

Series de Tiempo Difusas aplicadas al pronóstico de la remuneración por la fabricación del calzado

Marisol Gutiérrez¹, Luis Ernesto Mancilla Espinoza¹, Alfonso Gutiérrez Lugo² y
Marco A. Gutiérrez²

¹ División de Estudios de Posgrado e Investigación, Instituto Tecnológico de León,
Av. Tecnológico S/N, 37290 Guanajuato, México
marisol.gutierrez@ieee.org, lmancilla01@hotmail.com
<http://posgrado.itleon.edu.mx/>

² Ingeniería Industrial, Instituto Tecnológico Superior de Lagos de Moreno
Libramiento Tecnológico No. 5000, 47480 Jalisco, México
<http://www.teclagos.edu.mx/>
Paper received on 04/10/12, Accepted on 25/10/12.

Resumen En el presente trabajo, se utiliza el método llamado Series de Tiempo Difusas en tiempo variante, para realizar el pronóstico de la remuneración por la fabricación del calzado (cuyos datos son obtenidos por medio de la Encuesta Mensual de la Industria Manufacturera, en el ciclo de Enero-2008 a Julio-2011), donde los conjuntos manejados son difusos y basado en la diferencia de segundo orden para obtener la tendencia. Al final los resultados obtenidos con el método de series de tiempo difusas son comparados con el método de regresión lineal, respecto al error obtenido por ambos.

Palabras claves: Conjuntos Difusos, Series de Tiempo Difusas, Pronóstico.

1. Introducción

La toma de decisiones depende del nivel de exactitud de la predicción que se desee (por lo cual es un proceso muy complejo, ya que interviene el futuro). En la vida diaria se desempeña el proceso de predicción muy cotidianamente, como por ejemplo el más popular el pronóstico del clima por mencionar alguno.

En algunos casos los modelos clásicos no pueden pronosticar (predecir el futuro) si los valores están representados en términos lingüísticos, es decir, en conjuntos difusos (es un conjunto que puede contener elementos de forma parcial [1] [2], es decir, en los conjuntos clásicos, no pertenece = 0 y pertenece = 1, pero los conjuntos difusos permiten grados de pertenencia entre 0 y 1) introducidos por Zadeh en 1965 [3]. Song y Chissom [4] [5] [6] presentan la teoría de Series de Tiempo Difusas para mejorar los inconvenientes que presentan los métodos de Series de Tiempo Clásicas.

La predicción con Series de Tiempo Difusas tiene una amplia aplicación la cual resulta muy importante, por ejemplo: planificación económica y de negocios, inventarios y control de producción, etcétera.

Song y Chissom pronostican el número de matrículas en la Universidad de Alabama, basados en la teoría de Series de Tiempo Difusas. Chen [7] continuó con la investigación de Song y Chissom, el cual redujo el tiempo y proceso de cálculo. Huang [8] usó la diferencia de matrículas, para predecir y posteriormente en el artículo [9] extiende el trabajo [7] con la adición de reglas heurísticas. Chen [10] utilizó un método para ajustar la longitud de cada intervalo con Algoritmos Genéticos. Jilani [11] presenta un método propuesto que pertenece al orden k -ésimo. Pathak y Singh [12] continúan con el pronóstico del número de matrículas en la Universidad de Alabama, introducen el concepto de intervalo de banda ancha de $(4/3)\sigma$ para el pronóstico. Jasim [13] pronostica el número de matrículas, en base al método de primer orden en tiempo variante.

Chen y Hwang [14] presentan un método para predecir la temperatura, basado en Series de Tiempo Difusas. Shah [15] pronostica el capital de la India con uso de Series de Tiempo Difusa, donde calcula la tendencia con la diferencia de segundo orden para el pronóstico. Ecerkale [16] empleó un periodo de tres años de los datos de producción de combustibles de aviación de Turquía como datos experimentales para el pronóstico.

Song [5] utilizó el siguiente modelo para el pronóstico de la matrícula universitaria:

$$A_i = A_{i-1} \circ R \quad (1)$$

Donde A_i , se refiere a las inscripciones difusas del año i , el símbolo \circ denota el operador de composición Max-Min, y R es la relación difusa formada por las inscripciones difusas, extraídas de las Series de Tiempo Difusas. Este método tiene algunas desventajas: requiere una gran cantidad de cálculos para obtener la relación difusa R . La operación de composición de máximos y mínimos de la ecuación (1) realiza una gran cantidad de cálculos, cuando la relación difusa R es muy grande.

Esta investigación, es estructurada como a continuación se muestra. En la sección 2, se explican los conceptos básicos necesarios para comprender las Series de Tiempo Difusas. En la sección 3, se presentan los objetivos de utilizar Series de Tiempo Difusas en esta investigación. En la sección 4, se describe detalladamente la metodología utilizada. En la sección 5, se muestran los resultados obtenidos con las Series de Tiempo Difusas y son comparados con los resultados obtenidos con Regresión Lineal por medio de sus errores. Finalmente en la sección 5, se exponen las conclusiones.

2. Conceptos básicos de Series de Tiempo Difusas

Song [5] propone la definición de series de tiempo difusas, basadas en conjuntos difusos [1] [2] [3].

Sea U el universo en discurso. $U = \{u_1, u_2, \dots, u_n\}$, y sea A el conjunto difuso del universo en discurso U definido como a continuación:

$$A = \mu_A(u_1)/u_1 + \mu_A(u_2)/u_2 + \dots + \mu_A(u_n)/u_n \quad (2)$$

Donde μ_A es la función de pertenencia de A , $\mu_A : U \rightarrow [0, 1]$, $\mu_A(u_i)$ indica el grado de pertenencia de u_i en el conjunto difuso A . $\mu_A(u_i) \in [0, 1]$ para $1 \leq i \leq n$.

$X(t)$ (para $t = \dots, 1, 2, \dots$) se encuentra dentro del universo en discurso y es un subconjunto de R , y el conjunto difuso $f_i(t)$ (para $i = 1, 2, \dots$) está definido en $X(t)$.

Sea $F(t)$ la colección de $f_i(t)$ (para $i = 1, 2, \dots$). Entonces $F(t)$ es llamada Serie de Tiempo Difusa de $X(t)$ (para $t = \dots, 1, 2, \dots$).

Si $F(t)$ es producida por $F(t-1)$, que se denota por $F(t-1) \rightarrow F(t)$, Esta relación puede representarse por $F(t) = F(t-1)^o R(t, t-1)$, donde el símbolo o denota el operador de composición Max-Min, $R(t, t-1)$ es la relación difusa que hay entre $F(t)$ y $F(t-1)$ y es llamado modelo de primer orden de $F(t)$.

Sea $F(t)$ una Serie de Tiempo Difusa y sea $R(t, t-1)$ la relación difusa de un modelo de primer orden de $F(t)$. Si $R(t, t-1)$ es igual a $R(t-1, t, 2)$ para cualquier tiempo t , entonces $F(t)$ se llama Serie de Tiempo Difusa en tiempo invariante. Si $R(r, t-1)$ es dependiente del tiempo t , es decir, $R(t, t-1)$ puede ser diferente de $R(t-1, t-2)$ para cualquier t , entonces $F(t)$ se llama Serie de Tiempo Difusa en tiempo variante.

Song [6] propuso un modelo de Series de Tiempo Difusas, para tiempo variable y pronosticar el número de matrículas de la Universidad de Alabama. A continuación se muestra una breve reseña del método [6]:

Paso 1: Definir el universo en discurso U donde están definidos los conjuntos difusos.

Paso 2: Dividir el universo en discurso U en intervalos de longitudes iguales.

Paso 3: Determinar los valores lingüísticos representados por los intervalos de los conjuntos difusos del universo en discurso.

Paso 4: Cambiar a conjuntos difusos los datos históricos de las matrículas.

Paso 5: Elegir un adecuado parámetro w , donde $w > 1$, para calcular $R^w(t, t-1)$ y el pronóstico de las matrículas como $F(t) = F(t-1)^o R^w(t, t-1)$.

Donde $F(t)$ denota el pronóstico difuso de la matrícula del año t , $F(t-1)$ denota el dato de la matrícula difusa del año $t-1$.

$$R^w(t, t-1) = F^T(t-2) \times F(t-1) \cup F^T(t-3) \times F(t-2) \cup \dots \cup F^T(t-w) \times F(t-w+1) \quad (3)$$

Donde w es llamado el modelo base que indica el número de años antes de t , \times es el operador de producto cartesiano, y T es el operador de transposición.

Paso 6: Obtiene el pronóstico de la matrícula por medio de Redes Neuronales.

3. Objetivo

Los modelos de series de tiempo clásicos no pueden predecir los períodos próximos, si las series de tiempo están expresadas lingüísticamente. El objetivo de la investigación es pronosticar la Remuneración por la fabricación del calzado, de los períodos siguientes dados los períodos anteriores.

La Serie de Tiempo de la Remuneración, se tomó del sitio oficial en Internet de PROSPECTA [17] a su vez obtenidos del Instituto Nacional de Estadística y Geografía (INEGI) por medio de Encuesta mensual de la industria manufacturera (EMIM), con una muestra de 603 empresas de Fabricación de calzado, en el ciclo de Enero-2008 a Julio-2011 (comprendido por 43 períodos cada uno con duración de un mes).

4. Metodología

A continuación se describe la metodología utilizada para el pronóstico por medio Series de Tiempo Difusas.

Este método está compuesto por 5 pasos principales, el primero paso, donde se define el universo en discurso y es dividido en intervalos de igual tamaño, el segundo paso, el universo es re-dividido en sub-intervalos de tamaños difusos según la distribución estadística de los datos, el tercer paso, se define para cada conjunto difuso A un valor lingüístico y un grado de pertenencia para cada intervalo del conjunto, el cuarto paso, se establece la relación lógica difusa entre los datos, el quinto paso, con el fin de predecir el dato del próximo periodo se utiliza una serie de reglas para encontrar si la tendencia es a la alza, la baja o tiene un ritmo constante, el método se basa en diferencias de segundo orden para encontrar el valor del próximo periodo.

PASO 1:

Primero se define el universo en discurso de acuerdo con los datos históricos (ver Tabla 1). Se encuentra el valor máximo (356,000) y mínimo (225,360), de los datos históricos, después se busca un número redondeado mayor que el máximo (360,000), así como un número redondeado menor que el mínimo (220,000). Con el objetivo de poder dividir el universo en n intervalos, para que los límites de dichos intervalos sean enteros. Entonces se tiene el universo que se denota con la letra U .

$$U = [220000, 360000] \quad (4)$$

Se divide el universo en discurso U (ecuación 4), en varios intervalos (n) de igual longitud u_1, u_2, \dots, u_n . Para este trabajo el universo se divide en siete intervalos ($n = 7$) de la misma longitud (tamaño de la longitud es 20,000). Por lo tanto los intervalos quedan de la siguiente manera: $u_1 = [220000, 240000]$, $u_2 = [240000, 260000]$, $u_3 = [260000, 280000]$, $u_4 = [280000, 300000]$, $u_5 = [300000, 320000]$, $u_6 = [320000, 340000]$ y $u_7 = [340000, 360000]$.

PASO 2:

Obtener la distribución estadística de los datos históricos en cada intervalo (ver Tabla 2). Se contabilizan todos los datos que se encuentran en cada intervalo, es decir, la frecuencia en la que cada intervalo se repite en los datos históricos (ver Fig .1).

Una vez que se tiene la distribución estadística de los datos, los intervalos obtenidos en el paso 1 se dividen nuevamente en sub-intervalos de tamaño difuso. El intervalo que tiene menor frecuencia de datos se divide solamente entre la unidad, así, sucesivamente se incrementa el valor en que se dividen los intervalos, hasta llegar al intervalo con mayor frecuencia de datos, se toma en cuenta que si dos intervalos tienen la misma frecuencia de datos ambos se dividen entre el mismo valor.

Para este trabajo se cuenta con siete datos distintos de frecuencias, es decir, empezamos el intervalo u_5 con menor frecuencia de datos solo es dividido entre la unidad (es decir, el intervalo queda igual), el intervalo que continua con menor frecuencia de datos

Tabla 1. Total de remuneraciones por la fabricación del calzado [17].

Mes/año	Moneda Nacional	Mes/año	Moneda Nacional
1/2008	256,330	1/2010	246,000
2/2008	260,250	2/2010	245,540
3/2008	267,510	3/2010	281,110
4/2008	265,020	4/2010	254,300
5/2008	265,190	5/2010	264,800
6/2008	247,560	6/2010	268,450
7/2008	273,690	7/2010	286,980
8/2008	258,710	8/2010	280,900
9/2008	255,750	9/2010	283,950
10/2008	282,930	10/2010	289,050
11/2008	261,890	11/2010	270,490
12/2008	324,880	12/2010	356,000
1/2009	226,700	1/2011	259,850
2/2009	225,360	2/2011	254,910
3/2009	241,450	3/2011	280,850
4/2009	236,900	4/2011	268,360
5/2009	233,780	5/2011	270,200
6/2009	230,200	6/2011	277,750
7/2009	258,900	7/2011	280,130
8/2009	243,320		
9/2009	250,650		
10/2009	273,390		
11/2009	252,600		
12/2009	324,930		

Tabla 2. Distribución de los datos históricos.

Intervalos	u_1	u_2	u_3	u_4	u_5	u_6	u_7
Frecuencia	5	14	13	8	0	2	1

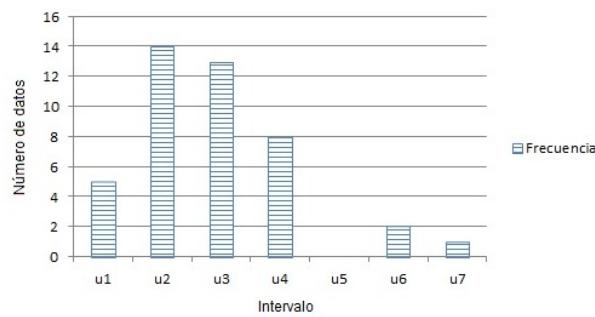


Figura 1. Histograma de la distribución de los datos históricos.

Tabla 3. Intervalos difusos.

$u_l = [valor_{l1}, valor_{l2}]$	
$u_{1,1} = [220000, 225000]$	$u_{3,4} = [270000, 273333.33]$
$u_{1,2} = [225000, 230000]$	$u_{3,5} = [273333.33, 276666.66]$
$u_{1,3} = [230000, 235000]$	$u_{3,6} = [276666.66, 280000]$
$u_{1,4} = [235000, 240000]$	$u_{4,1} = [280000, 284000]$
$u_{2,1} = [240000, 242857.14]$	$u_{4,2} = [284000, 288000]$
$u_{2,2} = [242857.14, 245714.28]$	$u_{4,3} = [288000, 292000]$
$u_{2,3} = [245714.28, 248571.42]$	$u_{4,4} = [292000, 296000]$
$u_{2,4} = [248571.42, 251428.57]$	$u_{4,5} = [296000, 300000]$
$u_{2,5} = [251428.57, 254285.71]$	$u_{5,1} = [300000, 320000]$
$u_{2,6} = [254285.71, 257142.85]$	$u_{6,1} = [320000, 326666.66]$
$u_{2,7} = [257142.85, 260000]$	$u_{6,2} = [326666.66, 333333.33]$
$u_{3,1} = [260000, 263333.33]$	$u_{6,3} = [333333.33, 340000]$
$u_{3,2} = [263333.33, 266666.66]$	$u_{7,1} = [340000, 350000]$
$u_{3,3} = [266666.66, 270000]$	$u_{7,2} = [350000, 360000]$

se divide entre dos sub-intervalos iguales, así sucesivamente, hasta llegar al intervalo u_2 con mayor frecuencia de datos es dividido en siete sub-intervalos iguales.

NOTA: Si algún intervalo no tiene ningún dato histórico, dicho intervalo puede ser eliminado ya que no tiene uso en el resto de la investigación.

PASO 3:

Se define cada conjunto difuso A_i basándose en el número total de sub-intervalos obtenidos en el paso anterior (28 sub-intervalos en total), se denota cada conjunto difuso A_i con un valor lingüístico, como se muestra a continuación:

$A_1 = \text{Muy}^{10} \text{ bajo}$, $A_2 = \text{Muy}^9 \text{ bajo}$, $A_3 = \text{Muy}^8 \text{ bajo}$, $A_4 = \text{Muy}^7 \text{ bajo}$, $A_5 = \text{Muy}^6 \text{ bajo}$, $A_6 = \text{Muy}^5 \text{ bajo}$, $A_7 = \text{Muy}^4 \text{ bajo}$, $A_8 = \text{Muy}^3 \text{ bajo}$, $A_9 = \text{Muy}^2 \text{ bajo}$, $A_{10} = \text{Muy bajo}$, $A_{11} = \text{Bajo}$, $A_{12} = \text{Por debajo de la media}$, $A_{13} = \text{Poco abajo de la media}$, $A_{14} = \text{Media}$, $A_{15} = \text{Poco arriba de la media}$, $A_{16} = \text{Por arriba de la media}$, $A_{17} = \text{Alto}$, $A_{18} = \text{Muy}^1 \text{ alto}$, $A_{19} = \text{Muy}^2 \text{ alto}$, $A_{20} = \text{Muy}^3 \text{ alto}$, $A_{21} = \text{Muy}^4 \text{ alto}$, $A_{22} = \text{Muy}^5 \text{ alto}$, $A_{23} = \text{Muy}^6 \text{ alto}$, $A_{24} = \text{Muy}^7 \text{ alto}$, $A_{25} = \text{Muy}^8 \text{ alto}$, $A_{26} = \text{Muy}^9 \text{ alto}$, $A_{27} = \text{Muy}^{10} \text{ alto}$ y $A_{28} = \text{Muy}^{11} \text{ alto}$.

Los conjuntos difusos A , tienen un grado de pertenencia para cada intervalo de longitud u_i . Por simplicidad el grado de pertenencia de los conjuntos difusos son 0, 0.5 y 1. El objetivo de representar los datos históricos (conjuntos clásicos), en datos difusos (conjuntos difusos), es obtener una serie de tiempo difusa. Para evitar acumular demasiados datos, si el grado de pertenencia es (0/u_i) se omitirá escribirlo.

$$\begin{aligned} A_1 &= 1/u_{1,1} + 0.5/u_{1,2} \\ A_2 &= 0.5/u_{1,1} + 1/u_{1,2} + 0.5/u_{1,3} \\ A_3 &= 0.5/u_{1,2} + 1/u_{1,3} + 0.5/u_{1,4} \\ A_4 &= 0.5/u_{1,3} + 1/u_{1,4} + 0.5/u_{2,1} \\ A_5 &= 0.5/u_{1,4} + 1/u_{2,1} + 0.5/u_{2,2} \\ A_6 &= 0.5/u_{2,1} + 1/u_{2,2} + 0.5/u_{2,3} \end{aligned}$$

$$\begin{aligned}
 A_7 &= 0,5/u_{2,2} + 1/u_{2,3} + 0,5/u_{2,4} \\
 A_8 &= 0,5/u_{2,3} + 1/u_{2,4} + 0,5/u_{2,5} \\
 A_9 &= 0,5/u_{2,4} + 1/u_{2,5} + 0,5/u_{2,6} \\
 A_{10} &= 0,5/u_{2,5} + 1/u_{2,6} + 0,5/u_{2,7} \\
 A_{11} &= 0,5/u_{2,6} + 1/u_{2,7} + 0,5/u_{3,1} \\
 A_{12} &= 0,5/u_{2,7} + 1/u_{3,1} + 0,5/u_{3,2} \\
 A_{13} &= 0,5/u_{3,1} + 1/u_{3,2} + 0,5/u_{3,3} \\
 A_{14} &= 0,5/u_{3,2} + 1/u_{3,3} + 0,5/u_{3,4} \\
 A_{15} &= 0,5/u_{3,3} + 1/u_{3,4} + 0,5/u_{3,5} \\
 A_{16} &= 0,5/u_{3,3} + 1/u_{3,4} + 0,5/u_{3,5} \\
 A_{17} &= 0,5/u_{3,4} + 1/u_{3,5} + 0,5/u_{3,6} \\
 A_{18} &= 0,5/u_{3,5} + 1/u_{3,6} + 0,5/u_{4,1} \\
 A_{19} &= 0,5/u_{3,6} + 1/u_{4,1} + 0,5/u_{4,2} \\
 A_{20} &= 0,5/u_{4,1} + 1/u_{4,2} + 0,5/u_{4,3} \\
 A_{21} &= 0,5/u_{4,2} + 1/u_{4,3} + 0,5/u_{4,4} \\
 A_{22} &= 0,5/u_{4,3} + 1/u_{4,4} + 0,5/u_{4,5} \\
 A_{23} &= 0,5/u_{4,4} + 1/u_{4,5} + 0,5/u_{5,1} \\
 A_{24} &= 0,5/u_{5,1} + 1/u_{6,1} + 0,5/u_{6,2} \\
 A_{25} &= 0,5/u_{6,1} + 1/u_{6,2} + 0,5/u_{6,3} \\
 A_{26} &= 0,5/u_{6,2} + 1/u_{6,3} + 0,5/u_{7,1} \\
 A_{27} &= 0,5/u_{6,3} + 1/u_{7,1} + 0,5/u_{7,2} \\
 A_{28} &= 0,5/u_{7,1} + 1/u_{7,2}
 \end{aligned}$$

PASO 4:

Establecer la relación difusa formada por los datos históricos, de la siguiente manera:

$$\begin{aligned}
 A_i &\rightarrow A_j \\
 A_j &\rightarrow A_k \\
 A_k &\rightarrow \dots \\
 \dots &\rightarrow A_m
 \end{aligned}$$

Donde A_i es el dato difuso del periodo $n - 1$ y A_j el dato difuso del periodo n , por lo tanto, la relación difusa es $A_i \rightarrow A_j$. El significado de esta relación lógica difusa este definida como: **SI** los datos difusos del periodo $n - 1$ es A_i **ENTONCES** para el periodo n sera A_j .

Se basa en los datos históricos de la tabla 1, para obtener la relación lógica difusa es la siguiente:

PASO 5:

Para averiguar si la tendencia es creciente, decreciente o casi lo mismo, se dividen los intervalos difusos (obtenidos en el paso 2), en cuatro sub-intervalos de igual longitud, donde el primer cuartil (o el punto 0.25 del intervalo) es tomado como el valor de

Tabla 4. Relación lógica difusa.

$r : A_i \rightarrow A_j$			
$r_1 : A_{10} \rightarrow A_{12}$	$r_{12} : A_{24} \rightarrow A_2$	$r_{23} : A_9 \rightarrow A_{24}$	$r_{34} : A_{20} \rightarrow A_{15}$
$r_2 : A_{12} \rightarrow A_{14}$	$r_{13} : A_2 \rightarrow A_2$	$r_{24} : A_{24} \rightarrow A_7$	$r_{35} : A_{15} \rightarrow A_{28}$
$r_3 : A_{14} \rightarrow A_{13}$	$r_{14} : A_2 \rightarrow A_5$	$r_{25} : A_7 \rightarrow A_6$	$r_{36} : A_{28} \rightarrow A_{11}$
$r_4 : A_{13} \rightarrow A_{13}$	$r_{15} : A_5 \rightarrow A_4$	$r_{26} : A_6 \rightarrow A_{18}$	$r_{37} : A_{11} \rightarrow A_{10}$
$r_5 : A_{13} \rightarrow A_7$	$r_{16} : A_4 \rightarrow A_3$	$r_{27} : A_{18} \rightarrow A_{10}$	$r_{38} : A_{10} \rightarrow A_{18}$
$r_6 : A_7 \rightarrow A_{16}$	$r_{17} : A_3 \rightarrow A_3$	$r_{28} : A_{10} \rightarrow A_{13}$	$r_{39} : A_{18} \rightarrow A_{14}$
$r_7 : A_{16} \rightarrow A_{11}$	$r_{18} : A_3 \rightarrow A_{11}$	$r_{29} : A_{13} \rightarrow A_{14}$	$r_{40} : A_{14} \rightarrow A_{15}$
$r_8 : A_{11} \rightarrow A_{10}$	$r_{19} : A_{11} \rightarrow A_6$	$r_{30} : A_{14} \rightarrow A_{19}$	$r_{41} : A_{15} \rightarrow A_{17}$
$r_9 : A_{10} \rightarrow A_{18}$	$r_{20} : A_6 \rightarrow A_8$	$r_{31} : A_{19} \rightarrow A_{18}$	$r_{42} : A_{17} \rightarrow A_{18}$
$r_{10} : A_{18} \rightarrow A_{12}$	$r_{21} : A_8 \rightarrow A_{16}$	$r_{32} : A_{18} \rightarrow A_{18}$	
$r_{11} : A_{12} \rightarrow A_{24}$	$r_{22} : A_{16} \rightarrow A_9$	$r_{33} : A_{18} \rightarrow A_{20}$	

tendencia a disminuir, y el tercer cuartil (o el punto 0.75 del intervalo) es tomado como el valor de tendencia al alza. Para predecir el valor del próximo periodo se utiliza el siguiente procedimiento:

Este procedimiento se basa en la diferencia de segundo orden. Donde la diferencia de segundo orden está dada por dS como a continuación se muestra:

$$dS = (S_{n-1} - S_{n-2}) - (S_{n-2} - S_{n-3}) \quad (5)$$

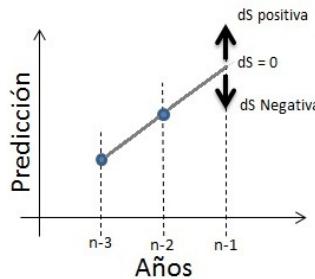


Figura 2. Si la proyección de una línea recta con los dos primeros datos ($n - 3$ y $n - 2$) es igual al valor de $n - 1$, se dice que la tendencia tiene un crecimiento constante ($dS = 0$), mientras que si es mayor (dS positivo), se dice que la tendencia es a la alza y de lo contrario, si es menor (dS negativo) la tendencia es a la baja.

Para predecir el valor futuro se aplicaran las siguientes 4 reglas:

Regla 1: Dado que se tiene una diferencia de segundo orden dS (ver ecuación 5), el cual es imposible calcular para los dos primeros datos históricos (ver Tabla 1), en el tercer dato, donde no está disponible el dato del periodo $n - 3$, pero si se encuentran disponibles los datos de los periodos $n - 2$ y $n - 1$, se utilizará:

Si $|S_{n-1} - S_{n-2}| >$ que la mitad de la longitud del intervalo correspondiente al conjunto difuso A_j , con valor de pertenencia igual a 1, entonces la tendencia del pronóstico es a la baja y el valor esperado sera el punto ubicado en $1/4$ (0.25) del intervalo.

Si $|S_{n-1} - S_{n-2}| <$ que la mitad de la longitud del intervalo correspondiente al conjunto difuso A_j , con valor de pertenencia igual a 1, entonces la tendencia del pronóstico es a la alza y sera el punto ubicado en $3/4$ (0.75) del intervalo.

Si $|S_{n-1} - S_{n-2}| =$ que la mitad de la longitud del intervalo correspondiente al conjunto difuso A_j , con valor de pertenencia igual a 1, entonces el valor esperado sera el punto medio (0.5) del intervalo.

Regla 2: Si la $S_{n-1} + (| (S_{n-1} - S_{n-2}) - (S_{n-2} - S_{n-3}) | \times 2)$ o $S_{n-1} - (| (S_{n-1} - S_{n-2}) - (S_{n-2} - S_{n-3}) | \times 2)$ cae dentro del intervalo correspondiente al conjunto difuso A_j , con valor de pertenencia igual a 1, entonces la tendencia es a la alza y el valor esperado sera el punto ubicado en $3/4$ (0.75) del intervalo.

Regla 3: Si no se cumple la regla 2, se continua con la siguiente condición. Si la $S_{n-1} + (| (S_{n-1} - S_{n-2}) - (S_{n-2} - S_{n-3}) | \div 2)$ o $S_{n-1} - (| (S_{n-1} - S_{n-2}) - (S_{n-2} - S_{n-3}) | \div 2)$ cae dentro del intervalo correspondiente al conjunto difuso A_j , con valor de pertenencia igual a 1, entonces la tendencia es a la baja y el valor esperado sera el punto ubicado en $1/4$ (0.25) del intervalo.

Regla 4: Si los datos no siguen la regla 2 o la regla 3, el pronóstico tomara el valor del punto medio (0.5) del intervalo correspondiente al conjunto difuso A_j , con valor de pertenencia igual a 1.

5. Resultados

Al usar la metodología mencionada anteriormente, la tendencia de la Serie de Tiempo se predice como se muestra en la tabla 5. El gráfico muestra la compración entre los datos históricos, regresión lineal (dada por la ecuación 6) y la predicción de la serie de tiempo difusa.

$$y_i = 626.22x_i + 252,212 \quad (6)$$

Para los periodos x_j y los pronósticos resultantes y_j . Donde $j = 1, 2, \dots$.

El error promedio se define como:

$$\text{Error promedio} = \frac{\left[\sum_{i=1}^m \left(\frac{|valor\ pronosticado_i - valor\ histríco_i|}{valor\ histríco_i} \right) \times 100 \right]}{m} \quad (7)$$

El error cuadrático medio (ECM) para comparar los resultados, el ECM es calculado de la siguiente manera:

Tabla 5. Resultados de la predicción con Series de Tiempo Difusa comparados con los datos históricos y la regresión lineal.

Mes/Año	Remuneración histórica	Tendencia	Predicción	Regresión lineal
1/2008	256,330			
2/2008	260,250			
3/2008	267,510	DOWN	267,500	254,090.66
4/2008	265,020	DOWN	264,166.66	254,726.88
5/2008	265,190	MIDDLE	265,000	255,343.1
6/2008	247,560	MIDDLE	247,142.85	255,969.32
7/2008	273,690	MIDDLE	275,000	256,595.54
8/2008	258,710	MIDDLE	258,571.42	257,221.76
9/2008	255,750	MIDDLE	255,714.28	257,847.98
10/2008	282,930	MIDDLE	282,000	258,474.2
11/2008	261,890	MIDDLE	261,666.66	259,100.42
12/2008	324,880	MIDDLE	323,333.33	259,726.64
1/2009	226,700	MIDDLE	227,500	260,352.86
2/2009	225,360	MIDDLE	227,500	260,979.08
3/2009	241,450	MIDDLE	241,428.57	261,605.3
4/2009	236,900	MIDDLE	237,500	262,231.52
5/2009	233,780	MIDDLE	232,500	262,857.74
6/2009	230,200	UP	233,750	263,483.96
7/2009	258,900	MIDDLE	258,571.42	264,110.18
8/2009	243,320	MIDDLE	244,285.71	264,736.4
9/2009	250,650	MIDDLE	250,000	265,362.62
10/2009	273,390	MIDDLE	275,000	265,988.84
11/2009	252,600	MIDDLE	252,857.14	266,615.06
12/2009	324,930	MIDDLE	323,333.33	267,241.28
1/2010	246,000	MIDDLE	247,142.85	267,867.5
2/2010	245,540	MIDDLE	244,285.71	268,493.72
3/2010	281,110	MIDDLE	282,000	269,119.94
4/2010	254,300	MIDDLE	255,714.28	269,746.16
5/2010	264,800	MIDDLE	265,000	270,372.38
6/2010	268,450	MIDDLE	268,333.33	270,998.6
7/2010	286,980	MIDDLE	286,000	271,624.82
8/2010	280,900	MIDDLE	282,000	272,251.04
9/2010	283,950	DOWN	282,000	272,877.26
10/2010	289,050	MIDDLE	289,000	273,503.48
11/2010	270,490	MIDDLE	271,666.66	274,129.7
12/2010	356,000	MIDDLE	355,000	274,755.92
1/2011	259,850	MIDDLE	258,571.42	275,382.14
2/2011	254,910	MIDDLE	255,714.28	276,008.36
3/2011	280,850	MIDDLE	282,000	276,634.58
4/2011	268,360	MIDDLE	268,333.33	277,260.8
5/2011	270,200	MIDDLE	271,666.66	277,887.02
6/2011	277,750	DOWN	277,500	278,513.24
7/2011	280,130	DOWN	281,000	279,139.46
Error promedio (ecuación 7)		0.33	6.55	
Error cuadrático medio (ecuación 8)		1,138.39	24,616.69	

$$ECM = \sqrt{\frac{\sum_{i=1}^m (\text{valor histríco}_i - \text{valor promosticado}_i)^2}{m}} \quad (8)$$

Donde m es el número de datos pronosticados.

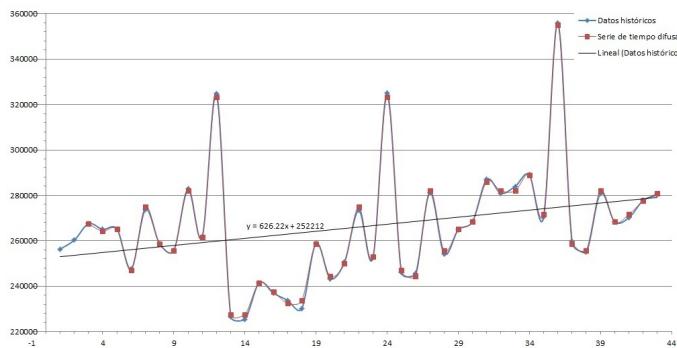


Figura 3. Predicción con Series de Tiempo Difusas.

6. Conclusiones

El uso de técnicas difusas, proporciona mejoras para la predicción de series de tiempo, ya que se puede trabajar el lenguaje natural (o lingüístico), por medio de conjuntos difusos.

Se puede observar que la Serie de Tiempo Difusa, tiene mejores resultados para el pronóstico de la remuneración por la fabricación del calzado, que la Regresión Lineal, ya que ambos errores son inferiores. También se puede observar en la gráfica, la Serie de Tiempo Difusa tiene una mayor aproximación a los datos históricos que la Regresión Lineal.

Por lo tanto, se concluye que la técnica de Series de Tiempo Difusas es capaz de predecir la remuneración por la fabricación del calzado, para los próximos periodo, así como identificar la tendencia.

Agradecimientos Al Consejo Nacional de Ciencia y Tecnología (CONACYT).

Referencias

1. Klir, G.J., Yuan, B.: *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall, New Jersey (1995).
2. Kaufmann, A.: *Introducción a la teoría de los subconjuntos borrosos. Para el uso de ingenieros*. Continental, México (1982).
3. Zadeh, L.A.: *Fuzzy Sets*. *Information and Control* 8, 338–353 (1965).

4. Song, Q.: Fuzzy Time Series and Its Models. *Fuzzy Sets ans System* 54, 269–277 (1993).
5. Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series - Part I. *Fuzzy Sets and System* 54, 1–9 (1993).
6. Song, Q., Chissom, B.S.: Forecasting enrollments with fuzzy time series - Part II. *Fuzzy Sets and System* 62, 1–8 (1994).
7. Chen, S.M., Hsu, C.C.: A New Method to Forecast Enrollments Using Fuzzy Time Series. *Inter. Journal of Applied Science and Engineering* 2004. 2, 3: 234–244 (2004).
8. Hwang, J.R., Chen, S.M., Lee, C.H.: Handling forecasting problems using fuzzy time series. Elsevier 100, 217–228 (1998).
9. Huarng, K.: Heuristic models of fuzzy time series for forecasting. *Fuzzy Sets and Systems* 123, 369–386 (2001).
10. Chen, S.M., Chung, N.Y.: Forecasting enrollments of students by using fuzzy time series and genetic algorithms. *International journal of information and management sciences* 17, 1–17 (2006).
11. Jilani, T.A., Burney, S.M.A., Arbil, C.: Fuzzy metric approach for fuzzy time series forecasting based on frequency density based partitioning. *Proceedings of World Academy of Science, Engineering and Technology* 23, 333–338 (2007).
12. Pathak, H.K., Singh, P.: A New Bandwidth Interval Based Forecasting Method for Enrollment Using Fuzzy Time Series. *Applied Mathematics* 2011. 2, 504–507 (2011).
13. Jasim, H.T., Salim, A.G.J., Ibraheem, K.I.: A Novel Algorithm to Forecast Enrollment Based on Fuzzy Time Series. *Applications and Applied Matematics: An international Journal (AAM)* 7, 385–395 (2012).
14. Chen, S.M., Hwang, J.R.: Temperature prediction using fuzzy time series. *IEEE Trans. on Systems, Man, and Cybernetics-Part B: Cybernetics* 30, 263–275 (2000).
15. Shah, M.: Fuzzy Time Series: A Realistic Method to Forecast Gross Domestic Capital of India. *Analysis and Design of Intelligent Systems using Soft Computing Techniques*, 255–264 (2007).
16. Ecerkale, K., Küçükdeniz, T., Esnaf, S.: Comparison of fuzzy time series based on difference parameters and two-factor time-variant fuzzy time series models for aviation fuel production forecasting. *Journal of aeronautics and space technologies* 4, 57–63 (2010).
17. PROSPECTA, Centro De Innovación y Competitividad, <http://www.prospecta.org.mx>