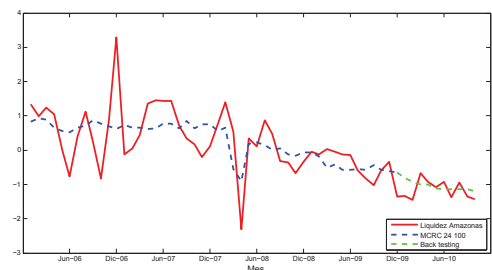


(b) 500 iteraciones

Figura 21. Banco Amazonas, mínimos cuadrados recursivo (MCR)

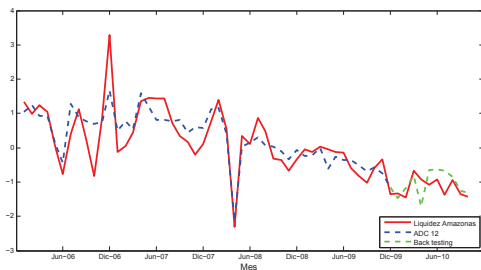


(a) Centros $c = 12$

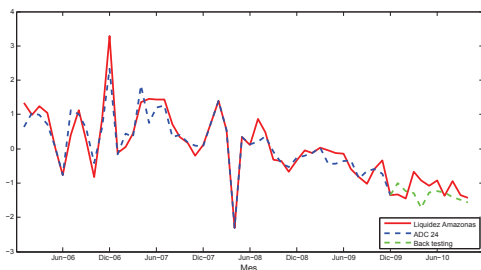


(b) Centros $c = 24$

Figura 23. Banco Amazonas, mínimos cuadrados recursivo combinado 100 iteraciones (MCRC) Gaussiana

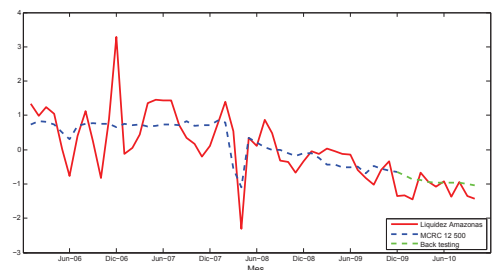


(a) Centros $c = 12$

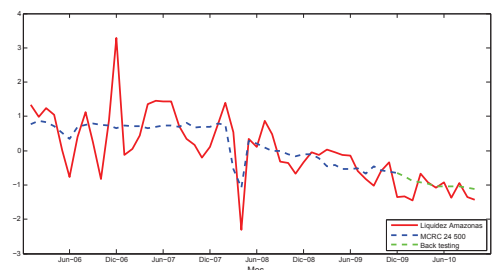


(b) Centros $c = 24$

Figura 22. Banco Amazonas, agrupamiento difuso combinado (ADC)



(a) Centros $c = 12$



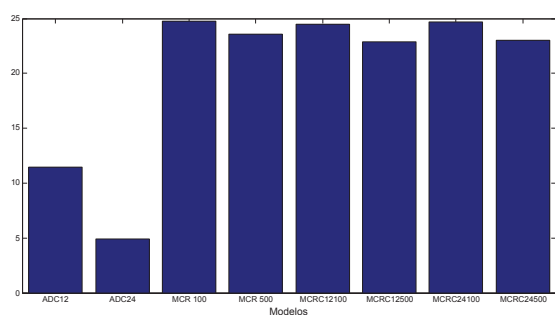
(b) Centros $c = 24$

Figura 24. Banco Amazonas, mínimos cuadrados recursivo combinado 500 iteraciones (MCRC) Gaussiana

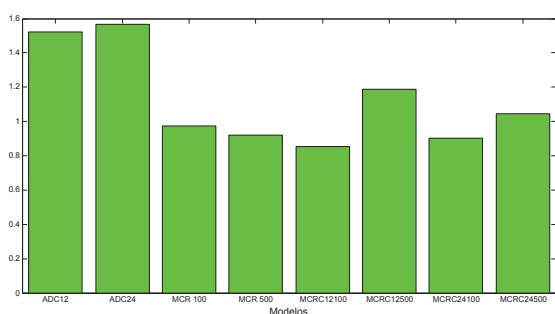
En el método de mínimos cuadrados recursivo combinado con 100 y 500 iteraciones (Figura 17 y Figura 18), se utilizaron 12 y 24 centros. Con 100 iteraciones, el ajuste del conjunto de entrenamiento es muy parecido para 12 y 24 centros; lo mismo ocurre cuando se calcula los pronósticos, el modelo logra ajustar adecuadamente la tendencia de la serie: ver Figura 17a. De la misma forma, cuando se aumenta a 500 iteraciones; no se observa ninguna mejora, Figura 18a.

La Figura 25 presenta para cada uno de los métodos, la suma de errores del modelo, y el *back testing*. Los mejores métodos, según estos criterios, son agrupamiento difuso combinado 24 (24 grupos) y mínimos cuadrados recursivo combinado 12-500 (12 centros, 500 iteraciones).

Al igual que en el caso anterior, se calcula un modelo SARIMA para el índice de liquidez del Banco Amazonas utilizando el mismo conjunto de entrenamiento, estos resultados muestran que el modelo más apropiado para su estimación y pronóstico es un SARIMA(2, 1, 1)(0, 0, 0), este modelo se compara con los mejores métodos automáticos, ver Figura 26. Se puede notar que el método de agrupamiento difuso combinado 24 y el modelo SARIMA son los mejores al momento de la estimación (residuos de los modelos); pero cuando se calculan los pronósticos (residuos del *back testing*), al igual que el índice de liquidez del Inter-nacional, el método mínimos cuadrados recursivo combinado 12-500 es más competitivo.



(a) Residuos de los modelos



(b) Residuos del *back testing*

Figura 25. Banco Amazonas, residuos de los métodos automáticos y *back testing*

4 Conclusiones

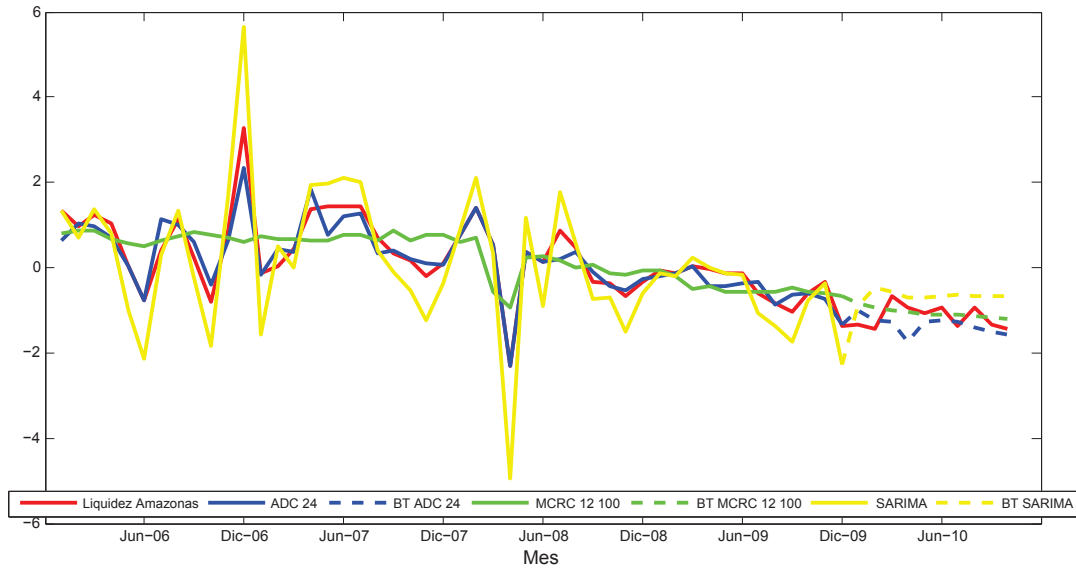
En este trabajo se describen los métodos automáticos de regresión difusa: mínimos cuadrados por lotes, mínimos cuadrados recursivo, aprendizaje desde el ejemplo modificado y agrupamiento difuso combinado. Se generalizan los métodos al caso de entrada y salida múltiple, que permite aplicarlos a una gama más amplia de problemas. Además, se propone un nuevo método, *mínimos cuadrados recursivo combinado*, el cual ha mostrado un gran desempeño, tal como se puede observar en las pruebas numéricas realizadas. En general, los resultados obtenidos corresponden a una medida cuantitativa del potencial de estos métodos en aplicaciones económicas, con resultados muy alentadores.

Como se visualiza en las pruebas numéricas, con un mayor número de observaciones (RILD e índice de liquidez), existe mayor eficiencia al caracterizar la tendencia y sobre todo el *back testing* en los métodos automáticos y, sobre todo, en los métodos combinados y el método propuesto. Esto se debe a que una mayor cantidad de información contribuye al mejoramiento en la construcción de reglas del sistema, lo que caracteriza de mejor manera al fenómeno de estudio. En general, el *back testing* del método propuesto proporciona estimaciones muy competitivas, comparadas con los modelos SARIMA y los métodos de la literatura.

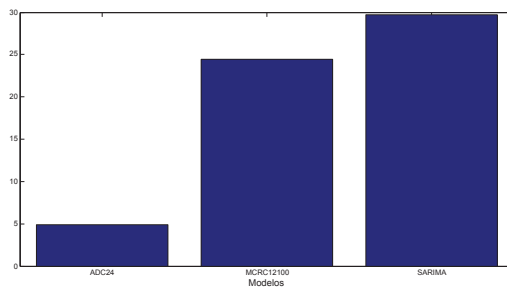
Finalmente, cabe señalar que la precisión requerida y el costo computacional son parámetros que se deben tener en cuenta al momento de decidir que método aplicar, *e.g.*, el método de mínimos cuadrados recursivo no requiere la inversión de una matriz y, por tanto, es más estable numéricamente. Sin embargo, su costo computacional es elevado comparado con el método de mínimos cuadrados por lotes. Por otra parte, la escala de los datos puede causar inestabilidad en los algoritmos. Por esta razón, en general es recomendable estandarizar los datos.

Referencias

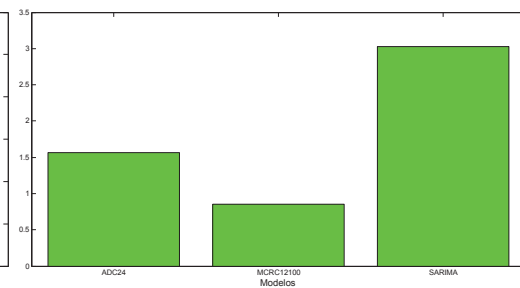
- [1] Banco Central del Ecuador <http://www.bce.fin.ec/>.
- [2] R. Cajamarca. *Regresión automática difusa aplicada a datos económicos*. Escuela Politécnica Nacional Tesis, (2011).
- [3] R. Cajamarca y H. Mena. *Métodos automáticos de regresión difusa - SCRIPT mard para R*. <http://www.math.epn.edu.ec/hmena/>, sección Software, (2011).
- [4] J. Bezdek, J. Keller and R. Krisnapuram. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Springer Science Business Media, (2005).
- [5] D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Applications*. Mathematics in Science and Engineering, Volume 144, (1980).



(a) Contraste



(b) Residuos de los modelos



(c) Residuos del back testing

Figura 26. Banco Amazonas, contraste

- [6] F. Hayashi. *Econometrics*. Princeton University Press, (2000).
- [7] M. Hayes. *Statistical Digital Signal Processing and Modeling*. John Wiley and Sons, Inc., (1996).
- [8] E. Laukonen and K. Passino. *Training Fuzzy Systems to Perform Estimation and Identification*. Engineering Applications Artificial Intelligence Vol. 8 No. 5 pp. 499-514. Elsevier Science Ltd. Printed in Great Britain, (1995).
- [9] S. Medina and G. Paniagua. *Fuzzy Inference Systems to Creditworthiness Analysis*. *Dyna*, Año 75, No. 154, pp. 215-229, (2006).
- [10] D. Montgomery and G. Runger. *Applied Statistics and Probability for Engineers*. Third Edition, John Wiley and Sons, Inc., (2003).
- [11] R. Nagamune. *Intelligent Control*. Department of Mechanical Engineering, University of British Columbia, (2008).
- [12] A. Palit and D. Popovic. *Computational Intelligence in Time Series Forecasting, Theory and Engineering Applications*. Springer-Verlag, (2005).
- [13] K. Passino and S. Yurkovich. *Fuzzy Control*. Addison Wesley Longman, (1998).
- [14] R Project for Statistical Computing, www.r-project.org/.
- [15] T. Ross. *Fuzzy logic with engineering applications*. Second Edition. Wiley, England, (2004).
- [16] Superintendencia de Bancos y Seguros. *Normas Generales para la Aplicación de la Ley General de Instituciones del Sistema Financiero*. Libro I, Título X, Capítulo IV, (2003).
- [17] R. Xu and D. Wunsch II. *Clustering*. John Wiley and Sons, Inc., (2009).
- [18] L. Zadeh. *Fuzzy Sets*. University of California, Berkeley, California, (1965).

