



193315

**LA LOGICA DIFUSA COMO HERRAMIENTA EN LOS
PRONOSTICOS.**

TESIS QUE PARA OPTAR EL GRADO DE
MAESTRO EN CIENCIAS COMPUTACIONALES
PRESENTA

HECTOR GUEVARA ISLAS

Asesor: Dr. NEIL HERNÁNDEZ GRESS

Jurado: Dr Edgar Emmanuel Vallejo Clemente
Dr. Neil Hernandez Gress
Maestro. Eduardo A. Salas Rivas

Presidente
Secretario
Vocal

Atizapán de Zaragoza, Edo. Méx., Julio de 2003.

INDICE DE CONTENIDO

	Página
RESUMEN	3
INTRODUCCIÓN	4
1. LOS PRONOSTICOS	
1.1 La necesidad de pronosticar	7
1.1.1 Pasos a seguir en el pronóstico	7
1.2 Pronósticos tradicionales vs pronósticos modernos	11
1.3 Métodos tradicionales para pronosticar	13
1.3.1 Métodos de Promedio	13
1.3.2 Métodos de Atenuación	14
1.3.3 Regresión	16
1.3.8 Transformación.	17
1.4 Selección de Variables a través de La Técnica de Componentes Principales	18
2. LOGICA DIFUSA	
2.1 Lógica Difusa	20
2.2 Conjunto borroso	21
2.3 Funciones de Membresía.	24
2.4 Reglas Difusas	25
2.4.1 Fusificación	26
2.4.2 Evaluación de Reglas.	27
2.4.3 Defusificación de Salidas.	28
2.5 Modelo difuso tipo Sugeno.	30
2.6 ANFIS	31
2.6.1. Redes adaptables	31
2.6.2 Sistemas neurodifusos.	32
.3. DESARROLLO DEL SISTEMA DIFUSO.	
3.1 Conocimiento de las empresas que emiten las acciones	37
3.2 Sistema Difuso	38
3.2.1 Base de Datos	41
3.2 .2 Selección de la información	42
3.2.2.1 Modelo Correlación	42

3.2.2.2 Modelo ACP	47
3.2.3 ANFIS	49
3.2.4 Evaluación Difusa	52
3.2.5 Comparación del sistema Difuso Vs Real	54
4. RESULTADOS.	56
4.1 Pronóstico por Método Tradicional.	58
4.1.1 Pronóstico Telmex por método Tradicional	60
4.1.2 Pronóstico Cemex por método Tradicional	62
4.2 Pronóstico Sistema Difuso	63
4.2.1 Pronóstico Telmex por modelo Correlación	64
4.2.2 Pronóstico Telmex por modelo ACP	66
4.2.3 Pronóstico Cemex por modelo Correlación	69
4.2.4 Pronóstico Cemex por modelo ACP	71
4.3 Pronóstico Tradicional Vs Difuso	73
4.3.1 Pronóstico Telmex	74
4.3.2 Pronóstico Cemex	76
5. CONCLUSIONES.	80
5.1 Trabajo Futuro.	82
APÉNDICES	
Apéndice A..	83
Apéndice B.	84
Apéndice C	85
ANEXOS	
Anexo 1	86
BIBLIOGRAFÍA.	87

RESUMEN

Se desarrolló un sistema experto para realizar pronósticos en el espectro bursátil, basado en la lógica difusa, utilizando ANFIS como método de aprendizaje, teniendo una selección de entradas multivariantes por dos métodos (ACP, CORRELACION) y una salida monovariante (pronóstico deseado), la aplicación que se propuso en este trabajo es en un problema del área financiera, pero la generalidad del sistema construido permite que fácilmente puedan ocuparse en otros campos de aplicación diferentes. Los sistemas expertos basados en lógica difusa van minimizando su error en base al aprendizaje, selección de información, nuevas reglas basadas en la experiencia.

Se obtuvieron resultados muy aproximados a los reales con márgenes de error muy pequeños, y comparados con métodos tradicionales resulta una herramienta muy efectiva para realizar pronósticos en plataformas altamente no lineales. Este trabajo de tesis es una opción para cualquier financiero que desee realizar un pronóstico.

Se desarrolló dos tipos de selección de información Correlación y ACP, resultando más exacto el modelo ACP. Los resultados indicaron que el error cuadrático medio del sistema que desarrollamos es menor a un método tradicional, y lo mejor de todo es que este error cuadrático medio no es estático, es decir mediante la afinación del sistema difuso podemos disminuirlo.

Lo más interesante de este sistema difuso es su proceso de aprendizaje para minimizar el error, en base a la información histórica. Este tipo de sistema experto basado en la lógica difusa es el futuro en los pronósticos debido a su funcionalidad en plataformas no lineales.

INTRODUCCION.

La idea de tener el control en los datos estadísticos surge como un imposible, si alguna vez nos preguntamos ¿será posible tener la predicción exacta; cuándo lloverá y durante qué lapso de tiempo?, o aún más difícil ¿será posible saber o tener la predicción exacta si va a ocurrir un terremoto y de qué magnitud?, todas estas opciones solamente serían como resolver un imposible.

No importa si se trata de un negocio pequeño, una empresa sólida o simplemente una toma de decisión, el pronóstico llega a ser tan importante como el pensar en vanguardia, obviamente el pronóstico es una de las herramientas mas poderosas e innovadoras en la actualidad. El pronóstico se ha vuelto la parte comercial mas importante para poder competir en todos los sectores. En la actualidad se han desarrollado poderosas herramientas para hacer el pronóstico, basadas en matemáticas mediante un sistema computacional. Los pronósticos se llevan acabo en cualquier área de estudio, pero sobre todo en el área de mercado bursátil en donde resulta esencial el tener un pronóstico con el menor grado de error para maximizar los bienes o posesiones del financiero. Las casas de bolsa han financiado diversos sistemas para el desarrollo de pronósticos, pero la realidad es que no existe un sistema de pronóstico estándar, es decir que una casa de bolsa solamente ocupe un sistema para realizar sus pronósticos bursátiles, al contrario se utilizan diversos sistemas para realizar el pronóstico. El desarrollo de pronósticos bursátiles es una de las áreas de mayor investigación y desarrollo. Mientras que el espectro financiero crece, también se necesita de sistemas de pronósticos con menor grado de error [1].

La lógica difusa comúnmente conocida en el mundo científico como 'FUZZY LOGIC' ha resuelto un sin número de aplicaciones que parecían imposibles de desarrollar. La lógica difusa parte de un hecho muy sencillo, es actuar como nuestro cerebro lo puede resolver, existe la diferencia que nuestro cerebro no está acostumbrado a resolver mas de tres entradas (variables) al mismo tiempo, mientras que un sistema experto basado en lógica difusa lo podría resolver sin ningún problema. La lógica difusa puede realizar un pronóstico de manera diferente a lo que estamos acostumbrados pero sin lugar a duda con excelente grado de certeza, es una de las mejores herramientas para manejar sistemas inestables, como podría ser el mercado de acciones.

Actualmente existe la necesidad de emigrar conocimiento de unas ciencias a otras, es decir existe una conectividad de un TODO, por lo que a pesar de que lógica difusa fue aplicada en principio a los sistemas de control por medio de la electrónica existe una analogía con el espectro bursátil. Quizá lo mas seguro es que cada persona interesada en pronosticar tiene su propio método para hacerlo, no porque la estadística o la matemática no funcione al contrario hay que saber confiar en ésta, la respuesta está en el saber detectar y combinar métodos para pronosticar, la lógica difusa es el futuro de los sistemas de control moderno en los pronósticos[13].

Este trabajo resulta de gran interés para los financieros o cualquier tipo de actividad que necesite tener un pronóstico esperado, utilizando la lógica difusa[5].

Algo que surge muy interesante es que el financiero experto va construyendo su modelo de pronóstico solamente en la mente y en base a la experiencia adquirida en el campo, de esta manera resulta muy difícil poder heredar ese conocimiento o transmitirlo a sus colegas, en cambio los sistemas expertos se van alimentando en función al conocimiento de uno o varios expertos en la materia. Para alimentar un sistema experto se requiere: conocimientos matemáticos (estadísticos, regresiones, series de tiempo, transformadas, etc.) así como también acciones adquiridas con la experiencia del experto (intuiciones), es decir combinan estas dos partes y de esta manera se puede heredar este conocimiento que algunos financieros llaman intuición.

La lógica difusa ocupa todas estas acciones de una manera sencilla y han sido comprobadas experimentalmente con éxito por lo que se busca en este trabajo de investigación, dar una herramienta sencilla, efectiva, y eficaz al financiero así como dinámica, es decir el seguir actualizando para ir minimizando el error de aproximación.

En el mundo financiero la teoría y la realidad resultan en muchos casos incompatibles, es mas difícil analizar o utilizar un método de pronóstico que una simple intuición de una persona experta en el tema. En el espectro bursátil existen muchas variables que están relacionadas con la variable a pronosticar, algunas variables ni siquiera se tienen cuantificadas o se tiene un problema con la estandarización, nuestro trabajo pretende dar una predicción de datos en el espectro bursátil, basado en los métodos modernos y utilizando la lógica difusa, puesto que actualmente es uno de los métodos mas usados para resolver sistemas altamente no lineales[7].

Nuestro objetivo general en este trabajo: Desarrollar un sistema experto para el uso de pronósticos en el espectro del mercado bursátil, basado en la lógica difusa teniendo de entradas multivariantes y salida monovariante.

Nuestro objetivo específico en este trabajo: Partiendo de nuestro objetivo general, utilizaremos ANFIS como método de aprendizaje de lógica difusa, nuestro sistema tendrá una variable independiente que en nuestro caso específico será una acción, y tendremos entradas dependientes como pueden ser el complemento de acciones e indicadores; se realizará una selección de información mediante análisis de componentes principales, la predicción de estos datos no estará dada en tiempo real.

Nuestra hipótesis : Los métodos de pronóstico modernos basados en la lógica difusa responden con mayor precisión que los métodos tradicionales en sistemas altamente no lineales.

Nuestra hipótesis alterna: Los métodos de pronóstico modernos basados en la lógica difusa no responden con mayor precisión que los métodos tradicionales en sistemas altamente no lineales.

Partiendo de esto cabe mencionar que un sistema experto requiere de tiempo y retroalimentación para llegar a minimizar el error, con lo que ratificando nuestras hipótesis resulta efectivo que nuestro sistema experto empiece a funcionar desde el principio tratando de minimizar errores en las salidas.

Actualmente es necesario que los administradores que toman decisiones bursátiles sepan usar estos sistemas expertos o saber en base a qué se toma la decisión, las organizaciones necesitan de éstos sistemas expertos para estar a la vanguardia[12].

1.- LOS PRONOSTICOS.

1.1 La necesidad de pronosticar.

Los pronósticos han existido durante toda la vida del ser humano, desde tiempos ancestrales hasta nuestros días; el hombre siempre se ha interesado de alguna manera en poder predecir el futuro mediante datos del pasado, desde tiempos remotos donde el hombre necesitaba predecir las lluvias, cosechas, caza, temblores, sequías, etc., hoy en día se sigue necesitando esas variables y muchas otras, el hombre moderno tiene que resolver día a día el pronóstico, para el éxito en sus actividades laborales. Realmente toda actividad necesita un pronóstico en algún sector de la organización e inclusive todos necesitamos hacer pronósticos en nuestra vida, es incierto el decir que los pronósticos lo necesitan sólo los matemáticos o la gente con trabajos relacionados con la ingeniería, de alguna manera todos los que desarrollamos una actividad hemos encontrado la forma fácil y mas exacta de tener nuestros pronósticos. Hay que recordar que siempre que utilizamos un pronóstico tenemos la posibilidad de acertar o alejarnos del objetivo, obviamente el decir acertar es dependiendo nuestro grado de certeza que tengamos estandarizado. Una de las áreas con mayor necesidad de pronóstico sin lugar a duda es el mundo financiero, en donde se han desarrollado métodos modernos para poder pronosticar con cierto grado de certeza.

1.1.1. Pasos a seguir en el pronóstico

Todos los procedimientos formales de pronósticos comprenden la extensión de las experiencias del pasado al futuro incierto. Tenemos cuatro pasos generales[8]:

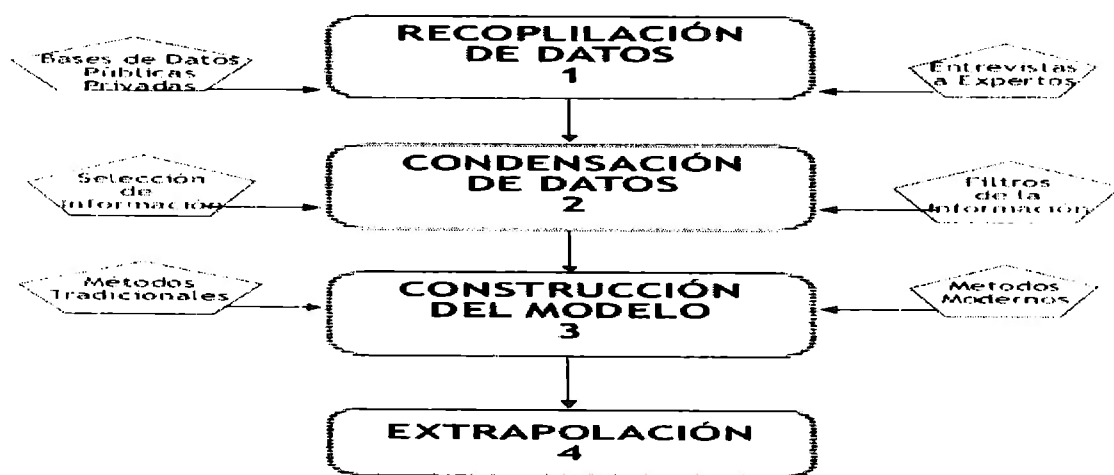


Figura 1.1 Pasos Generales del Pronóstico

El paso 1, es la recopilación de datos; dentro de este paso es muy importante obtener la información adecuada y sobre todo que no tengan valores alterados, es decir sean verídica; se tiene que consultar bases de datos, las cuales pueden ser privadas o públicas; generalmente las bases de datos privadas tienen un costo por obtener la información. Las bases de datos en el espectro financiero se pueden consultar en websites que tienen costos mensuales como “Finsat” que es el servicio electrónico de información especializado en finanzas, economía y negocios, también la obtención de información es a través de expertos en el tema; definimos al experto en el tema como la persona que ha trabajado en el área de interés con éxito y es capaz de tomar decisiones en tiempos cortos en base a su conocimiento o experiencia.

Es muy importante después de obtener la información adecuada poderla estandarizar así como homologar. En muchos casos se encuentran los valores de una variable en otro tiempo o espacio, es decir podemos tener acciones cotizadas al cierre del día y otras cotizadas al máximo precio del día, es aquí donde tendremos que tomar la decisión para estandarizar todas las variables.

El paso 2, la condensación de datos, es en este paso donde se realiza la selección de la información, buscamos tener la información necesaria de toda la información obtenida del paso anterior. Con lo cual se realiza el filtrado adecuado de la información, es obvio que si se trabaja con la información necesaria se puede manejar mejor cualquier tipo de sistema predictorio[8].

El paso 3, la construcción del modelo, implica el ajustar los datos reunidos en un modelo de pronósticos que sea el adecuado para minimizar el error en el pronóstico. Antes de decidir el modelo se tiene que tener presente la tendencia del sistema es decir el crecimiento o disminución en la serie sobre un periodo amplio así como también si existe alguna fluctuación en forma de onda alrededor de la tendencia, o si existe un patrón de cambio que se repita a través del tiempo, es en este paso donde se puede tomar la decisión de usar un método tradicional o un método moderno, definiendo al método tradicional como los procesos matemáticos ya sean regresiones, series de tiempo, estadística, transformaciones, etc, y los métodos modernos como: lógica difusa, redes neuronales, etc.

El paso 4 consiste en la extrapolación en sí del modelo de pronóstico, lo cual ocurre una vez que se recolectaron y tal vez redujeron, los datos adecuados y que se prueba el modelo del pronóstico a estimar, en este paso donde se verifica la exactitud del sistema de pronóstico, comparado con algunos de los modelos establecidos como son: tradicionales, modernos, Box-Jenkins, Delphi,

etc. Quizá lo mas indicado en este paso es analizar un error cuadrático medio por tramos en el pronóstico, para ir detectando en la parte que realmente el método sirve y en la parte que se aleja del real[13].

Los principales paquetes de computo que nos sirven como herramientas para realizar el pronóstico son los siguientes:

- AutoBox3.0 de Automatic Forecasting Systems, Inc.
- Autocast II de Levenbach Associates
- Business and Economic Forecasting: Decision Support System Software de Jonh Wiley & Sons
- Easy Forecaster Plus I o II del Institute of Business Forecasting
- Economics Software Program (ESP) de Chase Econometric.
- Express Easycast de Management Decisions Systems, Inc.
- ET (Econometrics Toolkit) de Econometrics Software, Inc.
- Forecast Master de Scientific System, Inc.
- 4 Cats de Scientific Systems, Inc.
- Forecasting de Hewlett Packard.
- Forecast Pro de Business Forecast Systems, Inc.
- Micranal de Gwilym Jenkins & Partners, Ltd.
- Micro-BJ de Stratix.
- MICCROFIT3 de Oxford University Press
- MTS (the Multiple Time Series Program) de Automatics Forecatsing Systems
- Pro*Cast de Fleming Software
- Forecast de Computer Software Consultants, Inc.
- RATS de Var Econometrics
- Smartforecasts II de Smart Software, Inc.
- SORITEC de sorites Group.
- Timestat DE Timestat, Inc.
- SCA Statistical System de Scientific Computing Associates Corp.
- The Forecaster de Wadsworth Electronic Publishing Company
- Ystat de Ming Telecomputing, Inc.
- Matlab

- ForeCalc de Business Forecast Systems
- Forecast! De index Solutions, Inc.
- Forecast/DSS y Lotus-FORMAN, ambos de John Willey & Sons, Inc
- Add a Stat y Multifin ambos de Scientific Software company
- Stat-packets de Statpac, Inc
- Spreadsheet Forecaster de LevenBach Associates
- Minitab
- Micro TSP
- Statpro
- Statpac
- Daisy
- Decision Support
- Modeling
- Visitrend/Plot
- Graph N
- Expert Choice
- Nuametrics
- Systat
- Xpolaitor
- X11ARIMA/88
- Sibyl Runner
- FuturCast
- The Total Forecasting System.

Matlab es uno de los paquetes mas completos puesto que se puede aplicar procesos de control así como también ofrece la simulación del sistema, aplicando herramientas matemáticas como métodos numéricos[8], transformadas, sistemas de ecuaciones diferenciales, desde luego en esta clase de paquetes uno puede aplicar métodos tipo ACP (Análisis de componentes principales), así como transformadas de Fourier; como herramientas para nuestro pronóstico, también ofrece un sistema de librerías especializadas en muchas áreas por ejemplo; series financieras, lógica difusa, redes neuronales, pronósticos, etc.

1.2 Pronósticos tradicionales vs pronósticos modernos

Los métodos de pronósticos tradicionales son aquellos que se basan en estadística, parametrización, estimación; como por ejemplo:, series de tiempo, regresiones, transformación, ARIMA, etc. Realmente los métodos tradicionales de pronósticos resultan a veces muy ajenos a la realidad, en muchos casos son poco flexibles al mundo de predicciones es decir, no es que no sirvan al contrario es matemática, lo que resulta es que no es posible minimizar el error ante una inestabilidad en el sistema, resulta bastante difícil el hacer el pronóstico en plataformas altamente no lineales. Cuando se especializa en salida monovariable, obviamente cada vez que sus entradas de variables dependientes al sistema han aumentado y por ende los parámetros asociados, los métodos tradicionales pueden llegar a ser tan complejos de manejar que se pierde la finalidad del pronóstico.

Los métodos de pronósticos modernos usan inteligencia artificial para llegar al éxito del pronóstico, aunque también se pueden alimentar sus reglas o aprendizajes por medio de los métodos tradicionales, pero no dejando que éstos últimos resuelvan todo el sistema[13] . Este tipo de métodos utilizan como puede ser redes neuronales, algoritmos genéticos, lógica difusa, data mining, etc., de alguna manera estos métodos simulan el pensamiento humano o la naturaleza de reproducción, el aprendizaje, etc.; han tenido éxito en la solución de sistemas inestables en donde es difícil encontrar un ecuación exacta, éstos han tenido éxito en donde la matemática ha sido difícil de resolver (ecuaciones diferenciales, curvas no suavizadas, etc) todo esto gracias al estudio del comportamiento humano, siendo ventaja para éstos métodos los algoritmos computacionales, combinaciones y rapidez de respuesta.

Para realizar pronósticos en el mercado bursátil tenemos que considerar lo siguiente: en mediano o largo plazo, la circulación de los precios de una emisora ó acción y los movimientos de los índices bursátiles no depende de algo fortuito o desordenado, sino que esta formado con raciocinio propio, de tal forma que podemos aplicar los métodos tradicionales o los métodos modernos para predecir lo que vendrá. Son varias las condiciones que se deben llevar a cabo para que el mercado bursátil se pueda comprender, se conozca lo que esta pasando y se pueda predecir lo que sucederá en base al pronóstico, y son las siguientes:

- Que los grupos humanos que se consideren sean grandes aunque no hay un número de individuos que sea suficiente para poder entender lo que se genere de sus decisiones, pero entre menor sea el grupo es menos confiable las decisiones y actitudes que se generen.
- Que la ejecución del problema esté sujeto al libre mercado
- Que las participaciones individuales no sean muy notorias en el resultado final.

Teniendo en cuenta estas condiciones se espera que el fenómeno generado se pueda estudiar bajo el paradigma del pronóstico moderno. Si no se llevan a cabo los movimientos de los precios no serán claros ni entendibles bajo estas proposiciones. En los métodos tradicionales como en los modernos se tiene que saber de qué manera están relacionadas las variables y esto se calcula con el coeficiente de correlación para una población significa el grado en que se relacionan dos variables, esto va de -1 a 1 obviamente el cero es cuando no tienen relación, -1 correlación negativa significaría que mientras una aumenta la otra disminuye, 1 significa igual[2].

$$\rho = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{N \sum X^2 - (\sum X)^2} \sqrt{N \sum Y^2 - (\sum Y)^2}} \quad (1.1)$$

No existe un método único que se pueda llevar a cabo para hacer un pronóstico, resulta que depende de la visión que tenga el experto para hacer el pronóstico, de la forma en que pueda dividir su curva de salida para obtener técnicas de pronósticos parciales, es decir combinar las técnicas de pronósticos, el experto tiene que tener la perspicacia de poder innovar, existen pronósticos los cuales usan herramientas matemáticas bastantes sofisticadas, pero no hay que tomar en cuenta el Costo-Beneficio.

1.3 Métodos tradicionales para pronosticar.

1.3.1 Métodos de Promedio

Este Tipo de Técnicas utilizan una forma de promedio ponderado de observaciones anteriores para atenuar fluctuaciones de corto plazo[8].

Método	Descripción	Pronóstico
Promedio Simple	Se obtiene encontrando la media de todos los valores aplicables y empleando después esa media para pronosticar el siguiente periodo, quizá es el más sencillo.	$\hat{Y}_{t+1} = \frac{\sum_{i=1}^n Y_i}{n} \quad (1.3)$
Promedio Móvil	Se obtiene encontrando la media de un conjunto específico de valores y usándola después para pronosticar el siguiente periodo.	$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + Y_{t-2} + \dots + Y_{t-n+1}}{n} \quad (1.3)$
Promedio Móvil Doble	El método hace lo que su nombre implica: se calcula un conjunto de promedios Móviles y después se calcula un segundo conjunto como promedio Móvil del primero. similar a la medición de una pendiente que cambia a través de la serie[16].	$M'_t = \frac{M_t + M_{t-1} + M_{t-2} + \dots + M_{t-n+1}}{n} \quad (1.4)$ $a_t = 2M_t - M'_t \quad (1.5)$ $b_t = \frac{2}{n-1} (M_t - M'_t) \text{ (Similar pendiente)} \quad (1.6)$ $\hat{Y}_{t+p} = a_t + b_t p \text{ (Ecuación de una recta)} \quad (1.7)$ <p>Donde: n = número de periodos en el promedio móvil; P = número de periodos a pronosticar hacia el futuro.</p>

Tabla 1.1 Métodos de Promedio

1.3.2 Métodos de Atenuación

Los métodos de atenuación por lo general son usados para pronosticar periodos cortos, en espectros con tendencias y tienen constantes de tendencia o de estacionalidad.

Métodos de Atenuación.

Método	Descripción	Característica
Atenuación Exponencial	Este método nos sirve, para llevar memoria sobre el pasado basado en una constante, este método puede ser muy útil cuando se desea que se atenúen las variaciones aleatorias o una respuesta rápida a las variaciones[7]. Si se desea una respuesta lenta ante variaciones aleatorias se requiere de un valor de α pequeño, si se desea una respuesta rápida al cambio real en el patrón de observaciones resulta mas apropiado un mayor valor de α .	$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$ $0 < \alpha < 1.$ <p style="text-align: right;">(1.8)</p>
Método de Brown	Se usa para pronosticar series de tiempo que tienen una tendencia lineal. Los conceptos básicos son similares a los de los promedios móviles dobles. Alfa se selecciona minimizando el error entre los valores reales y los pronósticos de estos valores utilizando el modelo. Se convierte en un proceso de prueba y error[14]. Este método tiene una propiedad bastante interesante, puesto que responde de una manera mas rápida ante variaciones en la curva.	$A_t = \alpha Y_t + (1-\alpha)A_{t-1}$ <p style="text-align: right;">(1.9)</p> $A'_t = \alpha A_t + (1-\alpha)A'_{t-1}$ <p style="text-align: right;">(1.10)</p> $a_t = 2A_t - A'_t$ <p style="text-align: right;">(1.11)</p> $b_t = \frac{\alpha}{(1-\alpha)}(A_t - A'_t)$ <p style="text-align: right;">(1.12)</p> $\hat{Y}_{t+p} = a_t + b_t p$ <p style="text-align: right;">(1.13)</p> <p>Coeficientes de INICIO:</p> $A_0 = a_0 - \frac{1-\alpha}{\alpha} b_0$ <p style="text-align: right;">(1.14)</p> $A'_0 = a_0 - 2\frac{1-\alpha}{\alpha} b_0$ <p style="text-align: right;">(1.15)</p>

Tabla 1.2 Métodos de Atenuación

Métodos de Atenuación.

Método	Descripción	Característica
Método Holt	La técnica de Holt atenúa en forma directa la tendencia y la pendiente empleando diferentes constantes de atenuación para cada una de ellas. En el enfoque de Brown, sólo se usaba una constante de atenuación y los valores estimados de la tendencia serán muy sensibles a variaciones aleatorias. La técnica de Holt proporciona mayor flexibilidad al seleccionar las proporciones a las que se rastrearán la tendencia y la pendiente.	$A_t = \alpha Y_t + (1 - \alpha)(A_{t-1} + T_{t-1}) \quad (1.16)$ <p>Tendencia:</p> $T_t = \beta(A_t - A_{t-1}) + (1 - \beta)T_{t-1} \quad (1.17)$ $\hat{Y}_{t+p} = A_t + pT_t \quad (1.18)$ <p> A_t = nuevo valor atenuado α = constante de atenuación de los datos ($0 \leq \alpha \leq 1$) Y_t = nueva observación o valor real de la serie, en el periodo t β = constante de atenuación de la estimación de la tendencia ($0 \leq \beta \leq 1$) T_t = estimación de la tendencia; p = periodos a pronosticar en el futuro $t+p$ = pronóstico de p periodos en el futuro </p>
Método Winter	Existe un patrón estacional[8], el modelo de atenuación exponencial lineal y estacional de tres parámetros de Winter[20], que es una extensión del modelo de Holt. En este método se utiliza una ecuación adicional para determinar la estacionalidad.	$A_t = \alpha \frac{Y_t}{S_{t-L}} + (1 - \alpha)(A_{t-1} + T_{t-1}) \quad (1.19)$ $T_t = \beta(A_t - A_{t-1}) + (1 - \beta)T_{t-1} \quad (1.20)$ $S_t = \gamma \frac{Y_t}{A_t} + (1 - \gamma)S_{t-L} \quad (1.21)$ $\hat{Y}_{t+p} = (A_t - pT_t) S_{t-L+p} \quad (1.22)$ <p> A_t = nuevo valor atenuado α = constante de atenuación Y_t = nueva observación β = constante de atenuación de estimación de la tendencia ($0 \leq \beta \leq 1$) T_t = estimación de la tendencia γ = constante de atenuación de la estimación de la estacionalidad ($0 < \gamma < 1$) S_t = estimación de la estacionalidad p = periodos a estimar a futuro L = longitud de la estacionalidad; Y_{t+p} = pronóstico de p periodos a futuro </p>

Tabla 1.3 Métodos de Atenuación

1.3.3 Regresión lineal

La línea que mejor se ajuste a un conjunto de puntos de datos $X - Y$, es aquella que minimiza la suma de las distancias al cuadrado de los puntos a la línea, medidas en dirección vertical o hacia Y . A esta línea se le conoce como la línea de regresión y su ecuación se denomina ecuación de regresión[6].

$$\hat{Y} = b_0 + bX \quad (1.23)$$

El problema radica en encontrar aquella recta que mejor ajuste a los datos. Tradicionalmente se ha recurrido para ello al método de mínimos cuadrados, que elige como recta de regresión a aquella que minimiza las distancias verticales de las observaciones a la recta. En la figura 1.2 se observa la nube de puntos y la regresión es la recta que pasa por casi todos los puntos.

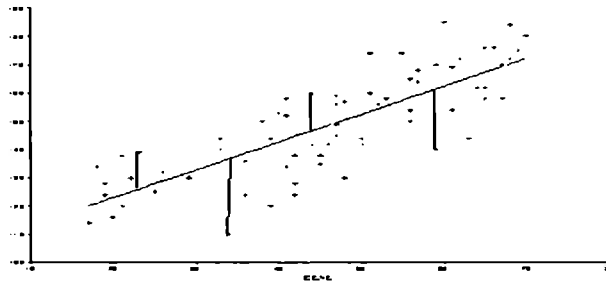


Figura 1.2 Línea de Regresión

Donde simplemente esta es la ecuación de una recta con pendiente b y desfasada del origen b_0 , donde se obtiene por mínimos cuadrados lo siguiente:

$$b = \frac{n \sum XY - \sum X \sum Y}{n \sum X^2 - (\sum X)^2} \quad (1.24)$$

$$b_0 = \frac{\sum Y}{n} - \frac{b \sum X}{n} \quad (1.25)$$

Con estas fórmulas (1.24) y (1.25) se encuentra de manera muy sencilla una aproximación lineal de nuestro pronóstico, en la práctica los sistemas no suelen ser lineales, pero definitivamente ésta podría ser una herramienta muy útil cuando se tenga un sistema de forma lineal[8].

1.3.4 Transformación

Cuando el conjunto de puntos dado no se aproxima a una recta, frecuentemente conviene “mapear” esos puntos en otro sistema, en otro espacio, para que se acerquen a una recta, en lugar de buscar la mejor curva que se ajuste a los puntos en el sistema original, por ejemplo si se tiene un conjunto de puntos en el sistema de referencia xy , al cual se les puede ajustar la curva de ecuación no lineal $y = f(x)$, es posible encontrar ecuaciones de transformación[14].

$$U = U(x,y) \tag{1.26}$$

$$V = V(x,y) \tag{1.27}$$

Con lo que se podría mapear esos puntos en otro sistema UV a los que se les ajuste la recta

$$V = A_0 + A_1 U \tag{1.28}$$

Uno de los sistemas de transformación más frecuentes utiliza logaritmos.

$$U = x \tag{1.29}$$

$$V = \log y \tag{1.30}$$

Y en el sistema UV el conjunto de puntos mapeados (U_i, V_i) lo que corresponden a una ecuación 1.28, en el sistema xy se definirá en forma semejante la relación:

$$\log y = A_0 + A_1 \log x = \log 10^{A_0} + \log x^{A_1} \tag{1.31}$$

$$y = 10^{A_0} x^{A_1} \tag{1.32}$$

$$\alpha = 10^{A_0} \tag{1.33}$$

$$\beta = A_1 \tag{1.34}$$

$$y = \alpha \cdot x^\beta \tag{1.35}$$

Una vez que se ha desarrollado la técnica del pronóstico por cualquier modelo , lo siguiente a realizar es compararla con el dato real, y para ello se mide el error. El Error medio cuadrado, el cual consiste en elevar al cuadrado cada residual, luego estos valores se suman y se divide entre el número de observaciones, este enfoque penaliza los errores mayores de pronóstico[8].

$$EMC = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n} \quad (1.37)$$

El Porcentaje medio de error se calcula como se indica en la fórmula 1.36 de manera que si el resultado es un porcentaje negativo el método del pronóstico esta sobreestimado mientras que si es positivo el método esta subestimado.

$$PME = \frac{\sum_{i=1}^n \frac{(Y_i - \hat{Y}_i)}{Y_i}}{n} \quad (1.36)$$

1.4 Selección de Variables a través de La Técnica de Componentes Principales

La técnica de componentes principales tiene por objetivo el reducir el número de indicadores en el análisis de un problema que involucra múltiples variables. Por ejemplo, para el caso del mercado bursátil, se ha sugerido el uso de una gran cantidad de indicadores asociados con el tema. La complejidad del análisis se incrementa a medida que el número de variables crece cada una de ellas cuenta una parte distinta de la historia aunque pueda parecer que dos o más "casi" son idénticos y, en consecuencia, nos dicen "casi" lo mismo. Es decir, parece haber alguna redundancia entre indicadores, y de hecho la hay, pero cada uno de ellos aporta información no contenida en los demás[18]. Cuando la complejidad del estudio causada por la presencia de un gran número de variables es muy grande, no es extraño intentar priorizar el uso de un subconjunto. De este modo, se deberá ignorar un buen número de ellas y, en consecuencia, la información adicional que las mismas contienen. Sin embargo, la estadística matemática ha propuesto una serie de alternativas de análisis que buscan ayudar a resolver el conflicto: complejidad en el análisis contra pérdida de información. En términos estadísticos, "casi idénticos" se mide con buena precisión a través del coeficiente de correlación lineal. Más precisamente, dicho coeficiente mide el grado de redundancia lineal existente entre dos indicadores. Cuando toma el valor 1, es posible decir que ambos indicadores son redundantes; lo mismo ocurre cuando su valor es igual a -1, salvo que en este caso ambos indicadores se

comportan en sentidos opuestos. Por lo contrario, cuando su valor es cero, se dice que no existe ninguna asociación lineal entre ambos indicadores.

El coeficiente de correlación lineal, o la medida asociada denominada covarianza, constituye la base del análisis por componentes principales. A partir de los valores que toman estas medidas de asociación cuando se consideran todas las parejas de indicadores involucradas en el estudio y de los valores de éstos, es posible definir nuevos indicadores denominados componentes principales.

El primer componente principal es aquel indicador resumen que explica la mayor heterogeneidad entre los casos; es decir, tiene la mayor varianza. El segundo componente principal es un indicador resumen que es a la vez el más heterogéneo después del primer componente y que no está correlacionado con ésta. El tercer componente no está correlacionado ni con el primer ni con el segundo. Las restantes componentes principales cumplen con condiciones semejantes, el método se puede ver en el apéndice A.

Se tienen dos situaciones extremas muy improbables en la práctica, cuando las correlaciones entre todas las parejas de variables son iguales a cero, lo que ocurre cuando no hay redundancia entre ellas, habrá tantas componentes principales con varianzas distintas de cero como indicadores originales. Por otro lado, cuando la redundancia es total, sólo la primer componente principal tendrá una varianza distinta de cero. En el primer caso, la reducción de dimensionalidad se logrará sólo sacrificando información. En el segundo bastará con realizar el análisis a partir de la primer componente principal.

En otras palabras, la primer componente principal resumirá toda la información sólo en el caso de redundancia perfecta entre todos los indicadores. En cualquier otra circunstancia se hace imprescindible considerar algunas más, siempre buscando que la pérdida de información sea lo más pequeña posible. Por lo tanto, las componentes principales no son el resultado final para llevar a cabo el análisis, sino el recurso estadísticamente riguroso para detectar y seleccionar los indicadores con mayor capacidad de sintetizar las diferencias habidas entre las observaciones. Una vez hecha esta selección de indicadores, después seguirá encontrar los distintos grupos homogéneos de observaciones que serán clasificados y ordenados jerárquicamente. ACP es un método estadístico que permite realizar una transformación lineal del espacio de muestras de modo de extraer las características óptimas de la distribución de entrada, en el sentido del Mínimo Error Cuadrático Medio. En ACP los vectores de entrada se representan en términos de los vectores propios de la matriz de covarianza asociada a las muestras de entrada[2].

2 LOGICA DIFUSA

2.1 Lógica Difusa.

En 1960, Lotfi Zadeh inventó la lógica difusa guiado por el principio de que las matemáticas pueden ser usadas para encadenar el lenguaje con la inteligencia humana[10]. Algunos conceptos pueden ser mejor definidos en términos de palabras, que por matemáticas; la lógica difusa y su expresión en conjuntos difusos, proveen una disciplina que puede construir mejores modelos de la realidad. La lógica difusa difiere de la dicotómica en este sentido. En lógica dicotómica se espera derivar una solución decidiendo por sí o por no si cada una de las restricciones o parámetros es verdadero o falso, pero en lógica difusa es admisible usar escalas de condiciones (restricciones) y matices (flexibilidad) en los valores numéricos. Por ello está algo menos interesado en la verdad y algo más interesado en la facilidad práctica, es así como nuestro cerebro va tomando decisiones, es decir no tenemos salidas solamente de verdadero o falso.

La lógica difusa que en un primer nivel puede decirse que es un lenguaje que permite trasladar sentencias sofisticadas del lenguaje natural a un formalismo matemático, lo podemos entender como aterrizar estas premisas del lenguaje humano. El conocimiento se adquiere y se manipula de una manera inferencial y deductiva, por medio del razonamiento simbólico, es decir la capacidad para manipular signos que sitúan a algo generalmente en una estructura o red de hechos. A este conjunto de hechos inciertos, que se requieren para la solución de un problema, se le ha dado el nombre de conjuntos difusos y a sus reglas lógica difusa.

Las técnicas difusas son necesarias, o benéficas en[3]:

- Sistemas complejos, donde es muy difícil o imposible crear un modelo. (como es el caso de las predicciones en la bolsa de valores.)
- Sistemas controlados por expertos humanos.
- Sistemas con complejas y continuas entradas y salidas.
- Sistemas que usan observaciones humanas como entradas o como reglas básicas.
- Sistemas que son naturalmente vagos como las ciencias sociales.
- Cuando se maneja en el sistema experto un porcentaje elevado de incertidumbre.

Los Sistemas difusos, pueden ser usados para realizar: clasificaciones, predicciones y control. Algunas de las aplicaciones de la lógica difusa ha sido[19]:

En el esquema de predicciones, la lógica difusa se ha utilizado en: el Fuji Bank, en Tokio, usa un sistema que maneja lógica difusa para efectuar transacciones con bonos a corto plazo. Este sistema ha estado en operación durante los dos últimos años, y consta de unas 200 reglas difusas que se basan en estrategias financieras de uso común. Dado que en la vida real este tipo de reglas tienen un tiempo de vida limitado, las reglas difusas del sistema son también actualizadas regularmente por expertos humanos, usando como base datos históricos y su propia experiencia. Se ha estimado que este sistema le ha permitido al Fuji Bank ganar unos 970,000 dólares mensuales.

En el esquema de clasificaciones la lógica difusa se ha utilizado: En la extracción convencional de información no se permite el tipo de flexibilidad a la que los humanos estamos acostumbrados, porque el tipo de búsqueda que se efectuaría podría pasar por alto detalles importantes que para un ser humano resultarían obvios. Por ejemplo, si queremos detectar a agentes de ventas jóvenes que sean prometedores, tendríamos que definir qué entendemos por "joven" y por "prometedor". Si, por ejemplo, para nosotros, alguien de menos de 25 años es joven, entonces podríamos pasar por alto el desempeño de vendedores de 26 años. Análogamente, si para nosotros es "prometedor" alguien que vende más de 100,000 pesos en un año, podríamos ignorar a quien vende 99,000 pesos. La lógica difusa permite definir de manera más flexible el tipo de búsqueda que deseamos realizar, y nos proporcionaría mucha mayor información que las técnicas tradicionales.

2.2 Conjunto borroso.

Un conjunto difuso es un conjunto que no tiene límites claramente definidos o precisos, a diferencia de los conjuntos clásicos, en los conjuntos difusos la transición de la pertenencia o no-pertenencia de un elemento a un cierto conjunto, es gradual, y esta transición está caracterizada por funciones de membresía, las cuales les da a los conjuntos difusos flexibilidad para modelar expresiones lingüísticas empleadas cotidianamente.

Sea X una colección de objetos denotados genéricamente por x ; entonces, un conjunto difuso A en X se define como el siguiente conjunto de pares ordenados: $A = \{(x, \mu_A(x)) \mid x \in X\}$

donde $\mu_A(x)$ se conoce como la función de membresía (abreviado FM) de A. La FM transforma cada elemento de X hacia un grado de membresía (o valor de membresía) entre 0 y 1. Generalmente, X es llamado el universo de discurso o, simplemente, el universo, el cual puede consistir de objetos discretos (ordenados o no ordenados) o ser un espacio continuo. Por ejemplo, el universo ordenado discreto definido por la puntuación que cualquier persona puede realizar en un juego de video, se expresa como: $X = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. De esta forma, el conjunto $A = \text{“puntuación ideal en un juego de video”}$ se puede describir como:

$A = \{(0, 0.1), (1, 0.3), (2, 0.7), (3, 0.8), (4, 0.9), (5, 1), (6, 0.9), (7, 0.8), (8, 0.7), (9, 0.3), (10, 0.1)\}$ donde los grados de membresía asignados, son medidas subjetivas . Un ejemplo de universo continuo de 1 hasta 30 años. $B = \{\text{conjunto de gente joven}\}$ se puede expresar de la siguiente manera, en la figura 2.1 se muestran las FM correspondientes a los conjuntos difusos[9].

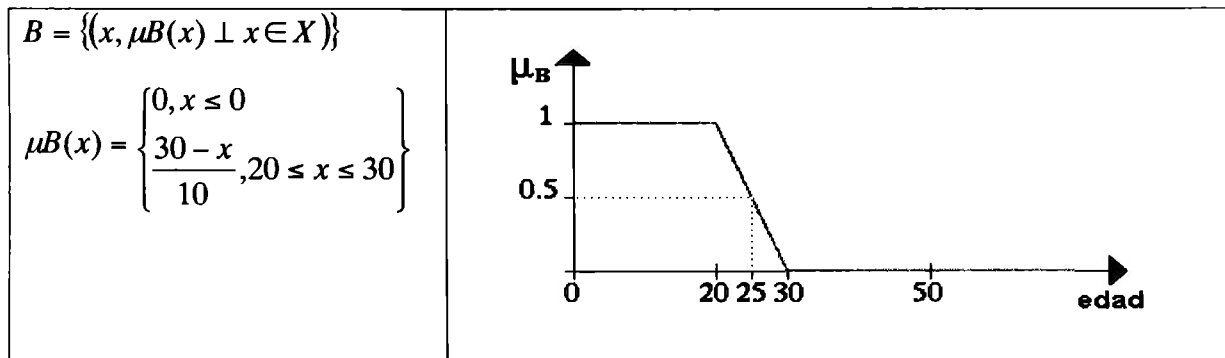


Figura 2.1 FM para un universo continuo

En la práctica, cuando el universo de discurso X es un espacio continuo, generalmente se divide en varios conjuntos difusos cuyas FM cubren a X de una manera más o menos uniforme. Estos conjuntos se denominan valores lingüísticos o etiquetas lingüísticas y, normalmente, se les asignan nombres de adjetivos utilizados en el lenguaje común, tales como “grande”, “frío”, “alto”, etc., la figura 2.2 proporciona el ejemplo de FM etiquetadas en todo el universo de discurso. Una variable lingüística se define mediante el quinteto de elementos siguiente: (x, T(x), X, G, M), donde x es el nombre de la variable; T(x), el conjunto de términos de x, esto es, el conjunto de sus valores lingüísticos o términos lingüísticos; X es el universo de discurso; G, una regla sintáctica que genera los términos en T(x), M, una regla semántica, la cual le asocia a cada valor lingüístico A un valor correspondiente $M(A)$, que denota un conjunto difuso en X.

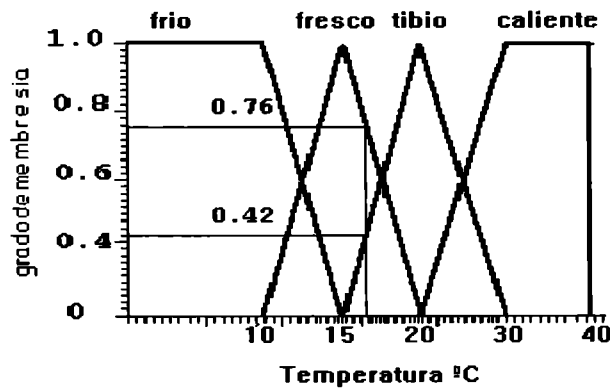


Fig. 2.2. Funciones de Membresía traslapadas

Por ejemplo, si temperatura se interpreta como una variable lingüística, entonces su conjunto de términos $T(\text{temperatura})$ podría ser: $T(\text{temperatura}) = \{ \text{frío, no frío, muy frío, no muy frío, ..., tibio, medio tibio, algo tibio, ..., caliente, no caliente, muy caliente, no muy caliente, ... } \}$ donde cada término en $T(\text{temperatura})$ se caracteriza por un conjunto difuso de un universo de discurso que puede ser: $X=[0, 40]$. Normalmente se dice “la temperatura es fría” para denotar la asignación del valor lingüístico “frío” a la variable lingüística temperatura. Por otro lado, cuando la variable temperatura se interpreta como una variable numérica, se emplea la expresión “temperatura=30°C” para asignarle el valor numérico “30” a la variable numérica temperatura. La regla sintáctica se refiere a la forma en que se generan los valores lingüísticos en el conjunto de términos $T(\text{temperatura})$. La regla semántica define la función de membresía de cada valor lingüístico del conjunto de términos. En la figura 2.2 hemos definido la función frío comprendiendo el rango de entre 0 a 15 grados ($a=0, d=15$); Los valores comprendidos entre 0 y 10 tendrán la máxima membresía de 1.0 ($b=0, c=10$). La membresía para fresco es un poco diferente. Aquí se desearía que fresco tuviera su máxima con los puntos $b = c$.

Las funciones de membresía se traslapan, el grado de membresía de una temperatura de entrada se determina mediante donde cae su valor dentro de la(s) función (es). En ejemplo, una entrada de digamos 16 grados caerá dentro de frío y de fresco, esto es, existe una zona de traslape para las funciones de membresía. Esto significa que un determinado valor de entrada puede tener cierto grado de membresía en más de una función de membresía.

En el ejemplo, los 16 grados tendrá un grado de membresía de 0.76 en la función fresco y de 0.42 en la función tibio. Esto es, la temperatura de entrada de 16 grado se considera a la vez tanto fresco como tibio. Por tanto, los conjuntos difusos y sus funciones de membresía permiten que adopte varios valores, no únicamente blanco y negro, si no un rango continuo de grises. Todo lo

que se logra con los conjuntos difusos, es categorizar el mundo en la forma como la hace la mente humana. La utilidad de la lógica difusa, son las reglas lógicas que operan sobre los conjuntos difusos.

2.3 Funciones de membresía

Las funciones de membresía mas comunes son[4]:

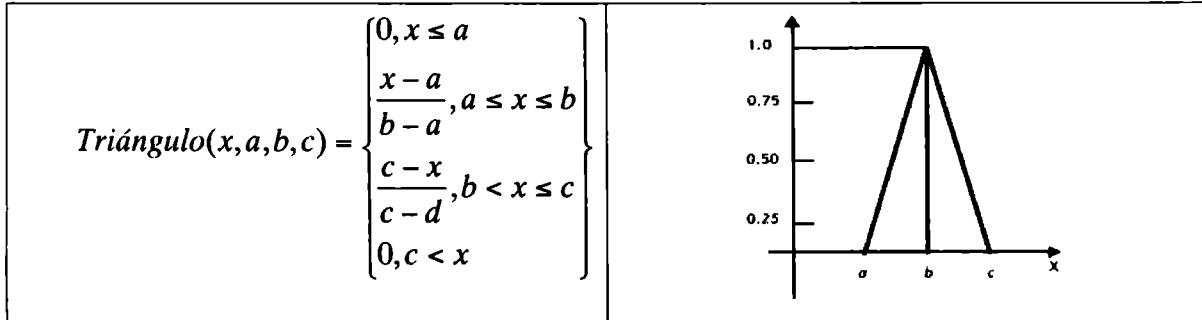


Figura 2.3 Triángulo

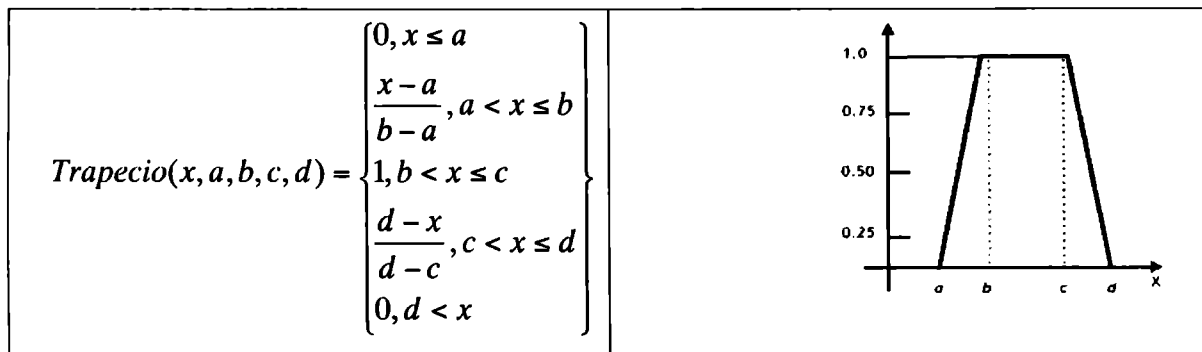


Figura 2.4 Trapecio

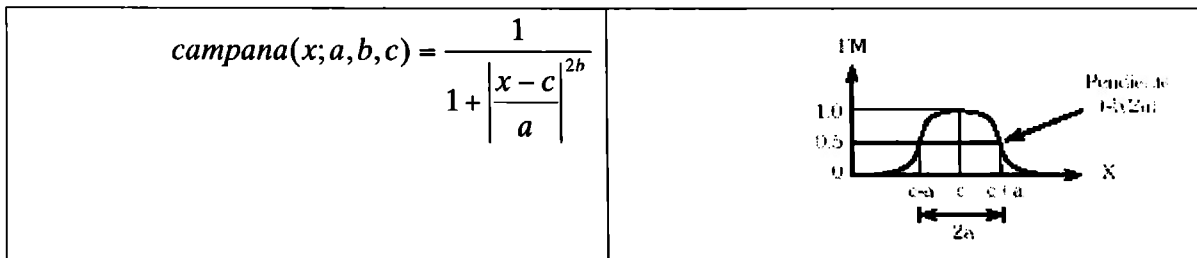


Figura 2.5 Campana

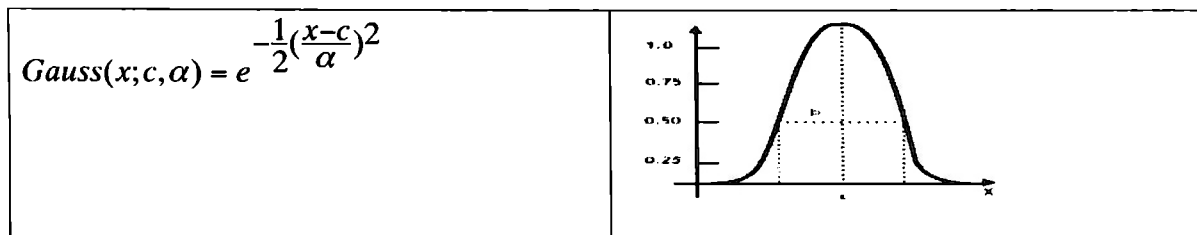


Figura 2.6 Gauss

Otras FM se pueden crear para aplicaciones específicas, si ello se requiere. En general, cualquier tipo de función de distribución de probabilidad estadística puede emplearse como FM, podría ser con pendientes distintas en ambos lados, si así se requiere.

2.4 Reglas Difusas

Operador norma-T[15].

Un operador norma-T es una función que se aplica a una pareja de valores, de la forma $T(*, *)$, y satisface las siguientes condiciones:

$$T(0, 0) = 0, T(a, 1) = T(1, a) = a \text{ (límites)}$$

$$T(a, b) \geq T(c, d) \text{ si } a \geq c \text{ y } b \geq d \text{ (monotonicidad)}$$

$$T(a, b) = T(b, a) \text{ (conmutatividad)}$$

$$T(a, T(b, c)) = T(T(a, b), c) \text{ (asociatividad)}$$

La primera condición establece la generalización para los conjuntos certeros; la segunda, implica que un decremento en los valores de membresía de A o B, no producirá un incremento en el valor de membresía de $A \cap B$. La tercera condición indica que el orden es indiferente al orden de los conjuntos difusos a ser combinados; la cuarta, permite tomar la intersección de cualquier cantidad de conjuntos, en cualquier orden, agrupándolos por parejas. Algunos de los operadores norma-T más comunes son:

$$\text{Minimum (Min)} = T_{\text{Min}}(a, b) = \min(a, b)$$

$$\text{Producto algebraico} = T_{\text{pa}}(a, b) = ab$$

$$\text{Producto Limitado} = T_{\text{pl}}(a, b) = 0 \vee (a + b - 1)$$

$$\text{Producto Drástico} = T_{\text{pd}}(a, b) = \begin{cases} a, b = 0 \\ b, a = 1 \\ 0; a, b < 1 \end{cases}$$

Un operador conorma-T (o norma-S) es una función que se aplica a un par de valores, de la forma $S(*, *)$, y satisface lo siguiente:

$$S(1, 1) = 1, S(a, 0) = S(0, a) = a \text{ (límites)}$$

$$S(a, b) \leq S(c, d) \text{ si } a \leq c \text{ y } b \leq d \text{ (monotonicidad)}$$

$$S(a, b) = S(b, a) \text{ (conmutatividad)}$$

$$S(a, S(b, c)) = S(S(a, b), c) \text{ (asociatividad)}$$

La justificación de estos requerimientos es similar a la de los correspondientes para los operadores norma-T. En forma similar, se presentan a continuación cuatro operadores norma-S comunes:

$$\text{Maximum (Max)} = S_{Max}(a, b) = \max(a, b)$$

$$\text{Suma algebraico} = S_{Sa}(a, b) = a + b - ab$$

$$\text{Suma Limitada} = S_{Sl}(a, b) = 1 \wedge (a + b)$$

$$\text{Suma Drástica} = S_{Sd}(a, b) = \begin{cases} a, b = 0 \\ b, a = 1 \\ 1; a, b > 0 \end{cases}$$

2.4.1 Fusificación

Las reglas definen las relaciones entre los conjuntos difusos de entrada y salida. Las reglas reflejan como una persona normalmente pensaría sobre el proceso o la aplicación, tanto en forma lógica como lingüística, cada regla está en la forma de una sentencia SI ... ENTONCES. La porción correspondiente al SI, o antecedente, especifica la condición que deberá cumplirse. La porción ENTONCES, o consecuencia, especifica la consecuencia del antecedente. Una regla difusa del tipo si ... entonces[4] (también llamada simplemente regla difusa, implicación difusa o declaración condicional difusa) asume la siguiente forma: si x es A entonces y es B, donde A y B son valores lingüísticos, definidos por conjuntos difusos en los universos de discurso X y Y, respectivamente. Frecuentemente a “x es A” se le llama el antecedente o premisa, mientras que a “y es B” se le llama el consecuente o conclusión. Una regla simple podría escribirse como:

Si TEMPERATURA es CALIENTE entonces VELOCIDAD DE VENTILADOR es ALTA.

Es posible formar reglas lógicas más complejas al utilizar el operador lógico Y. La sentencia SI...ENTONCES, junto con el operador Y, pueden considerar la gran mayoría de las relaciones requeridas en sistemas y procesos.

Por ejemplo, se considera además de la temperatura la humedad relativa del medio ambiente definida por los conjuntos difusos: seco, normal y húmedo, podría definirse la regla.

Si TEMPERATURA es CALIENTE y HUMEDAD es HÚMEDO entonces VELOCIDAD DE VENTILADOR es MEDIA

Por supuesto, la lógica estándar tiene una base de reglas semejante. Lo que hace que la lógica difusa sea única es la propia evaluación de las reglas dada la naturaleza de los conjuntos difusos[3].

Ya se ha considerado con algún detalle la componente de fusificador. Este es el mecanismo de conversión de una entrada a valores difusos mediante la definición de conjuntos difusos y sus funciones de membresía. A este procedimiento se le denomina fusificación.

2.4.2 Evaluación de Reglas

El segundo paso es la evaluación de estas reglas donde se pueden usar operadores (AND u OR)

$$\mu_{A \cup B}(X) = \max[\mu_A(x), \mu_B(X)] \quad (2.5)$$

$$\mu_{A \cup B}(X) = \text{probor}[\mu_A(x), \mu_B(X)] = \mu_A(X) + \mu_B(X) - \mu_A(X) \times \mu_B(X) \quad (2.6)$$

$$\mu_{A \cap B}(X) = \min[\mu_A(x), \mu_B(X)] \quad (2.7)$$

$$\mu_{A \cap B}(X) = \text{prod}[\mu_A(x), \mu_B(X)] = \mu_A(X) \times \mu_B(X) \quad (2.8)$$

Generalmente los paquetes que en la práctica se ocupa para un OR la ecuación 2.5 y para un AND la ecuación 2.7. Recordando que para una determinada entrada pueden lograrse distintos grados de membresía, esto es, puede existir traslape. Esto es importante al evaluar las sentencias en las reglas lógicas[9].

Las reglas lógicas en un sistema de lógica difusa tienen una cierta "fortaleza" o grado de influencia sobre el resultado de la salida final. Esta fortaleza depende del grado de membresía de sus variables de entrada y se salida.. Entre más alta sea su fortaleza, en mayor grado la regla es cierta, Por ejemplo, la regla: Si TEMPERATURA es CALIENTE entonces VEL VENTILADOR es RÁPIDO. Será verdadera siempre que la temperatura de entrada se encuentre dentro de la función de membresía "CALIENTE". Esto es, la regla se aplicará o se "disparará", siempre que el

valor de la entrada para CALIENTE sea mayor a 0.0; así, si se tiene una entrada (en la función de membresía CALIENTE) de 0.50, la fortaleza de la regla será menor que cuando la entrada tuviera un valor de membresía de 0.60.

Para el caso de reglas compuestas en las que se emplea el operador Y, la fortaleza de la regla se determina al tomar de los dos términos del antecedente aquél que tiene el menor grado de membresía. Por ejemplo, para la regla: Si TEMPERATURA es CALIENTE y HUMEDAD es ALTA entonces VEL VENTILADOR es RAPIDO. Existen aquí dos entradas a considerar TEMPERATURA y HUMEDAD. Si la temperatura tiene un grado de membresía de 0.50 en la función de membresía CALIENTE, y la humedad tienen un 0.40 de membresía en la función ALTA, la fortaleza de la regla será tomada como el menor de los valores que intervienen en la regla (0.40).

Como frecuentemente los valores de las funciones de entrada, caen en más de una función de membresía, es posible obtener reglas aparentemente contradictorias. Por ejemplo, dada una sola entrada, dos reglas pueden disparar: Si TEMPERATURA es CALIENTE entonces VEL VENTILADOR es ALTA ; Si TEMPERATURA es TIBIO entonces VEL VENTILADOR es MEDIO. Cuando, debido al traslape, el valor proporcionado como entrada cae en las dos funciones de membresía CALIENTE Y TIBIO, dispararán ambas reglas.

Durante la evaluación de la regla, las fortalezas de cada una de ellas se calculan basándose en sus antecedentes y entonces asignando a la salida de la regla, utilizando el antecedente menos cierto o más débil. Cuando mas de una regla se aplica a una acción específica , la salida se resuelve tomando la más fuerte o regla más verdadera:

Regla 1: si A y B entonces Z y X Regla 2: si C y D entonces Z y Y ;
 Fortaleza de Regla 1 = $\min(A,B)$; Fortaleza de Regla 2 = $\min(C,D)$

X = Fortaleza de Regla 1 Y = Fortaleza de Regla 2

Z = Salida = $\max(\text{Fortaleza de Regla 1}, \text{Fortaleza de Regla 2})$; Z = $\max(\min(A,B), \min(C,D))$

2.4.3 Defusificación de Salidas

El defusificador, es el mecanismo empleado para convertir los valores difusos de la salida a un valor tajante utilizable por la aplicación. Aún cuando el proceso de evaluación de reglas asigna fortalezas a cada acción específica, se requiere otra etapa mas o "defusificación", por dos razones:

Primero, para descifrar el significado de acciones "vagas", tal como la velocidad del ventilador es alta, empleando funciones de membresía. Y segundo, para resolver conflictos entre dos acciones que compiten entre si tales como "la velocidad es mediana" y "la velocidad en alta", que pudieran haber sido disparadas bajo ciertas condiciones durante la evaluación de las reglas. El proceso de defusificación emplea técnicas para resolver posibles conflictos y vaguedad resultantes durante la evaluación de las reglas[10].

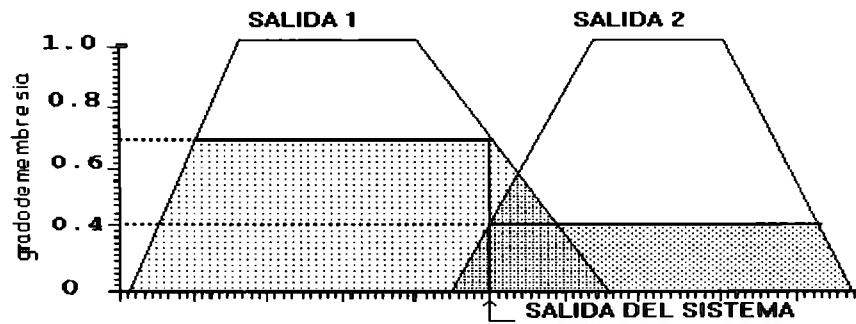


Figura 2.7 Defusificación

La técnica usada es el "método del centro de gravedad", este consiste en varios pasos. Inicialmente, se calcula el centroide (centro de gravedad), para cada función de membresía; entonces, las funciones de membresía se limitan en su altura (ver figura 2.3) dependiendo de la fortaleza de la regla aplicada, y se evalúan las áreas trapezoidales de las funciones de membresía. Finalmente, se obtiene la salida defusificada mediante el cálculo de un promedio ponderado de los puntos correspondientes a los centros de gravedad y las áreas calculadas, siendo estas últimas los factores de peso o ponderación.

Es la manera de dar al sistema de control un número, a final de cuenta sea que se tenga sistema experto para controlar velocidad de coches o un sistema experto para dar pronósticos en la bolsa de valores, se debe tener salidas reales con un número para cada entrada de variables.

El método del centro de gravedad (COG) es el más usado para dar una salida a la figura que se ha formado a partir de las entradas, matemáticamente la fórmula es:

$$COG = \frac{\int_a^b \mu_A(x) x dx}{\int_a^b \mu_A(x) dx} \tag{2.9}$$

Aunque en la práctica se puede hacer esta bajo la formula:

$$COG = \frac{\sum_a^b \mu_A(x)x}{\sum_a^b \mu_A(x)} \quad (2.10)$$

El procedimiento completo se denomina controlador de Mamdani y se puede ver en el apéndice B

2.5 Modelo difuso tipo Sugeno.

Este modelo fue propuesto inicialmente por Takagi, Sugeno y Kang , en un esfuerzo por desarrollar un método sistemático para generar reglas difusas a partir de un conjunto de datos de entrada/salida. Mandani-style inference, requiere encontrar el centro de gravedad de una forma de dos dimensiones (obviamente en un esquema de x-y) , en general este proceso pierde eficiencia, es por esta razón que esta sugeno, en el cual solo tenemos un spike o un “singleton”, como la función de membresía de la regla consecuente, un “singleton” es un conjunto difuso con una función de membresía que es la unidad[10], sugeno es básicamente lo mismo que Mamdani, solo cambia en la regla consecuente es decir:

IF x es A AND y es B THEN z es f(x,y)

Donde x, y y z son variables lingüísticas; A y B son conjuntos fuzzy, y f(x,y) es una función matemática, el mas común usado es el sugeno de orden cero en lo que tenemos:

IF x es A AND y es B THEN z es K

Donde K es una constante

En este caso la salida de cada regla fuzzy es una constante, en otras palabras todas las funciones de membresía son representadas por “singleton”, en el apéndice C tenemos el diagrama sugeno de orden cero y los pasos a seguir para realizar un controlador tipo Sugeno

2.6 ANFIS

2.6.1. Redes Adaptables.

Como su nombre lo indica, una red adaptable es una estructura cuyo comportamiento global de entrada/salida queda determinado por una colección de parámetros modificables. Específicamente, la configuración de una red adaptable se compone de un conjunto de nodos conectados por enlaces direccionados, donde cada nodo realiza un tipo de función de nodo estática en las señales que entran a él, para generar una sola salida de nodo, y cada enlace especifica la dirección del flujo de señales de un nodo a otro. Generalmente, una función de nodo es una función parametrizada con parámetros modificables; cambiando dichos parámetros, se modifica la función de nodo y, en consecuencia, el comportamiento global de la red adaptable. La regla de aprendizaje especifica la forma en que estos parámetros deben ajustarse para minimizar una medida de error dada, la cual es una expresión matemática que mide la desviación entre la salida real de la red y una salida deseada. Aquí se debe mencionar que si la regla de aprendizaje se aplica empleando pares de datos de entrada/salida, esto es, que para cada dato de entrada se le especifica a la red la respuesta que debe tener en su salida, se habla de una regla de aprendizaje supervisado y se le denomina una red de aprendizaje supervisado (o, simplemente, red supervisada); por otro lado, si únicamente se utilizan vectores de entrada, se tendrá un método de aprendizaje sin supervisión, y se hablará de una red de aprendizaje no supervisado. Puesto que el sistema empleado en esta tesis utiliza aprendizaje supervisado, el enfoque será en esa dirección[4]. La regla básica de aprendizaje de una red adaptable es el método descendiente abrupto (steepest descent method), en el cual el vector gradiente se deriva mediante invocaciones sucesivas de la regla de la cadena en el cálculo de derivadas de una función continua. Este es el método que se usó originalmente para hallar el gradiente en una red neuronal multicapa, bajo el nombre de regla de aprendizaje de retropropagación (backpropagation learning rule), nombre que actualmente resulta muy conocido en los tópicos de redes neuronales. Los parámetros de una red adaptable se distribuyen hacia sus nodos, de tal forma que cada nodo tiene un conjunto local de parámetros. La unión de estos conjuntos locales de parámetros genera el conjunto global de parámetros de la red. Si el conjunto de parámetros de un nodo tiene elementos ajustables, entonces su función de nodo depende de los valores de los parámetros; en este trabajo se usará un cuadrado para representar este tipo de nodo adaptable. Por otra parte, si un nodo tiene un conjunto de parámetros sin elementos ajustables, entonces su función es fija y se representa con

un círculo (nodo fijo). Los enlaces de una red adaptable se utilizan meramente para señalar la dirección de propagación de las salidas de los nodos; generalmente no hay pesos o parámetros asociados con ellos. La figura 2.8 ilustra lo dicho arriba para una arquitectura de dos entradas, tres nodos en una capa oculta y dos nodos en la capa de salida (abreviadamente: 2-3-2).

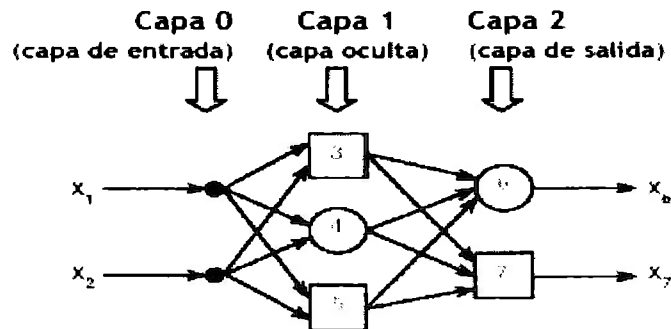


Fig. 2.8 Red Feedward

Las redes adaptables generalmente se clasifican en dos categorías de acuerdo al tipo de enlaces que tienen: Redes de avance (feedforward) y redes recurrentes (recurrent). La figura 2.8 es un ejemplo de red de avance, puesto que la salida de cada nodo se propaga del lado de entrada (izquierda) hacia el lado de salida (derecha). Si existe algún enlace de retroalimentación que forme una trayectoria circular en una red, entonces la red es recurrente. Conceptualmente, una red adaptable de avance realiza una transformación estática entre sus espacios de entrada y de salida. Esta transformación puede ser una relación lineal simple o una altamente no lineal, dependiendo de la estructura de la red y la función de cada nodo. El objetivo es lograr construir una red para realizar una transformación no lineal que esté regulada por un conjunto de datos de pares deseados de entrada/salida de un sistema que se quiera modelar. A este conjunto de datos normalmente se le llama el conjunto de datos de entrenamiento, y a los procedimientos que se siguen para ajustar los parámetros para optimizar el desempeño de la red, normalmente se les llama reglas de aprendizaje o algoritmos de adaptación. El desempeño de una red se determina midiendo el error entre la salida deseada y la salida de la red bajo las mismas condiciones de entrada. A esta variable se le llama la medida del error y puede tomar diferentes formas para diferentes aplicaciones. En general, una regla de aprendizaje se deriva al aplicar una técnica de optimización específica a una medida de error dada.

2.6.2 Sistemas neurodifusos

Hasta este punto ya se han descrito todos los elementos necesarios para construir un sistema neurodifuso propiamente dicho. En los sistemas puramente difusos, el problema de encontrar las

FM apropiadas así como las reglas difusas para un problema específico, frecuentemente es un proceso tedioso de prueba y error. Por lo tanto, la idea de aplicar algoritmos de aprendizaje a estos sistemas surgió inicialmente con los llamados controladores difusos auto-organizativos. Sin embargo, otra posibilidad más aceptada actualmente para el ajuste de los parámetros de los sistemas difusos surge de los métodos de aprendizaje empleados en redes neuronales. Los sistemas neurodifusos modernos normalmente se representan como una red neuronal multicapa de avance. Esto se hace simplemente para facilitar el análisis en el momento de aplicar el proceso de optimización e ilustrar el flujo de datos dentro del sistema, así como su naturaleza en paralelo. Los nodos en estas redes difieren de los utilizados en las redes neuronales en cuanto al tipo de función que realizan y la interpretación que se les da. Aunque existen muchos métodos diferentes, el término de sistemas neurodifusos se puede restringir a aquellos sistemas que presentan las siguientes propiedades: Un sistema neurodifuso es un sistema difuso que se entrena por medio de un algoritmo de aprendizaje (comúnmente) derivado de la teoría de las redes neuronales[9]. El procedimiento de aprendizaje opera en información local, y origina solo modificaciones locales en el sistema difuso de base. El proceso de aprendizaje no se basa en conocimiento, sino en un conjunto de datos dado. Un sistema neurodifuso puede verse como una red neuronal multicapa de avance. Las unidades en sus nodos emplean operaciones norma-T o conorma-T en lugar de las funciones de activación que normalmente usan las redes neuronales. La primera capa representa las variables de entrada; las FM y las reglas difusas se representan mediante enlaces y capas ocultas; y, la última capa, representa las variables de salida. Sin embargo, esta representación en forma de red neuronal no es un requisito indispensable para aplicar un procedimiento de aprendizaje, simplemente se hace por conveniencia. Un sistema neurodifuso siempre puede interpretarse como un sistema de reglas difusas, esto es, antes, durante y después del entrenamiento. Es posible crear el sistema a partir de datos de entrenamiento únicamente, o se pueden inicializar sus parámetros mediante un conocimiento previo en forma de reglas difusas. El procedimiento de entrenamiento de un sistema neurodifuso toma en consideración las propiedades semánticas del sistema difuso base. Esto resulta en restricciones en las posibles modificaciones de los parámetros del sistema. Un sistema neurodifuso aproxima una función de n dimensiones (desconocida) que se define parcialmente mediante los datos de entrenamiento. Las reglas difusas codificadas dentro del sistema representan muestras vagas, y se pueden interpretar como prototipos vagos de los datos de entrenamiento. Un sistema neurodifuso no debe entenderse como un tipo de sistema (difuso) experto, y no se relaciona con la lógica difusa en un sentido estrecho[11]. Los sistemas neurodifusos surgen debido a la falta de métodos estándar para transformar el conocimiento

humano o su experiencia hacia la base de las reglas y datos de un sistema de inferencias difuso, para lograr el objetivo anterior es necesario ajustar sus funciones de membresía apropiadas para minimizar la medida del error de salida. Con esta idea surge ANFIS (Sistema de inferencias difuso basado en redes adaptables), la cual sirve como base para construir un conjunto difuso si....entonces con funciones de membresía apropiadas para generar los pares de datos de entrada/salida óptimos. Estos se logra usando una regla de aprendizaje híbrida, basada en los métodos de optimización comunes de las redes adaptables.(el descendente abrupto y el estimador de mínimos cuadrados).

Sistema de inferencia difuso basado en redes adaptables: anfis.

Esta arquitectura soporta modelos difusos de sugeno de orden 0, y 1.

Para explicar el funcionamiento de la arquitectura, se considera un sistema de inferencias con dos entradas (x, y) y una salida (f). Para un modelo difuso del tipo Sugeno de primer orden, un conjunto de reglas con dos reglas difusas si ... entonces, se define como:

Regla 1: Si x es A1 y y es B1, entonces $z_1 = p_1 x + q_1 y + r_1$,

Regla 2: Si x es A2 y y es B2, entonces $z_2 = p_2 x + q_2 y + r_2$.

La figura 2.9 ilustra el mecanismo de razonamiento para este modelo de Sugeno; la arquitectura ANFIS correspondiente se muestra en la figura 2.10, en la cual, los nodos en una misma capa realizan funciones similares. A continuación se da una descripción de la función que realiza cada capa en la arquitectura mostrada (en donde la salida del i-ésimo nodo en la capa I se denota como $O_{1,i}$)

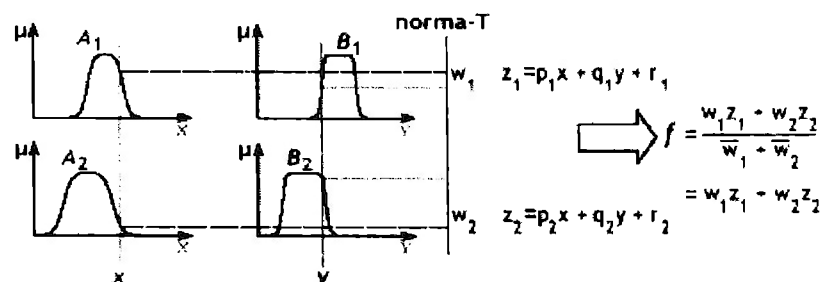


Figura 2.9 Sugeno

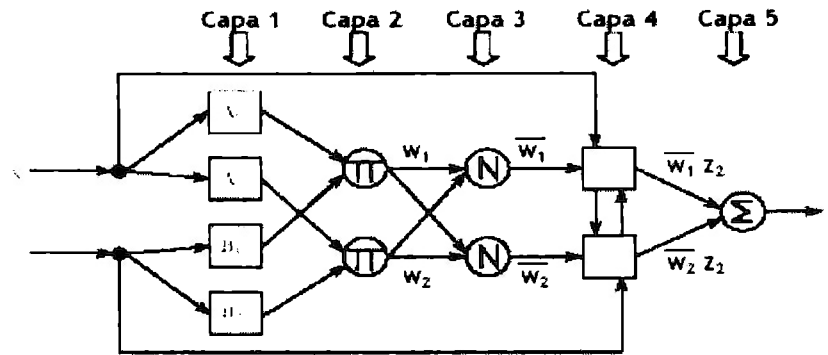


Figura 2.10 Red ANFIS

Capa 1. Cada nodo i en esta capa es un nodo adaptable con una función de nodo definida por:

$$O_{1,i} = \mu_{A_i}(x), \text{ para } i=1, 2, \text{ ó}$$

$$O_{1,i} = \mu_{B_{i-2}}(y), \text{ para } i=3, 4$$

(2.11)

donde x (o y) es la entrada al nodo i , y A_i (o B_{i-2}) es una etiqueta lingüística (tal como “pequeño” o “alto”) asociada con este nodo. En otras palabras, $O_{1,i}$ es el grado de membresía de un conjunto difuso A ($= A_1, A_2$, o $B = B_1, B_2$), el cual especifica el grado en que una entrada (x o y) satisface al cuantificador A . La función de membresía de A , debe ser una función parametrizada que sea continua y derivable a tramos, tal como la función campana generalizada, ver figura 2.5[11]

Capa 2. Los nodos de esta capa son nodos fijos, y se etiquetan con el símbolo Π ; la salida aquí es la operación norma-T *Minimum(MIN)*: $T_{\min}(a,b) = \min(a,b) = a \wedge b$ de todas las señales que entran a él:

$$O_{2,i} = W_i = T(\mu_{A_i}(x), \mu_{B_i}(y)), i = 1, 2.$$

(2.12)

Cada salida de estos nodos representa la intensidad de disparo de una regla. En general, se puede emplear cualquier operador norma-T como función de nodo en esta capa.

Capa 3. Esta capa se compone de nodos fijos etiquetados con N. El i-ésimo nodo calcula la relación de la intensidad de disparo de la i-ésima regla con respecto a la suma de las intensidades de disparo de todas las reglas:

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2}, i = 1,2 \quad (2.13)$$

A cada salida de esta capa se le llama intensidad de disparo normalizada.

Capa 4. Cada nodo i en esta capa es un nodo adaptable con una función definida por:

$$O_{4,i} = \bar{w}_i z_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (2.14)$$

donde w_i es una salida de intensidad normalizada de la capa 3, y $\{p_i, q_i, r_i\}$ es el conjunto de parámetros que define a los polinomios z_i en este nodo. Los parámetros en esta capa son referidos como parámetros del consecuente.

Capa 5. El único nodo presente en esta capa, el cual calcula la salida global como la suma de todas las señales que entran a él:

$$Salida_Global = O_{5,i} = f = \sum_i \bar{w}_i z_i = \frac{\sum_i w_i z_i}{\sum_i w_i} \quad (2.15)$$

De esta forma, lo que se obtiene es una red adaptable, la cual es, funcionalmente, equivalente a un modelo difuso de Sugeno. La estructura mostrada no es única, ya que se pueden hacer otras construcciones equivalentes, por ejemplo combinando las capas 3 y 4, con lo que la red tendría sólo cuatro capas. El sistema completo que se tomó como base para este trabajo, es una arquitectura ANFIS equivalente a un modelo difuso tipo Sugeno de 1er orden

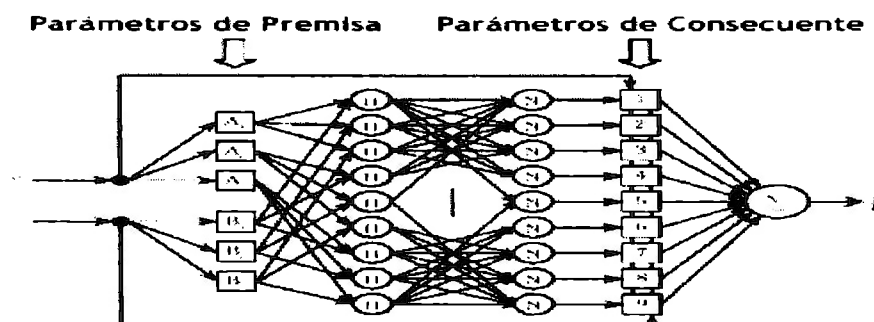


Figura 2.11 Red Sugeno

3. DASARROLLO DEL SISTEMA DIFUSO

Se ha elegido las acciones de interés a pronosticar: TELMEX L y CEMEX A; antes de desarrollar el sistema difuso es importante responder a la pregunta ¿cuál es el negocio de la compañía que tratamos de pronosticar, es decir se puede adelantar escenarios, por ejemplo si TELMEX es una compañía de telecomunicaciones en donde uno de sus negocios es TELBIP, el Pager, y estamos observando que hubo una estrategia de marketing a comprar teléfonos celulares en el sistema “el que llama paga”, en vez de un pager, creeríamos que la acción TELMEX se va para bajo debido a la demanda, pero a su vez analizamos que TELMEX lanzó este marketing a través de “TELCEL” que es una división de TELMEX y se especuló que TELBIP no llegó a tener acuerdos con UNICOM, siendo UNICOM la compañía que vende los pagers mas baratos en el mundo. Por otro lado la competencia directa de TELBIP era BIPER el cual tenía una exclusividad con UNICOM; su resolución de TELMEX colocar en el mercado el sistema celular “el que llama paga”, dando por terminado la industria del pager. Todo lo anterior es un ejemplo clásico, todas estas noticias tienen que ir cerca del experto para poder confiar en el sistema experto, es decir el experto tiene que incluir en la central del algoritmo ciertas alarmas ante eventos inesperados.

3.1 Conocimiento de las empresas que emiten las acciones.

TELMEX es la empresa líder de telecomunicaciones en México, y figura entre los 20 principales operadores a nivel mundial. A partir de su privatización en 1990, TELMEX ha invertido cerca de 28 mil millones de dólares en su modernización y crecimiento; actualmente es una empresa de vanguardia, con una de las infraestructuras tecnológicas más amplias y avanzadas a nivel mundial, con capacidad para transmitir señales de voz, datos e imágenes, con alta calidad y confiabilidad, y certificada con estándares internacionales de calidad y eficiencia. Las acciones de TELMEX cotizan en la Bolsa Mexicana de Valores (TELMEX A y TELMEX L), en el Mercado ADR de Nueva York (TMX) y en NASDAQ (TFONY), en diciembre de 1990, el Gobierno Federal vende las acciones de control de TELMEX a un grupo encabezado por Grupo CARSO, S.A. de C.V., SBC International, Inc., una subsidiaria de SBC Communications Inc. y France Télécom Financière Internationale, una subsidiaria de France Télécom. En abril de 1996, CARSO transfiere sus acciones de TELMEX a Carso Global Telecom, S.A. de C.V. Actualmente, es líder en la provisión de los servicios local, de larga distancia y celular en México. Asimismo, TELMEX provee otros servicios de telecomunicaciones relacionados tales como, el servicio de

directorios telefónicos, transmisión de datos, acceso al servicio de internet, radiolocalización e interconexión a operadores de larga distancia.

CEMEX, el producto que ha vendido la empresa mexicana Cemex por cerca de un siglo no tiene gran ciencia: una mezcla de arcilla y piedra forma el cemento. Lo llamativo es que la compañía latinoamericana no sólo se haya transformado en pocos años en una de las más grandes del mundo, sino el hecho de que esté hoy exportando productos tecnológicos y soluciones basadas en internet. El crecimiento fue espectacular. A mediados de los 80', la cementera se ubicaba en el lugar 19 del ranking mundial y hoy, menos de 23 años más tarde, no sólo es la empresa más rentable del mundo, sino la tercera más grande del planeta -detrás de la francesa Lafarge y la suiza Holcim. Las utilidades que ha tenido Zambrano el dueño de CEMEX se ha visto muy beneficiado y hoy es considerado el segundo hombre más rico de México, después del zar de la empresa de telecomunicaciones Telmex, Carlos Slim.

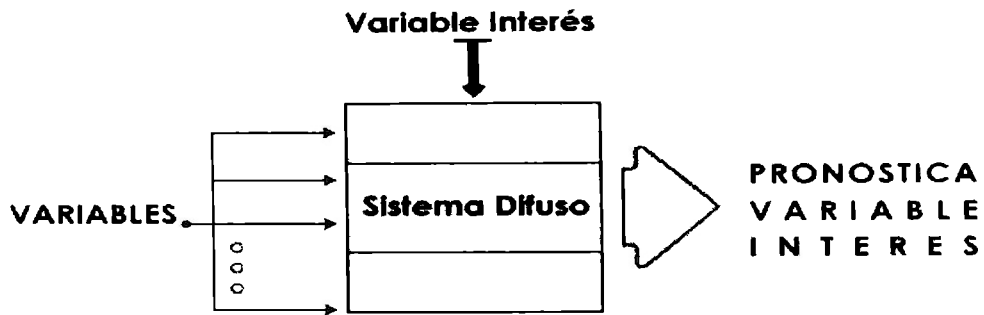
Parte de esta rápida expansión se debe a que Cemex ha enfocado sus esfuerzos en países en vías de desarrollo. De hecho, está presente en Chile, aunque a través de su participación en Cementos Bío Bío, que es de 11.92 %. En mercados como éstos el crecimiento y las utilidades son superiores, porque la mayoría del cemento se vende en sacos para construcciones pequeñas, lo que permite hacer mayores ganancias, y no tanto al por mayor como ocurre en Europa y Estados Unidos. La filosofía Cemex es: aplicar nuevas técnicas y tecnología a los procesos productivos de la compañía, pero también en la relación con proveedores y clientes.

Del Análisis anterior podemos llegar a pensar que tanto TELMEX como CEMEX son empresas vanguardistas que tienen tecnología de punta puesto que su mercado así lo requiere.

3.2 Sistema Difuso.

Nuestro modelo para realizar pronósticos basado en la lógica difusa se puede apreciar en la figura. 3.1, en este esquema observamos que tenemos un bloque que hemos llamado Sistema Difuso, dentro del cual llevamos a cabo muchas operaciones, entran [n] variables junto con su variable de interés a pronosticar y tenemos una salida de la variable de interés pronosticada en el lapso o intervalo de tiempo requerido, para nuestro trabajo de investigación hemos elegido dos variables a pronosticar las cuales son Cemex A y Telmex L, el principal motivo es el

conocimiento de las acciones. Cada una de las acciones las tomaremos por separado, es decir nuestro sistema será monosalida.



3.1 a .Bloque General

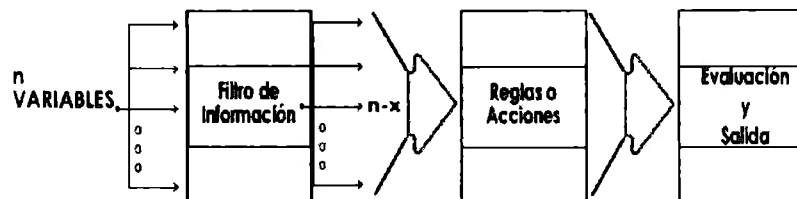


Figura 3.1 b Bloque Interno

Dentro del Sistema difuso tenemos a su vez un esquema interno, en la figura 3.1b se puede apreciar los dos bloques principales los cuales son el filtrado de información, ya que el espectro de información de entrada puede ser lo suficientemente grande, este bloque es muy importante puesto que se trata de trabajar con la información necesaria, aquí se tienen que realizar básicamente la acción de reducir la información para poder manipular fácilmente los datos; el otro bloque: utilizar las reglas o acciones a seguir para el procesamiento de esta información, es decir generar las reglas de lógica difusa para hacer la evaluación de salida del sistema, finalmente tenemos el bloque de evaluación y salida en el cual se realizará la evaluación de todas las reglas generadas o disparadas en nuestro sistema y una salida. La figura 3.2 se puede apreciar con mas detalle el sistema difuso así como una explicación de cada bloque:

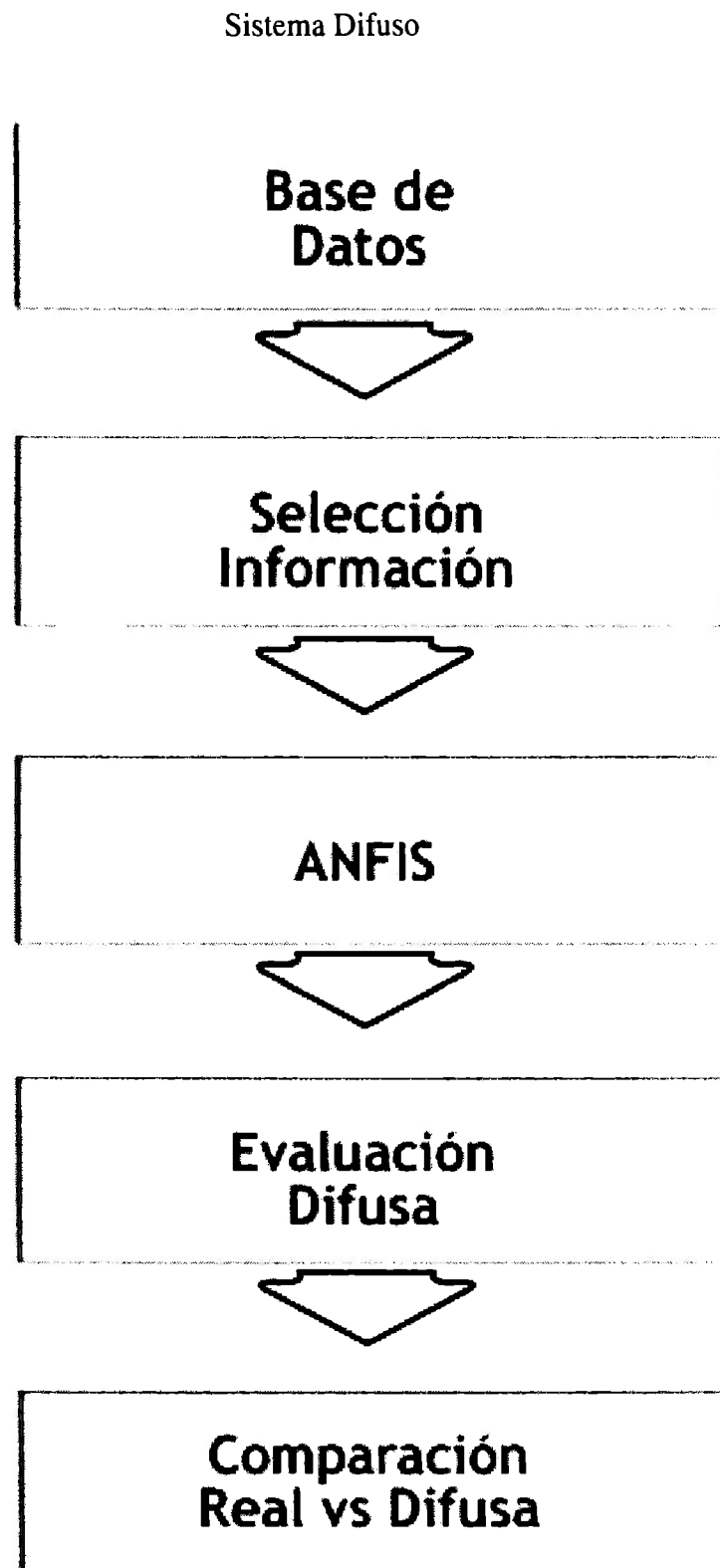


Figura 3.2 Sistema Difuso

3.2.1 Base de Datos

La base de datos con la que nuestro sistema se va a alimentar debe de provenir de una fuente confiable y verídica todo esto con la finalidad de poder tener un pronostico adecuado

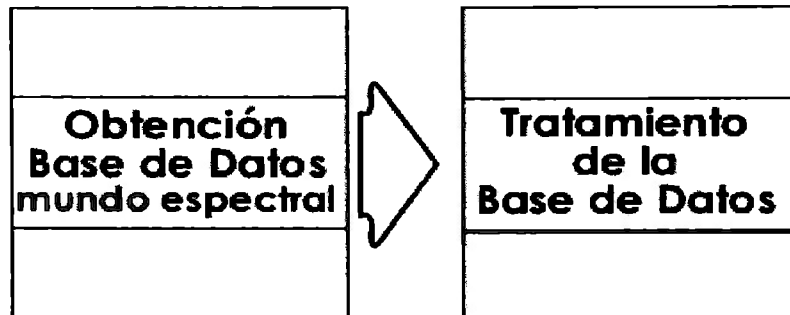


Figura 3.3 Bloque base de datos

Obtención de la base de datos, esta se obtuvo de una fuente segura FINSAT, sus propiedades de esta base de datos son las siguientes:

Propiedad	Concepto.
Ancho	Universo de 31 elementos (las 31 acciones mas bursátiles)
Profundidad	Con mas de 4 años (1400 stock exchang)
Formato	Xls, (facilidad de exportarla a formato txt)
Homologar	Usar fechas y periodos iguales para todas las acciones
Precios de las acciones	Interpolación del max y min, durante el día

Anexo archivo: Basededatos.xls

Tabla 3.1 Base de datos

Se tiene una fecha de inicio de: 15 de Mayo de 1996 y una fecha de termino de 15 de Agosto de 2001, e intervalos de 1 día.

Tratamiento de la base de datos. Una vez que ya tenemos esta base de datos confiable podemos empezar a tratarla, algunos de los pasos que debemos llevar a cabo son los siguientes:

Acciones	Concepto
Estandarización	Unidades de precio estandarizadas.
Normalización (números relativos)	Hacer la normalización a la cotización máxima por acción dentro del mismo periodo.
Formato Txt	Formato TXT delimitado por espacios para poderla exportar a cualquier plataforma
Resolución	6 cifras decimales.

Tabla 3.2 Tratamiento de la Base de datos

Cabe mencionar que en donde no se tuvo la información en el día de la acción por causas técnicas de la base de datos, se optó por hacer una interpolación basada en $t-1$, $t+1$, donde t es el periodo que no se obtuvo información. Con este tratamiento se pretende tener una base de datos en formato txt para poderla usar en cualquier plataforma de software. En el caso que los datos de las columnas sean muy diferentes es decir el precio de una acción de Telmex sea muy diferente a Cemex podemos calcular la desviación de cada columna desde su media esto es $(V - \text{media}(V)) / \text{desviación estándar}(V)$.

3.2.2 Selección de la información

Dentro de este bloque se tienen dos tipos de modelos para poder seleccionar la información

3.2.2.1 Modelo Correlación.

El esquema dentro de este modelo es el siguiente:

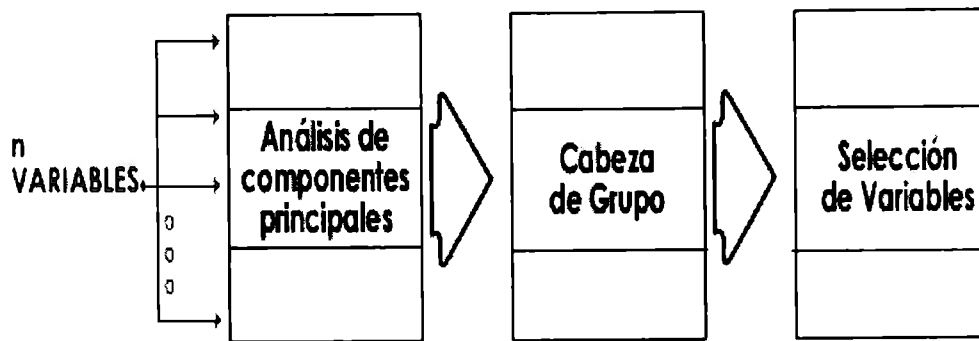


Figura 3.4 Modelo de Correlación

Análisis de componentes principales. Aquí se pretende hacer una correlación entre cada acción y el vector ortogonal que es proporcionado al realizar el análisis de componentes principales, recordando que la estrategia del ACP trata de encontrar la mejor representación posible de dicho objeto, es decir, aquella que es capaz de dar la mayor información de él, el objetivo del análisis de componentes principales es transformar el espacio de representación P en un nuevo espacio, P' , en el que los datos estén no correlacionados (la matriz de covarianza en ese espacio será diagonal). Con otras palabras, se trata de encontrar un nuevo conjunto de ejes ortogonales en el que la varianza de los datos sea máxima. El objetivo final es reducir la dimensionalidad del problema una vez realizada la transformación.(ver en anexo 1 el programa ACP)

Cabeza de Grupo. Teniendo la matriz de correlación entre X (acciones) y Cmp (Componentes Principales) y a partir de este momento ya podemos hacer una correlación entre el vector ortogonal y la acción, esto nos servirá para obtener nuestra cabeza de grupo, lo que se trata es de seleccionar grupos de correlación entre los valores de las acciones relativos y su correlación con el vector ortogonal proporcionado en el ACP; dentro de cada grupo elegimos una variable, la cual llamaremos "Cabeza de Grupo", y con lo cual podemos decir que de alguna manera este puede representar al grupo, si la variable que se seleccionó ya pertenecía a otra cabeza de grupo, se tiene que seleccionar otra variable (la siguiente de correlación alta) en la cabeza de grupo para evitar que se estén repitiendo información. Es importante pensar que si dos variables tienen una correlación elevada en el ACP solo se obtiene variables que están correlacionadas con el vector ortogonal, finalmente las cabezas de correlación nos van a indicar qué variables son las que podemos tomar en cuenta con toda su información para usar el siguiente paso ver figura 3.5

En la tabla 3.3 observamos que se han creado ciertos grupos con correlaciones altas, tenemos la variable con su factor de correlación en el primer componente, notamos que hay muchas correlacionadas, entonces tenemos que seleccionar una representativa que hable por todo el grupo de acciones y esta la llamaremos cabeza de grupo lo mismo ejecutamos para el segundo componente y así sucesivamente, si ya se eligió una variable para el siguiente componente se buscará no repetirla puesto que a final de cuentas se tendría la misma información, todo este procedimiento lo lleva el programa metodo2 ver anexo1 , a partir del componente principal 4 hasta el 31 no existen correlaciones mas altas.

Selección de Variables.

Se han creado la matriz grupo de cabeza con los elementos siguientes, estos grupos de cabezas funcionan tanto para el CEMEX como TELMEX:

GRUPO DE CABEZA
KOFL
VITRO
HYLSAMX BCP

Tabla 3.3 Grupo de Cabeza.

Diagrama de Flujo para encontrar Cabezas de Grupo.

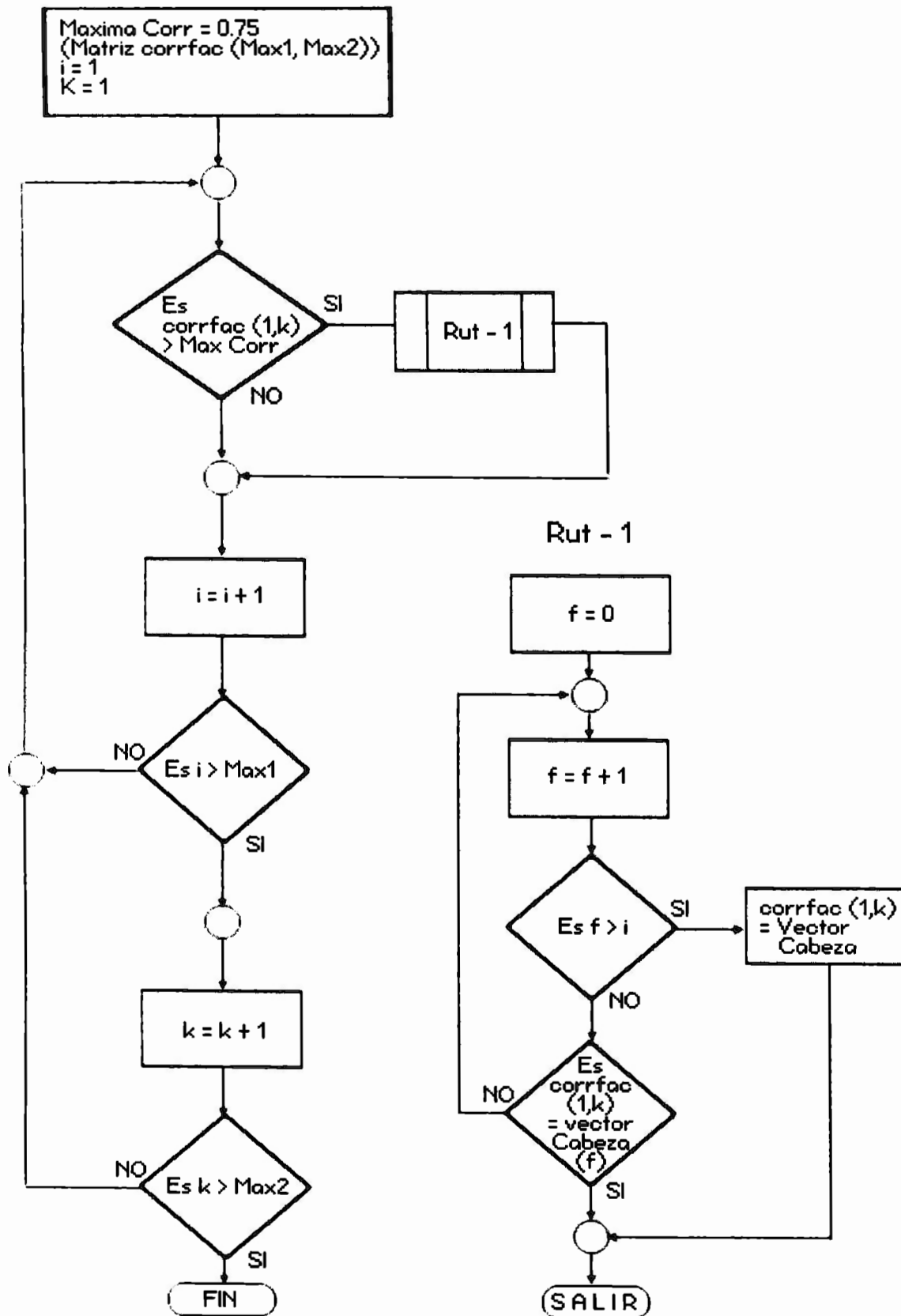


Figura 3.5 Cabeza de Grupo

Correlación entre el componente principal y la variable.

Variable	Componente 1	Componente 2	Componente 3
IPC	-0.6717	0.7051	-0.1574
ALFA A	0.7844	0.5485	0.088
ATY*	0.7705	-0.0909	-0.4541
APASCO*	-0.018	0.7711	-0.1433
CYDSASA A	0.9577	0.0847	-0.0689
CEMEX A	-0.5159	0.7143	-0.1552
CEMEX CPO	-0.7122	0.6028	-0.2743
CEMEX B	-0.6902	0.6052	-0.3168
CIFRA C	0.7883	0.4794	0.2179
CIFRA V	0.1159	0.7444	-0.468
DESC B	-0.5884	0.7779	-0.0093
LECTRA CPO	0.7186	0.6365	0.0575
FEMSA UBD	0.6952	0.3326	-0.3906
GCARSO A1	0.5714	0.3381	-0.5076
GCC B	0.806	0.2817	0.0459
GFB L	-0.7222	0.5118	0.2445
KOF L	0.9704	-0.0859	-0.1689
PEPSIGX CPO	0.4231	0.7566	0.2143
GMODELO C	0.858	-0.1001	0.2794
HYLSAMX BCP	-0.0386	0.2263	0.8276
ICA	0.4291	0.7312	-0.217
KIMBER A	0.8882	-0.0159	0.3199
MASECA B	-0.3745	0.7953	0.4161
MODERNA A	-0.7645	0.5321	0.1215
PE&OLES *	-0.683	0.5921	0.0052
TELMEX L	0.8197	0.4735	-0.0509
TLEVISA CPO	-0.7532	0.5618	0.1678
VITRO	0.0115	0.856	-0.0111
GFNORTE B	0.714	0.6113	0.1732
TAMSA	0.6426	0.6147	0.2093
CINTRA A	0.6155	0.6123	0.1921

Tabla 3.4

3.2.2.2 Modelo ACP

El esquema dentro de este modelo es el siguiente:

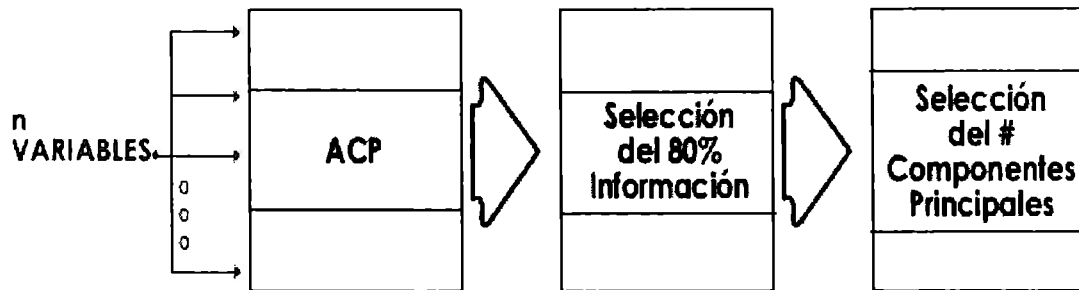


Figura 3.6 Modelo ACP

Análisis de ACP. El grado de bondad de dicha representación vendrá determinada por el porcentaje de información total del objeto puesta de relieve en ella. En algunos casos puede llegar al 90% mientras que en otros no pase del 30%, lo cual nos indicará el grado de orden del objeto (datos) representado (la información que contiene), el objetivo del análisis de componentes es encontrar un nuevo conjunto de ejes ortogonales en el que la varianza de los datos sea máxima. El objetivo final es reducir la dimensionalidad del problema una vez realizada la transformación, utilizamos el ACP para obtener nuestros vectores ortogonales y vamos a trabajar con todo el espectro de información de la base de datos arrojada por el ACP, lo que nosotros aprovechamos mediante este método es la extracción de datos que nos proporcione la mayor información posible del sistema, es decir mediante el ACP podemos saber con X componentes principales tenemos casi toda la información del sistema, es decir con un porcentaje de información en los componentes principales podemos trabajar en nuestro sistema, y estos componentes principales que nos dan un acumulado de información será nuestra base de datos, según los expertos para las ciencias exactas como la astrología ó física se requiere por lo menos un 95% [18] de información acumulada mientras que para otras ciencias es suficiente con un 70%, para nuestro sistema nosotros hemos propuesto un 80%. Por el método de ACP, lo que nos interesa es usar los vectores ortogonales que puedan tener casi toda la información posible y con esto nos podemos dar cuenta que tenemos prácticamente 3, los cuales son:

Información de cada Componente Principal

Componente	Información	Inf. Acumulada
1	44.7395	44.7395
2	31.7721	76.5116
3	8.1356	84.6472
4	4.3136	88.9608
5	2.553	91.5137
6	2.0554	93.5691
7	1.4359	95.005
8	0.8213	95.8263
9	0.7144	96.5407
10	0.599	97.1397
11	0.4994	97.6391
12	0.391	98.03
13	0.2993	98.3294
14	0.2398	98.5691
15	0.2082	98.7773
16	0.1932	98.9705
17	0.1639	99.1344
18	0.1421	99.2764
19	0.1123	99.3888
20	0.1043	99.4931
21	0.0755	99.5686
22	0.0683	99.6369
23	0.0655	99.7025
24	0.0626	99.765
25	0.0568	99.8218
26	0.0475	99.8693
27	0.0403	99.9096
28	0.0395	99.9491
29	0.032	99.9811
30	0.0113	99.9924
31	0.0076	100

Tabla 3.5 Información acumulada ACP

Selección Variables.

En el caso de TELMEX y CEMEX tenemos nuestra matriz con 3 componentes Principaln con mas del 80% de la información del mercado bursátil(acciones).

ANFIS
Componente Principal 1
Componente Principal 2
Componente Principal 3

Tabla 3.6 Componentes para ANFIS

3.2.3 ANFIS.

El esquema general de ANFIS es el siguiente:

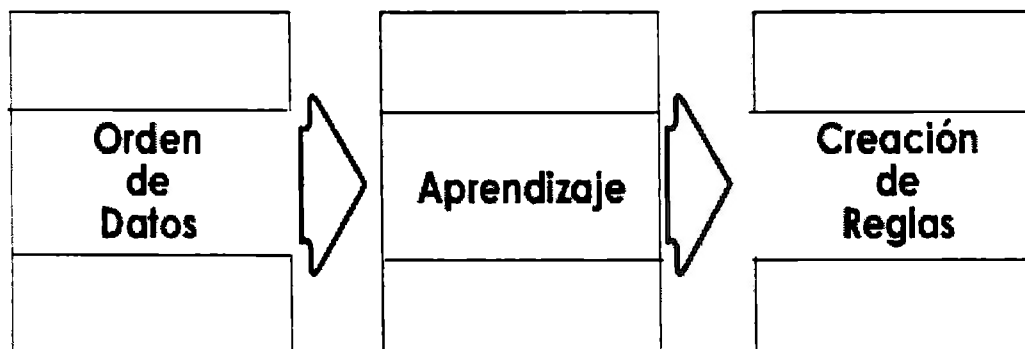


Figura 3.7 Boque ANFIS

Orden de datos. Se tiene que ordenar los datos en forma matricial, crear dos matrices una de entrenamiento y otra para comprobar el funcionamiento de los datos, la matriz de entrenamiento sería con las cabezas de grupo hasta una profundidad de 1300 que es el valor que nosotros hemos

Información de cada Componente Principal

Componente	Información	Inf. Acumulada
1	44.7395	44.7395
2	31.7721	76.5116
3	8.1356	84.6472
4	4.3136	88.9608
5	2.553	91.5137
6	2.0554	93.5691
7	1.4359	95.005
8	0.8213	95.8263
9	0.7144	96.5407
10	0.599	97.1397
11	0.4994	97.6391
12	0.391	98.03
13	0.2993	98.3294
14	0.2398	98.5691
15	0.2082	98.7773
16	0.1932	98.9705
17	0.1639	99.1344
18	0.1421	99.2764
19	0.1123	99.3888
20	0.1043	99.4931
21	0.0755	99.5686
22	0.0683	99.6369
23	0.0655	99.7025
24	0.0626	99.765
25	0.0568	99.8218
26	0.0475	99.8693
27	0.0403	99.9096
28	0.0395	99.9491
29	0.032	99.9811
30	0.0113	99.9924
31	0.0076	100

Tabla 3.5 Información acumulada ACP

Selección Variables.

En el caso de TELMEX y CEMEX tenemos nuestra matriz con 3 componentes Principaln con mas del 80% de la información del mercado bursátil(acciones).

ANFIS
Componente Principal 1
Componente Principal 2
Componente Principal 3

Tabla 3.6 Componentes para ANFIS

3.2.3 ANFIS.

El esquema general de ANFIS es el siguiente:

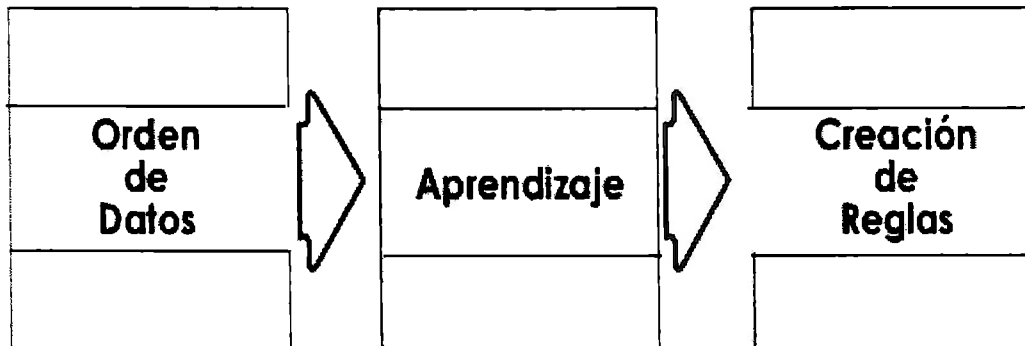


Figura 3.7 Boque ANFIS

Orden de datos. Se tiene que ordenar los datos en forma matricial, crear dos matrices una de entrenamiento y otra para comprobar el funcionamiento de los datos, la matriz de entrenamiento sería con las cabezas de grupo hasta una profundidad de 1300 que es el valor que nosotros hemos

delimitado para el aprendizaje, la matriz de chequeo seria: (1300:1400). Esto lo generamos mediante el programa generadorcito2 ver anexo1:

Para el caso de Modelo de Correlación:

Generadorcito2, matriz(a1,a2,...,acción a pronosticar, acción a pronosticar + t)

Donde a: es la acción cabeza de grupo; t: es el periodo a pronosticar 3 días.

Para el caso de Modelo de ACP:

Generadorcito2, matriz(a1,a2,...,acción a pronosticar, acción a pronosticar + t)

Donde

a: es el componente principal (solo el 80% de la información); t: es el periodo a pronosticar 3 días.

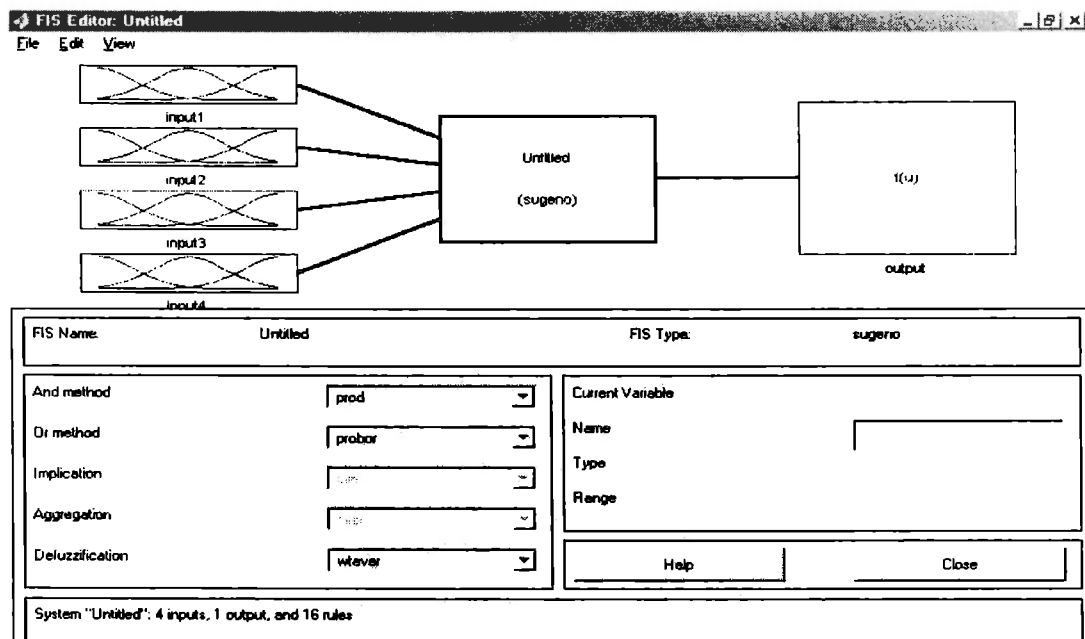


Figura 3.8 Matriz de Entrada

APRENDIZAJE . Es una arquitectura que es funcionalmente equivalente a una base de reglas difusas de tipo Sugeno. Lo que se pretende al usar ANFIS (Adaptive Neuro Fuzzy Inference System). es sintonizar una base de reglas existentes mediante un algoritmo de aprendizaje basado en la recolección de datos de entrenamiento, lo que permite a la base de reglas adaptarse a su entorno. Hacemos un entrenamiento entrada-salida ya que afortunadamente tenemos en nuestra

base de datos una información de aproximadamente NX1400 datos por lo que podemos tomar la decisión de hacer un entrenamiento de 1300 datos para tener 100 para comparar. Este paso es sumamente importante puesto que se generan las reglas difusas para poder formar nuestro sistema de predicciones, es decir entrenamos al sistema con información histórica, entonces para cualquier entrada vamos a tener una salida. Antes de realizar el aprendizaje mediante la función ANFIS se debe generar las matrices de entrenamiento y chequeo.

Las funciones de Membresía al igual que las reglas se pueden modificar para obtener una curva lo mas pegada a la real, se eligió la campana de gauss en las funciones de membresía puesto que tiene un cambio menos brusco en los traslapes, y después de hacer pruebas es la que mejor se ha comportado, también se eligió tener un numero de funciones de membresía bajo (>4) puesto que se puede llegar a un sobre-aprendizaje con lo que el sistema se dispara en puntos de extremos.

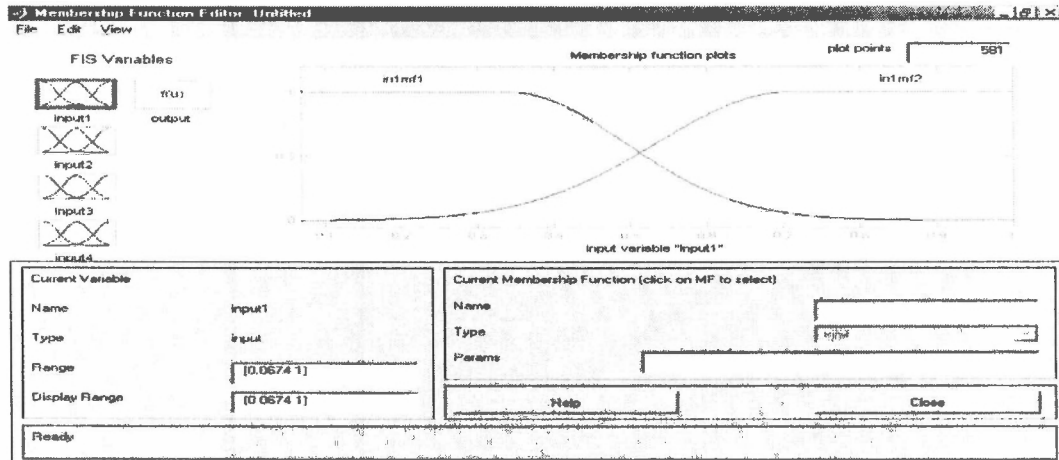


Figura 3.9 Funciones de Membresía

Creación de las reglas. Son las reglas difusas hechas a base del aprendizaje de 1300 puntos, estos aprendizajes llevan tiempo puesto que para 3 funciones de membresía, matrices de (1300X7); se tarda aproximadamente 3 hrs en una pentium a 2 Ghz, cabe señalar que estas reglas son hechas en función al aprendizaje y a los datos de entrada.

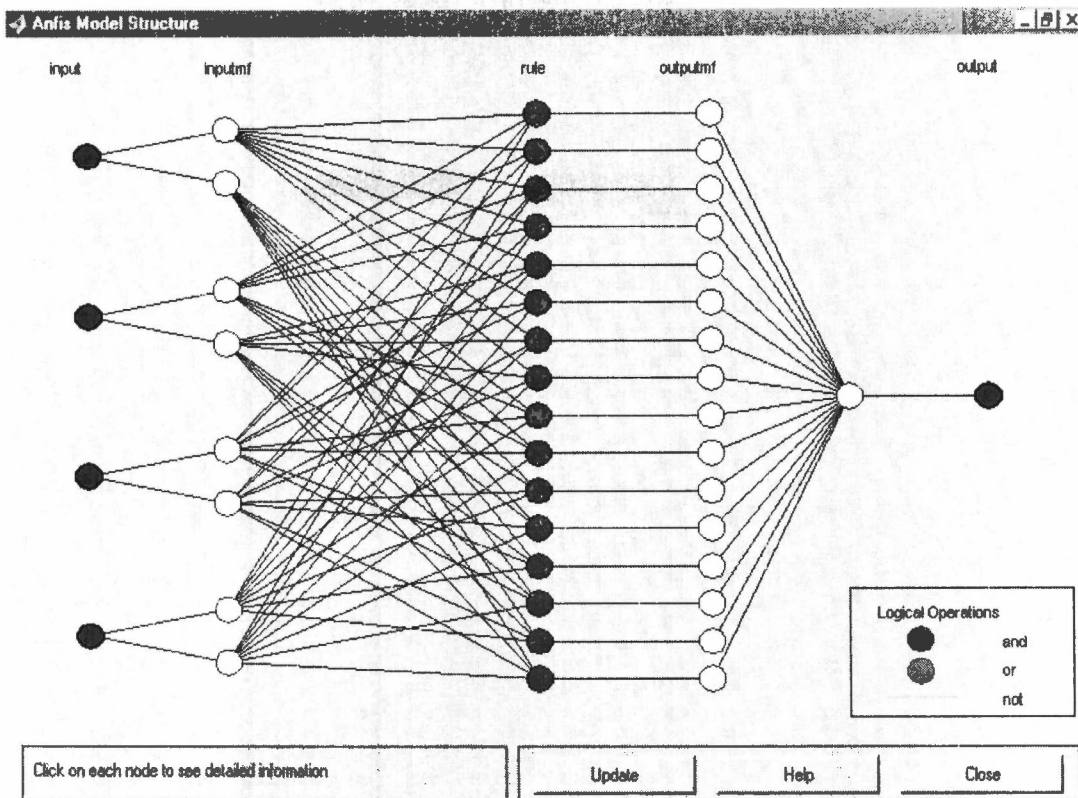


Figura 3.10 Red de Reglas

3.2.4 Evaluación Difusa

Creadas las reglas difusas se puede realizar la evaluación, es decir ya se puede tener una respuesta a cualquier entrada; en nuestro caso en específico tenemos una base de datos de profundidad 1400 con lo que podemos tener nuestras entradas ($n \times 1400$) y nuestra salida Difusa vector (1×1400); donde se tienen 100 puntos a comparar y 1300 en aprendizaje, es decir las reglas en la evaluación se pueden exportar para la evaluación de otro escenario bursátil.

Para el caso de Modelo de correlación:

matriz($a_1, a_2, \dots, \text{acción a pronosticar}, \text{acción a pronosticar} + t$)

donde:

a: es la acción cabeza de grupo; pueden ser después del aprendizaje

t: es el periodo a pronosticar , en nuestro caso tenemos que es 3

Para el caso de Modelo de ACP:

Matriz(a1,a2,...,acción a pronosticar, acción a pronosticar + t)

Donde

a: es el componente principal (solo el 90% de la información);

t: es el periodo a pronosticar 3 días.

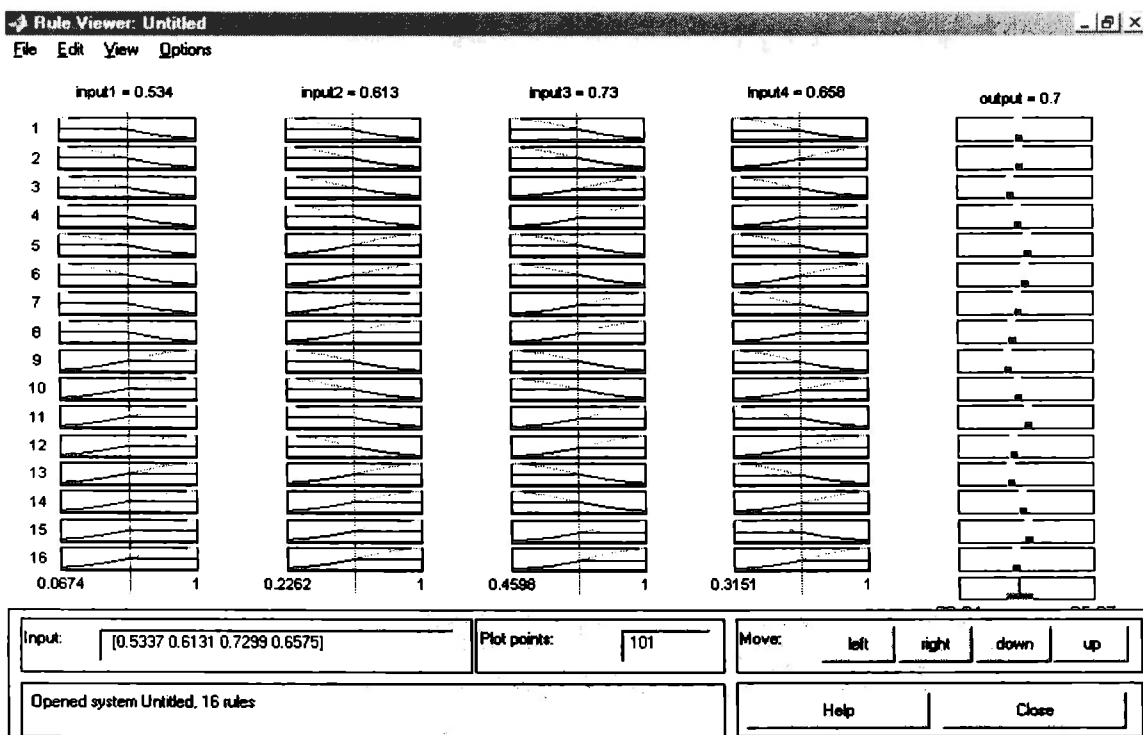


Figura3.11 Salidas Difusas

3.2.5 Comparación del sistema Difuso Vs Real.

Se puede comprobar la efectividad de sistema mediante el error cuadrático medio entre las salidas difusas y los valores reales, es obvio que vamos a tener un error cuadrático medio muy bajo en la parte del aprendizaje es decir en los primeros 1300 puntos y probablemente un error cuadrático medio elevado en los 100 puntos restantes que es en realidad nuestra prueba del sistema, recordando que existirá un error, esta salida solo puede ser una herramienta mas para un financiero en el momento de tomar decisiones, debemos tener por lo menos 22 puntos confiables (lo equivalente a un mes de cotizaciones), es necesario en este paso poder analizar las gráficas de salida, encontrar analogías en tramos y poder dar la explicación del sistema difuso.

Analizar nuestros resultados para obtener una conclusión del pronostico obtenido, recordando que el sistema que se ha propuesto se puede ir alimentando para su aprendizaje mediante índices, factores políticos, factores económicos, variables estadísticas con la finalidad de disminuir el error del pronóstico, resulta en un gráfico normal difícil el poder comparar el error de salida, puesto que realmente nuestro pronóstico va casi pegado al original, cabe mencionar que se tuvieron 1300 puntos de entrenamiento y 100 de prueba. En el siguiente capítulo se explican todos los resultados obtenidos.

Cabe mencionar que el sistema difuso requiere mantenimiento, es decir actualización del sistema y esto puede ser principalmente, ajustando reglas, mayor intervalo de datos para el aprendizaje, modificación de las funciones de membresía.. En la figura 3.12 se pude apreciar el diagrama de flujo para monitorear el pronóstico difuso.

Monitoreo del Pronóstico del sistema difuso

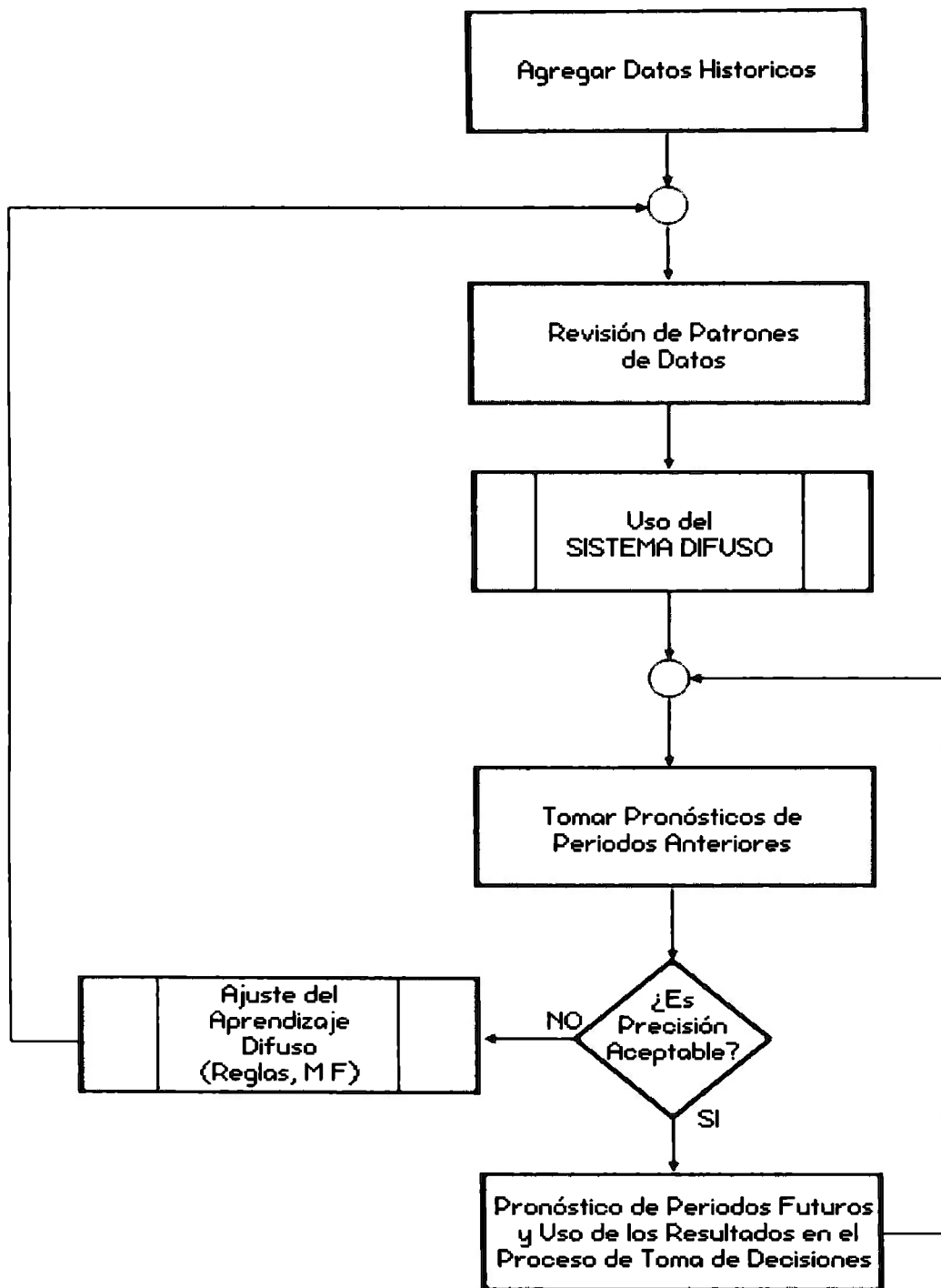


Figura 3.12 Proceso de Actualización del Pronóstico Difuso

4 RESULTADOS

Como se ha mencionando en capítulos anteriores el sistema experto trata de definirse como un sistema basado en los conocimientos que imita el pensamiento de un experto, para resolver problemas en un terreno particular de aplicación; lo que nos lleva primero a pensar, en tratar de familiarizarnos en el espectro de datos que tratamos de descifrar, realmente en nuestro caso como propósito de tesis vamos a enfocarnos a dos tipos de acciones a pronosticar como lo mencionamos en el capítulo anterior, las cuales son: Telmex 'L' y Cemex 'A', estas acciones manejan un porcentaje elevado en las transacciones que se reflejan diario en la bolsa mexicana de valores. Las figuras 4.1 y 4.2 son las gráficas de los precios de las acciones correspondientes, en donde se tiene un intervalo de 1400 puntos o días hábiles para la bolsa Mexicana de valores.

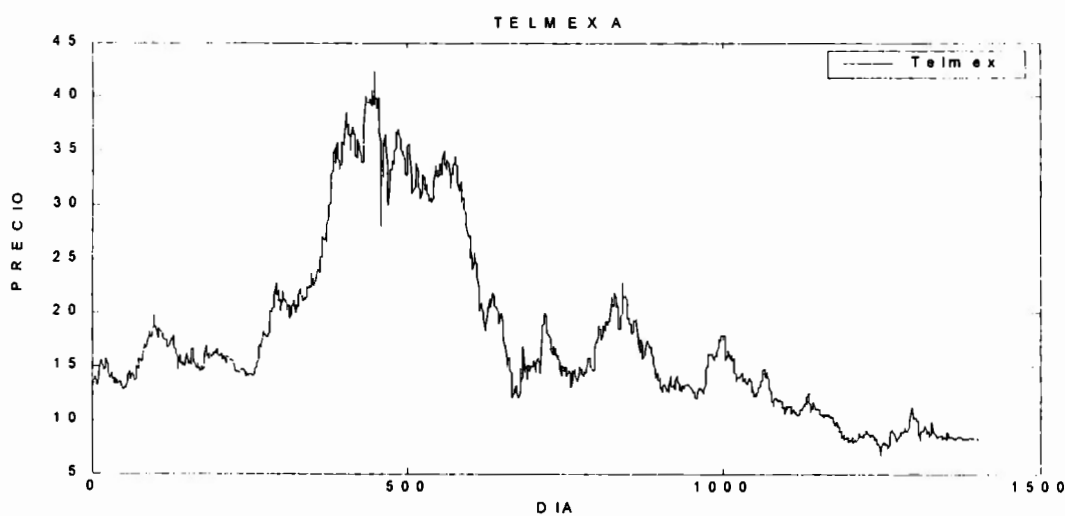


Figura 4.1 Precio de Telmex

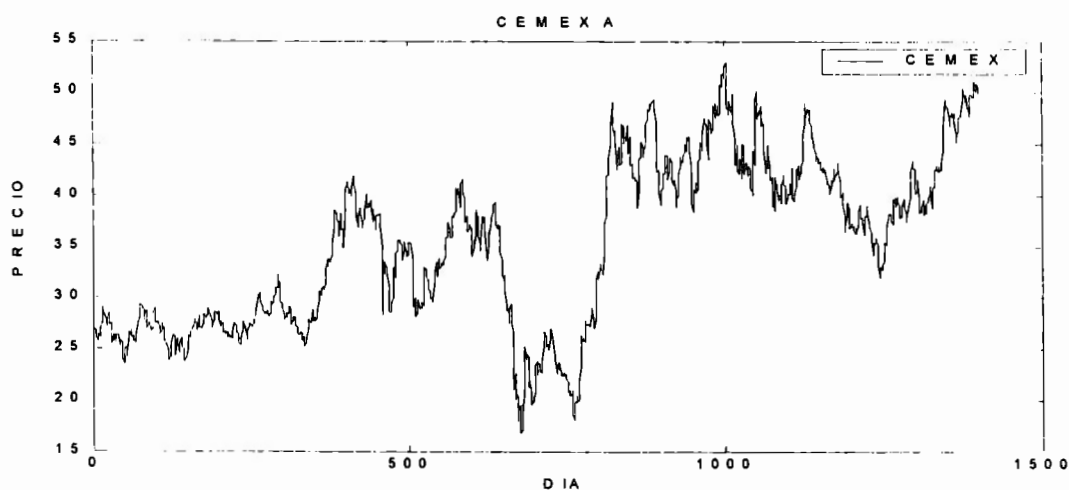


Figura 4.2 Precio de Cemex

En la tabla 4.1 se observa que se tiene una fecha de inicio en nuestra base de datos de: 2-Ene-96 y una final de: 2-Ago-01 así como se tienen 1400 puntos (días apertura de la bolsa) entre estas dos fechas, se ha decidido dejar 1300 puntos para el aprendizaje y 100 puntos para poder comprobar el sistema (elección en base a la experiencia) .

Punto	Fecha	Punto	Fecha
1 (INICIO)	2-Ene-96	1000	24-Dic-99
250	23-Dic-96	1250	28-Dic-01
500	29-Dic-97	1400(FIN)	2-Ago-01

Tabla 4.1

De la figuras 4.1 y 4.2 podemos observar que no son estacionarias, no existe una ciclicidad bien definida, se tiene bastantes puntos de inflexión por lo que significa que nuestras acciones a estudiar son bursátiles; en la tabla 4.2 se muestra algunas razones estadísticas.

Concepto	TELMEX	CEMEX
Media:	17.71	35.25
Desviación Estándar:	8.236	8.15
Precio Max:	42.35	53
Precio Min	6.6	16.7
Rendimiento Max.	.2226	.1591
Rendimiento Min.	-.2429	-.2134

Tabla 4.2 Valores Estadísticos

Es difícil el proponer un método exacto tradicional, el hacer un modelo matemático para que pueda dar como salida cualquiera de las figuras 4.1 o 4.2, una manera es seccionar la gráfica para su estudio especializado, el problema radica en que muchos puntos de inflexión, tarda en responder cualquier modelo a la estabilización. El peligro en el uso de los métodos tradicionales consiste en que comprenden la manipulación de datos históricos para generar el pronóstico, sería

un procedimiento válido si el pronóstico es para el futuro cercano, pero se hace más sospechosa al incrementarse el horizonte del pronóstico. No existe una forma sencilla tradicional para poder hacer un pronóstico, si se piensa en seccionar la curva se tiene que realizar correlaciones de hechos políticos, factores económicos del pasado con el presente, en cada sección, lo cual resulta difícil. El manejar un método de series de tiempo como atenuación exponencial solo nos lleva a pensar si la acción tuviera un comportamiento o tendencia bien definida.

Lo difícil es tener una clara percepción en donde se tienen tendencias, como se mencionó en capítulos anteriores se puede hacer una combinación de métodos matemáticos por tramos si de alguna manera se sabe la tendencia de la curva.

Se realizará básicamente tres experimentos los cuales consisten en: utilizar el pronóstico por métodos tradicionales a 4 días para las acciones TELMEX y CEMEX; estos métodos tradicionales son: Regresión lineal, Regresión múltiple y Promedios móviles, ya que fueron los métodos tradicionales que mejor respondieron al pronóstico, y posteriormente comparará con la señal original, estos se pueden apreciar en la sección 4.1. En sección 4.2 se realizará el pronóstico de las acciones Telmex y Cemex a 4 días por la lógica difusa mediante dos modelos ACP y Correlación para ambas acciones; posteriormente comparará el EMC y PME, con la señal original, para poder obtener el mejor método que ajusta al pronóstico. En la sección 4.3 se tomará el pronóstico a 12 días y se realizará por el método tradicional que mejor se ajuste a cada acción, así como también se utilizará la lógica difusa por el modelo que mejor ajusta al pronóstico y se comparará los EMC y PME arrojados para poder tener una conclusión satisfactoria.

4.1 Pronóstico por métodos tradicionales

Se seleccionó tres métodos tradicionales que arrojan un error cuadrático medio bastante bajo. En algunos tramos de la curva probablemente ajusta mejor uno de los tres métodos tradicionales. Para los promedios móviles se optó por un intervalo de pronóstico de 7, solo los 15 anteriores se utilizan para calcular el nuevo promedio (tablas 4.3 y 4.5). Para la regresión múltiple se seleccionó las variables por el método de correlación (tabla 3.3) es decir obtener la cabeza de grupo, ya que de esta manera se puede representar el espectro. Para la regresión simple solo se usa la acción a pronosticar (línea recta), en los métodos de regresión y regresión múltiple se utilizan los 1300 puntos para su obtención de coeficientes ($a_0 + a_1X_1 + a_2X_2 + a_3X_3 + \dots$) y los 100 últimos puntos es la generalización del sistema.

Pronóstico de TELMEX por modelos Tradicionales.

Método Usado	Acción a pronosticar	Datos Históricos	Generalización	Intervalo	Variables Utilizadas	Coefficientes Regresión
Regresión Múltiple (ROJO)	TELMEX	1300	100	4	TELMEX KOFL VITRO HYLSAMX (Tabla 3.3)	A0=0.0069 A1=0.0086 A2= -.0104 A3=0.0012 A4=0.9873
Regresión (AMARILLO)	TELMEX	1300	100	4	TELMEX	A0=0.0029 A1=0.9931
Promedios Móviles (AZUL)	TELMEX	15	100	7	TELMEX	Promedio Móvil Lineal

Tabla 4.3 Métodos tradicionales para Telmex

Pronóstico de CEMEX por modelos tradicionales

Método Usado	Acción a pronosticar	Datos Históricos	Generalización	Intervalo	Variables Utilizadas	Coefficientes Regresión
Regresión Múltiple (ROJO)	CEMEX	1300	100	4	CEMEX KOFL VITRO HYLSAMX (Tabla 3.3)	A0= 0.0220 A1= -0.0103 A2= 0.0173 A3= 0.0016 A4= 0.9564
Regresión (AMARILLO)	CEMEX	1300	100	4	CEMEX	A0= 0.0142 A1= 0.9792
Promedio Móvil (AZUL)	CEMEX	15	100	7	CEMEX	Promedio Móvil Lineal

Tabla 4.5 Métodos tradicionales para Cemex

4.1.1 Pronóstico TELMEX por método tradicional

La figura 4.3 nos muestran los pronósticos por métodos tradicionales de las acciones TELMEX, se tiene: la señal original (verde) y las salidas de los métodos tradicionales : promedios móviles (azul), regresión múltiple (rojo) y regresión (amarillo); lo que se observa es que la regresión en diferentes puntos de inflexión se aleja de la señal original.

En la figura 4.3 no se llega a percibir el error entre cada método por esta razón es conveniente hacer la ampliación de un tramo del pronóstico de la acción (figura 4.4), en donde se percibe la separación de las curvas reflejado en el EMC y PME; de la gráfica se puede observar que la regresión múltiple está muy pegada a la real, mientras que la regresión lineal con un poco de separación con respecto a la real y los promedios móviles como era de esperarse tarda en suavizar la curva.

En la tabla 4.4 podemos observar que el método tradicional que mejor ajusta para pronosticar la acción de TELMEX es la regresión múltiple, ya que tiene un EMC en los 100 puntos a comprobar mas bajo y su PME de los 100 últimos puntos esta ligeramente sobrestimado. Cabe mencionar que el método tradicional de promedios móviles tiene un error absoluto mas grande que la regresión ó la regresión múltiple debido al ajuste de sus valores de inicio, el método de promedios móviles es el mas fácil de usar, no requiere mucho tiempo de ensayo.

Errores de TELMEX por métodos tradicionales

Método	EMC (1300) Aprendizaje	EMC (100) Generalización	EMC TOTAL (1400)	Max Error absoluto	PME 1300	PME 100
Regresión Múltiple	9.2236e-006	9.0950e-006	8.5712e- 006	0.0065	-3.5925e- 004	-0.0011
Regresión	1.7755e-005	2.1212e-005	1.7994e- 005	0.0040	-8.6815e- 004	-0.0072
Promedio Móviles	4.8577e-004	6.8682e-005	4.5590e- 004	0.2429	6.5565e- 004	-0.0095

Tabla 4.4 Errores de aproximación de Telmex a 4 días por métodos tradicionales.

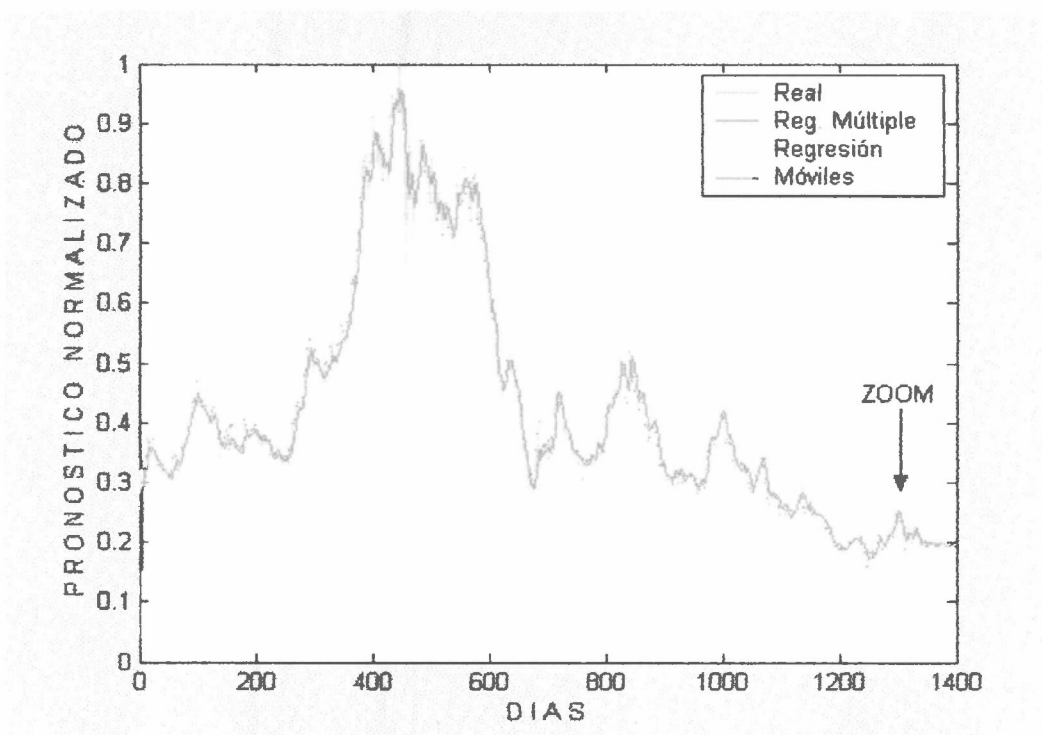


Figura 4.3 Telmex por método tradicionales a 4 días

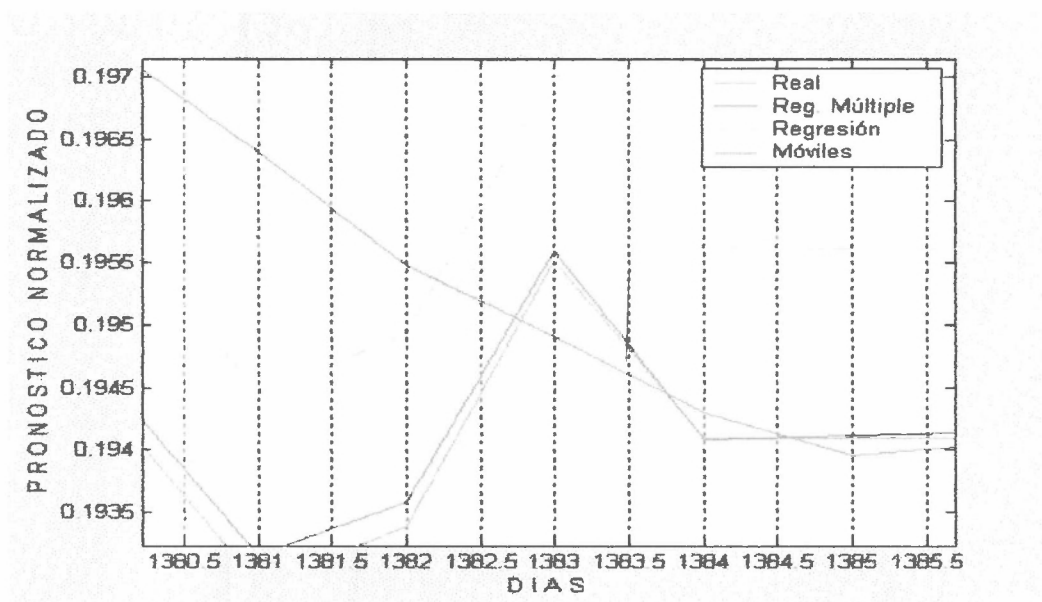


Figura 4.4 Zoom de Telmex por método tradicionales a 4 días.

4.1.2 Pronóstico CEMEX por método tradicional.

La figura 4.5 nos muestran los pronósticos por métodos tradicionales de las acciones CEMEX, se tiene: la señal original (verde) y las salidas de los métodos tradicionales : promedios móviles (azul), regresión múltiple (rojo) y regresión (amarillo); lo que se puede observar es que para esta acción los tres métodos van muy pegados al real sobre todo en los diferentes puntos de inflexión.

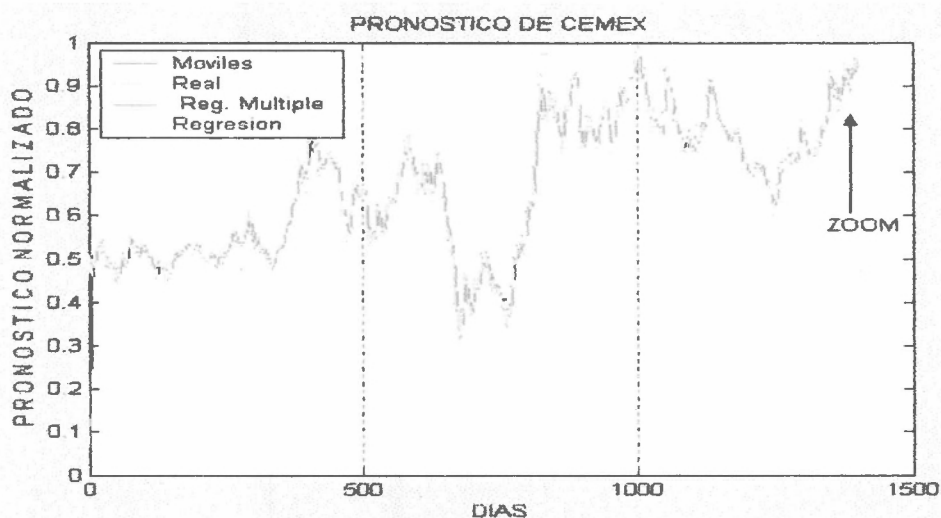


Figura 4.5 Cemex por métodos tradicionales a 4 días

En la figura 4.6 se percibe la separación de las curvas reflejado en el EMC y PME; de la gráfica se puede observar que la regresión lineal está muy pegada a la real, mientras que la regresión múltiple con un poco de separación con respecto a la real y los promedios móviles en diferentes tramos se separan mucho de la real.

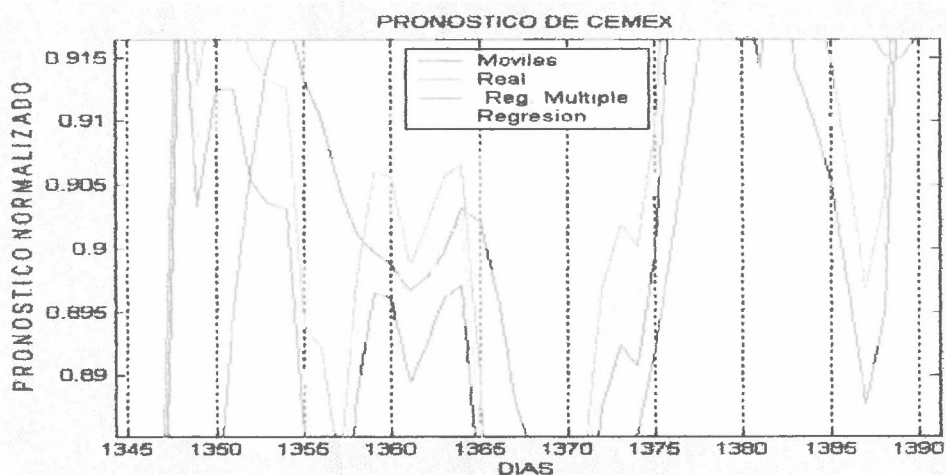


Figura 4.6 Zoom de Cemex por método tradicionales a 4 días

En la tabla 4.6 podemos observar que el método tradicional que mejor ajusta para pronosticar la acción de CEMEX es la regresión lineal ya que tiene un EMC en los 100 puntos a comprobar muy bajo y su PME de los 100 últimos puntos esta ligeramente sobrestimado.

Al igual que el pronóstico de TELMEX por método tradicionales los promedios móviles tiene un error absoluto mas grande que la regresión ó la regresión múltiple debido al ajuste de sus valores de inicio, este método es el mas fácil de usar, no requiere mucho tiempo de ensayo, es la ventaja práctica de este método sobre los anteriores.

Errores de CEMEX por métodos tradicionales

Método	EMC (1300) Aprendizaje	EMC (100) Generalización	EMC TOTAL (1400)	Max Error absoluto	PME 1300	PME 100
Regresión Múltiple	1.6200e-005	5.6566e-005	1.9072e-005	0.0124	-0.0023	0.0072
Regresión	9.9917e-006	1.9788e-005	1.5330e-005	0.0076	-0.0022	0.0040
Promedio Móviles	9.1850e-004	4.6587e-004	8.8617e-004	0.4001	0.0023	0.0045

Tabla 4.6 Errores de aproximación de Cemex a 4 días por métodos tradicionales

4.2 Pronóstico Sistema Difuso

Como se mencionó en el capítulo 3, en nuestro método difuso se tiene dos tipos de modelos el modelo Correlación y el modelo ACP ya que su principal diferencia es la selección de la información al sistema. (Ver la sección 3.2 .2). En ambos modelos tenemos que los primeros 1300 puntos serán para el aprendizaje del sistema ó la creación de la reglas difusas y los últimos 100 puntos serán para la generalización del sistema ó la salida del pronóstico.

4.2.1 Pronóstico Telmex por método Correlación

En la tabla 4.7 se puede observar que se seleccionó la cabeza de grupo conforme a la tabla 3.3, teniendo el tipo de función de membresía en forma de campana ya que fue la que mejor se ajustaba a la salida del aprendizaje, el pronóstico va a ser a cuatro días.

Pronóstico de TELMEX por el modelo Correlación a 4 días

Selección Información	Acción a pronosticar	Aprendizaje	Generalización	Cabeza de Grupo	Tipo de MF	#MF
Correlación	TELMEX (AZUL) Figs. 4.7 y 4.8	1300	100	KOFL VITRO HYLSAMX TELMEX (Tabla 3.3)	CAMPANA A	4,3,3,4

Tabla 4.7 Modelo Correlación a 4 días para Telmex

La figura 4.7 es el pronóstico por el método de la lógica difusa utilizando el modelo correlación para TELMEX; se tiene: señal original (verde) y la salida del pronóstico por el método de lógica difusa (azul), en esta figura no se aprecia con detalle el error, pero lo que se puede obtener en estas gráficas son los puntos colapsados, es decir en donde existen puntos verdes o azules encimados, son en estos puntos donde se está separando el pronóstico. Podemos observar en estas gráficas que en todos los puntos máximos o mínimos responde el sistema de lógica difusa tratando de suavizar la curva, es decir no exagerando estos puntos y obviamente existe mas exactitud en el tramo de aprendizaje.

La figura 4.8 es la ampliación en la zona del pronostico de la acción de TELMEX por el médelo de correlación, este tipo de gráfica nos sirve para poder observar con mas detalle como trata de suavizar cada método con respecto a la señal original.

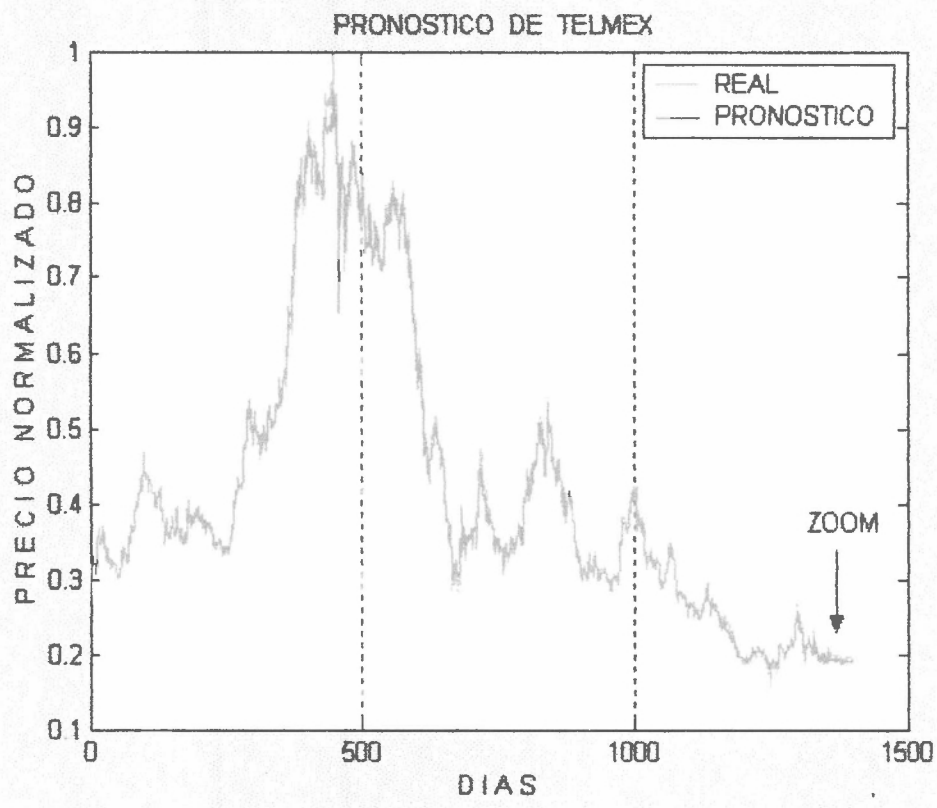


Figura 4.7 Telmex a 4 días por Correlación

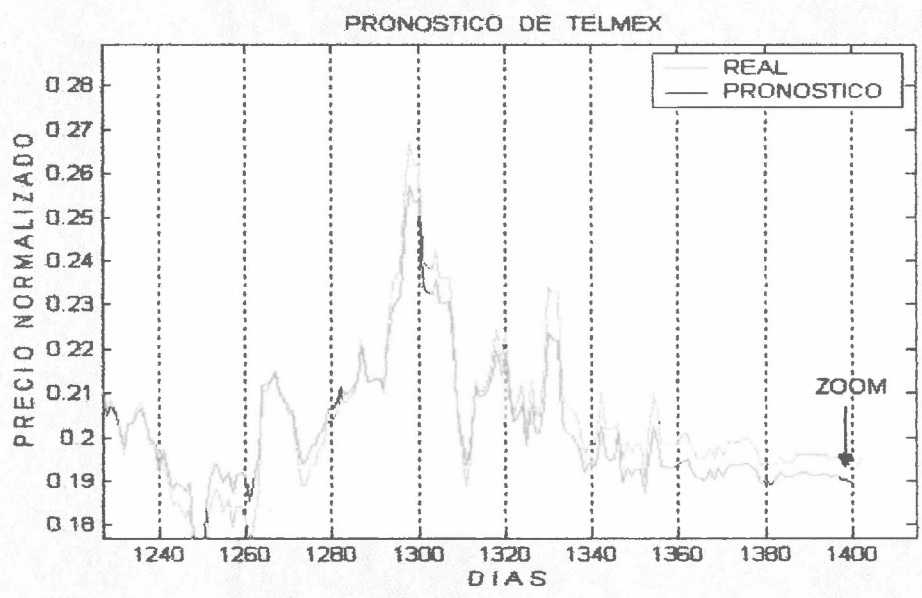


Figura 4.8 Zoom de Telmex a 4 días por correlación

De la tabla 4.8 se puede observar los errores generados por el sistema difuso en el modelo correlación, en todos los casos son menores al generado en la tabla 4.4 por los métodos tradicionales, para ser mas específico el método tradicional que mejor pronosticó para la acción TELMEX fue la regresión múltiple, comparado con el sistema difuso por correlación su error es mas elevado.

Errores de TELMEX por modelo Correlación a 4 días

Acción	EMC (Primeros 1300)	EMC (Últimos 100)	EMC TOTAL (1400)	Max Error absoluto	PME (Primeros 1300)	PME (Últimos 100)
TELMEX	7.1791e- 006	2.6556e- 006	6.8506e- 006	0.0533	-8.5676e-005	.000224

Tabla 4.8 Errores de aproximación de Telmex a 4 días por método Correlación

4.2.2 Pronóstico Telmex por método ACP

En la tabla 4.9 se puede observar que se seleccionó tres componentes principales conforme a la tabla 3.6, teniendo el tipo de función de membresía en forma de campana ya que fue la que mejor se ajustaba a la salida del aprendizaje.

Pronóstico de TELMEX a 4 días por el modelo ACP

Selección Información	Acción a pronosticar	Aprendizaje	Generalización	Componentes Principales	Tipo de MF	#MF
ACP	TELMEX (AZUL) Figs. 4.9 y 4.10	1300	100	3 (Tabla 3.6)	Campana	4,3,3,4

Tabla 4.9 Método ACP para Telmex a 4 días

La figura 4.9 es el pronóstico por el método de la lógica difusa utilizando el modelo ACP para TELMEX; se tiene: señal original (verde) y la salida del pronóstico por el método de lógica difusa (azul), en esta figura no se aprecia con detalle el error, pero lo que se puede obtener en estas gráficas son los puntos colapsados, es decir en donde existen puntos verdes o azules encimados, son en estos puntos donde se está separando el pronóstico. La figura 4.10 es la ampliación en la zona del pronóstico de la acción de TELMEX por el método de ACP, esta gráfica nos sirve para poder observar con mas detalle el error generado entre la salida difusa y la real.

De la tabla 4.10 se puede observar los errores generados por el sistema difuso en el modelo ACP, en todos los casos son menores al generado en la tabla 4.4 por los método tradicionales, para ser mas específico el método tradicional que mejor pronosticó para la acción TELMEX fue la regresión múltiple, comparado con el sistema difuso por ACP su error es mas elevado del pronóstico tradicional.

Errores de TELMEX a 4 días por modelo ACP

Acción	EMC (1300)	EMC (100)	EMC TOTAL (1400)	Max Error Absoluto	PME 1300	PME 100
TELMEX	5.9924 e-006	2.1588e- 006	5.5844e- 006	0.0581	-8.1743e-005	-9.7723e- 004

Tabla 4.10 Errores de aproximación de Telmex por método ACP

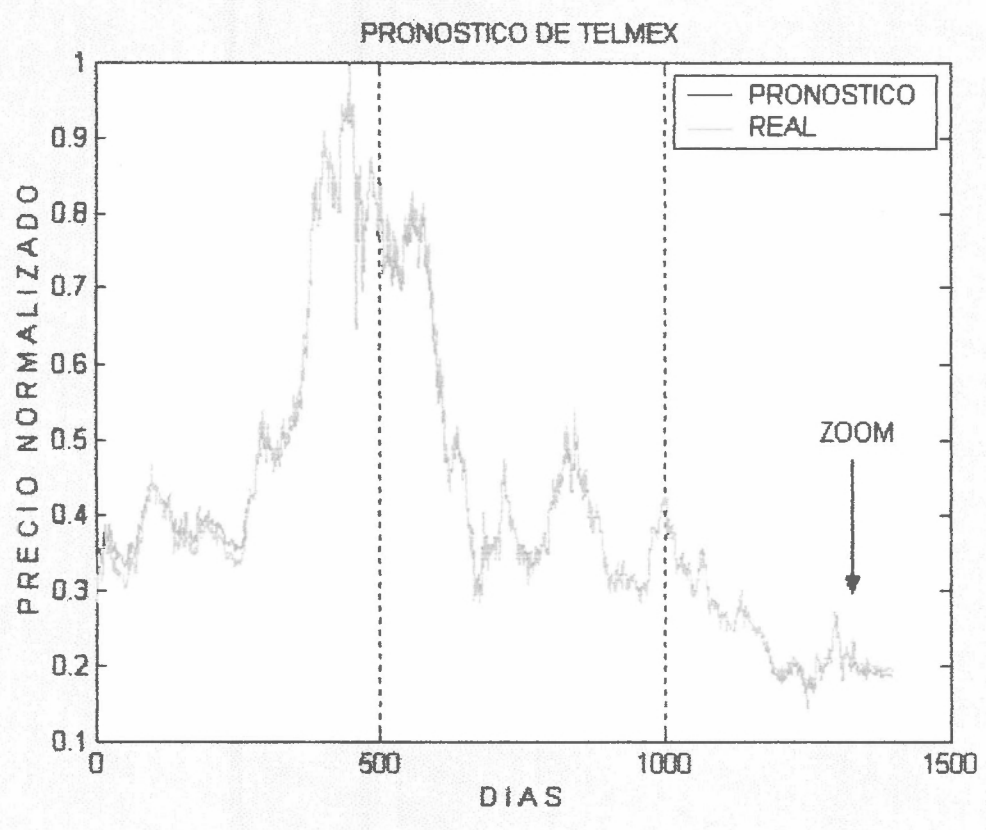


Figura 4.9 Telmex a 4 días por ACP

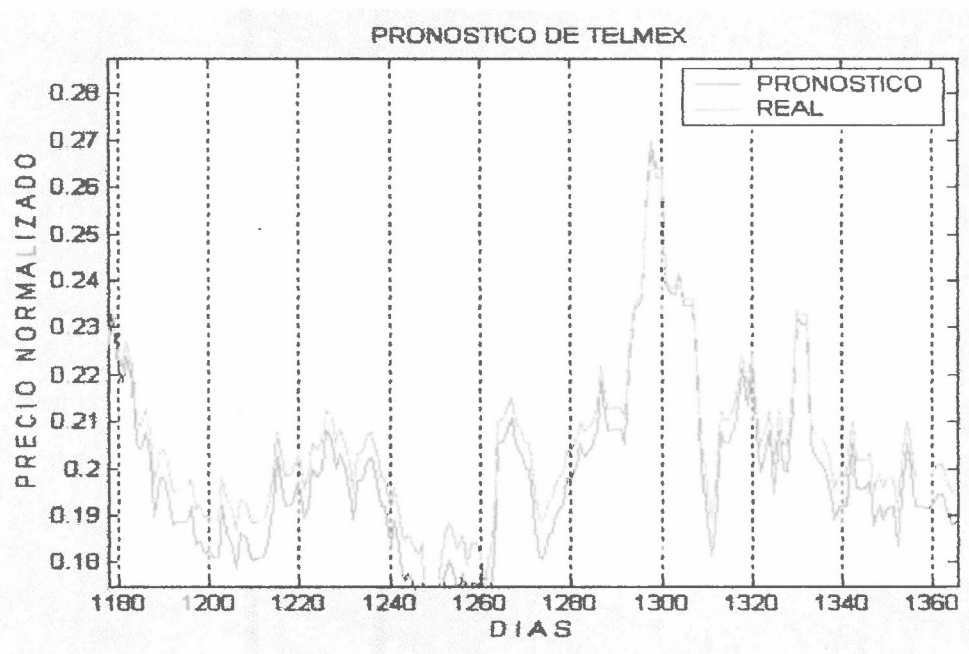


Figura 4.10 Zoom de Telmex a 4 días por ACP

4.2.3 Pronóstico Cemex por método Correlación.

En la tabla 4.9 se puede observar que se seleccionó las cabezas de grupo conforme a la tabla 3.3, teniendo el tipo de función de membresía en forma de GAUSS ya que fue la que mejor se ajustaba a la salida del aprendizaje.

Pronóstico de CEMEX a 4 días por el modelo Correlación

Selección Información	Acción a pronosticar	Aprendizaje	Generalización	Cabeza de Grupo	Tipo de MF	#MF
Correlación	CEMEX (AZUL) Figs 4.11 y 4.12	1300	100	KOFL VITRO HYLSAMX CEMEX (Tabla 3.3)	Gauss	4,3,3,4

Tabla 4.11 Método Correlación a 4 días para Cemex.

La figura 4.11 es el pronóstico por el método de la lógica difusa utilizando el modelo correlación para CEMEX; se tiene: señal original (verde) y la salida del pronóstico por el método de lógica difusa (azul). La figura 4.12 es la ampliación en la zona del pronóstico de la acción de CEMEX por el modelo de correlación, este tipo de gráfica nos sirve para poder observar con mas detalle como trata de suavizar cada método con respecto a la señal original. De la tabla 4.12 se puede observar los errores generados por el sistema difuso en el modelo correlación, en todos los casos son menores al generado en la tabla 4.6 por los métodos tradicionales; el método tradicional que mejor pronosticó para la acción CEMEX fue la regresión lineal, comparado con el sistema difuso por correlación su error es mas elevado.

Errores de CEMEX a 4 días por modelo Correlación

Acción	EMC (1300)	EMC (100)	EMC TOTAL (1400)	Max Error absoluto	PME 1300	PME 100
CEMEX	8.9725e-006	.000017	1.4337e-005	0.0651	-0.0021	.000383

Tabla 4.12 Errores de aproximación de Cemex por método Correlación

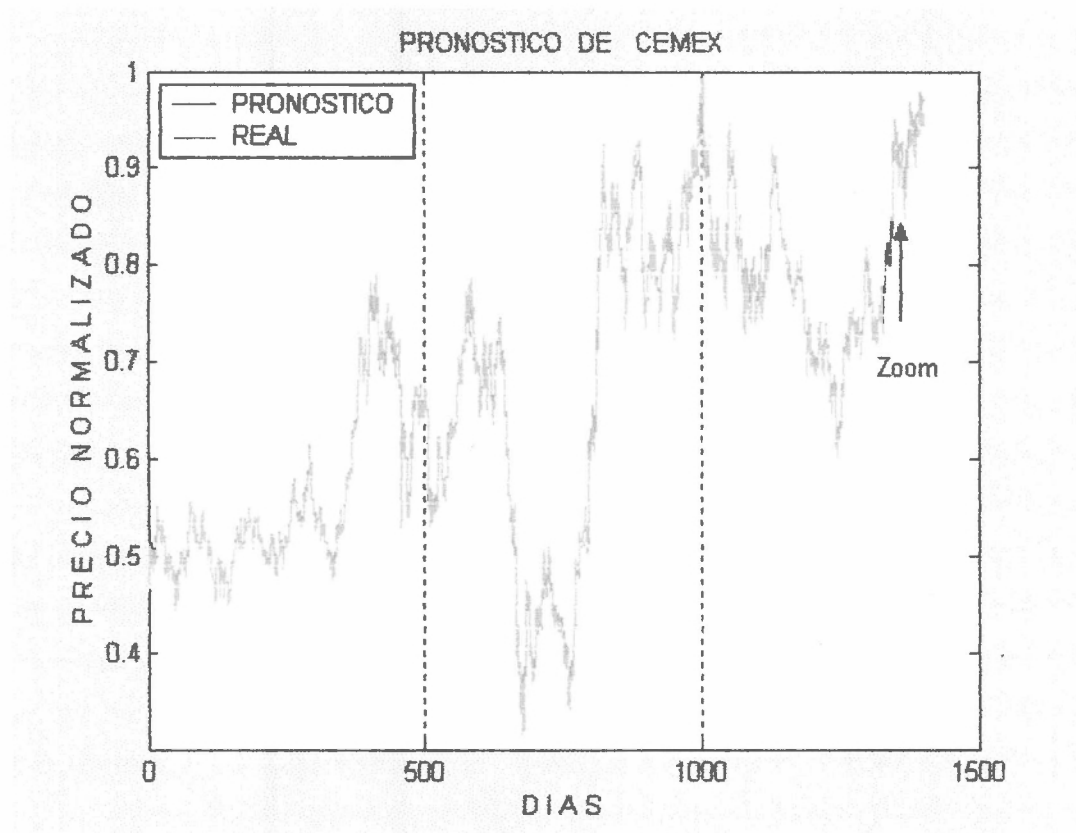


Figura 4.11 Cemex a 4 días por correlación

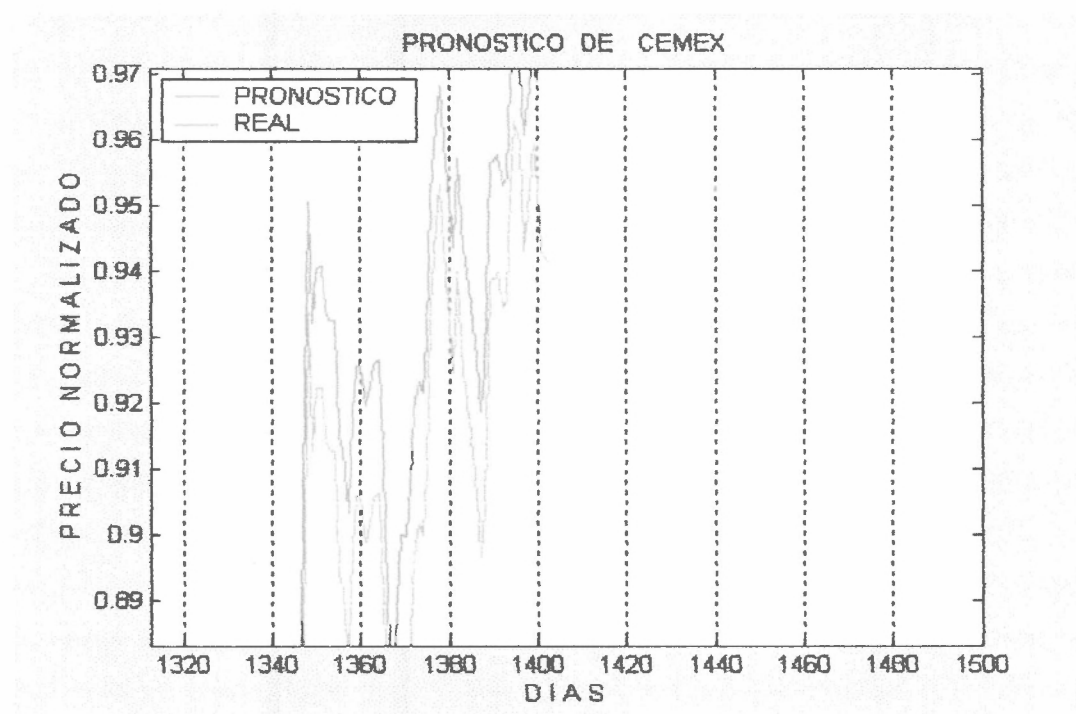


Figura 4.12 Zoom de Cemex a 4 días por correlación

4.2.4 Pronóstico Cemex por método ACP

En la tabla 4.13 se puede observar que se seleccionó tres componentes principales conforme a la tabla 3.6, teniendo el tipo de función de membresía en forma de campana ya que fue la que mejor se ajustaba a la salida del aprendizaje.

Pronóstico de Cemex por el modelo ACP

Selección Información	Acción a pronosticar	Aprendizaje	Generalización	Componentes Principales	Tipo de MF	#MF
ACP	CEMEX (AZUL) Figs 4.13 y 4.14	1300	100	3 (Tabla 3.6)	Campana	4,3,3,4

Tabla 4.13 Método ACP a 4 días para Cemex.

La figura 4.13 es el pronóstico por el método de la lógica difusa utilizando el modelo ACP para CEMEX; se tiene: señal original (verde) y la salida del pronóstico por el método de lógica difusa (azul), en esta figura no se aprecia con detalle el error, pero lo que se puede obtener en estas gráficas son los puntos colapsados, es decir en donde existen puntos verdes o azules encimados, son en estos puntos donde se está separando el pronóstico. Podemos observar en estas gráficas que en todos los puntos máximos o mínimos responde el sistema de lógica difusa tratando de suavizar la curva, es decir no exagerando estos puntos

La figura 4.14 es la ampliación en la zona del pronóstico de la acción de CEMEX por el método de ACP, este tipo de gráfica nos sirve para poder observar con mas detalle el error.

De la tabla 4.14 se puede observar los errores generados por el sistema difuso en el modelo ACP, en todos los casos son menores al generado en la tabla 4.6 por los método tradicionales, para ser mas específico el método tradicional que mejor pronosticó para la acción CEMEX fue la regresión lineal, comparado con el sistema difuso por ACP su error es mas elevado del pronóstico tradicional.

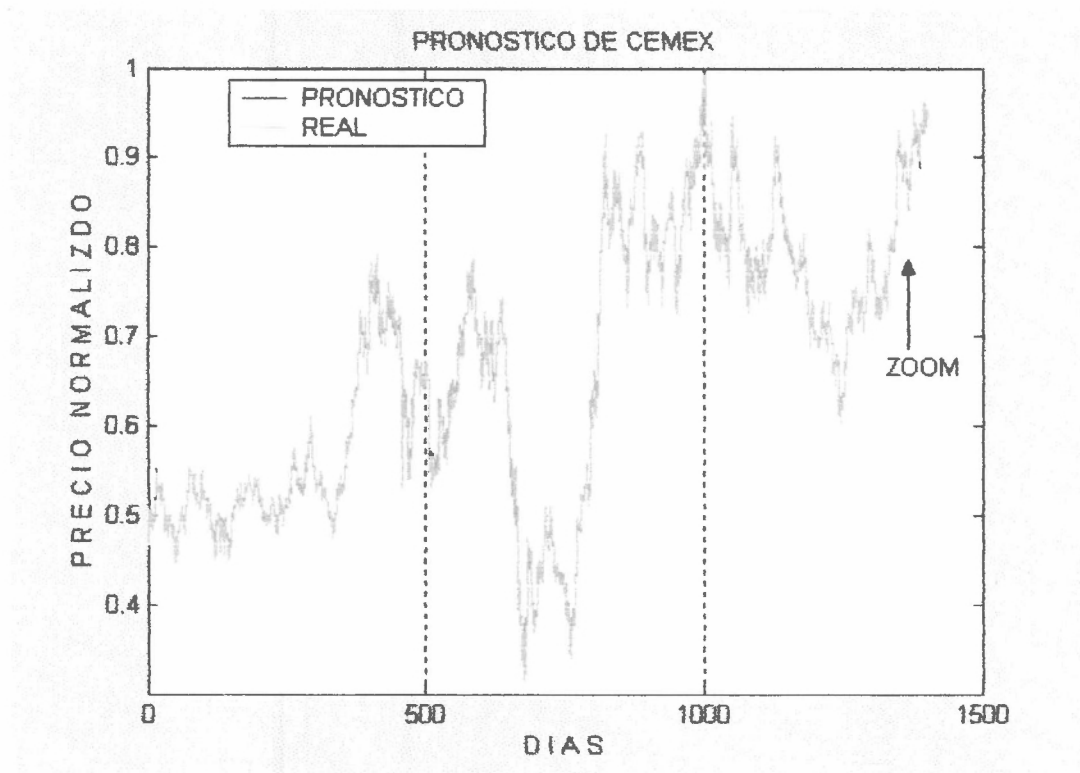


Figura 4.13 Cemex a 4 días por ACP

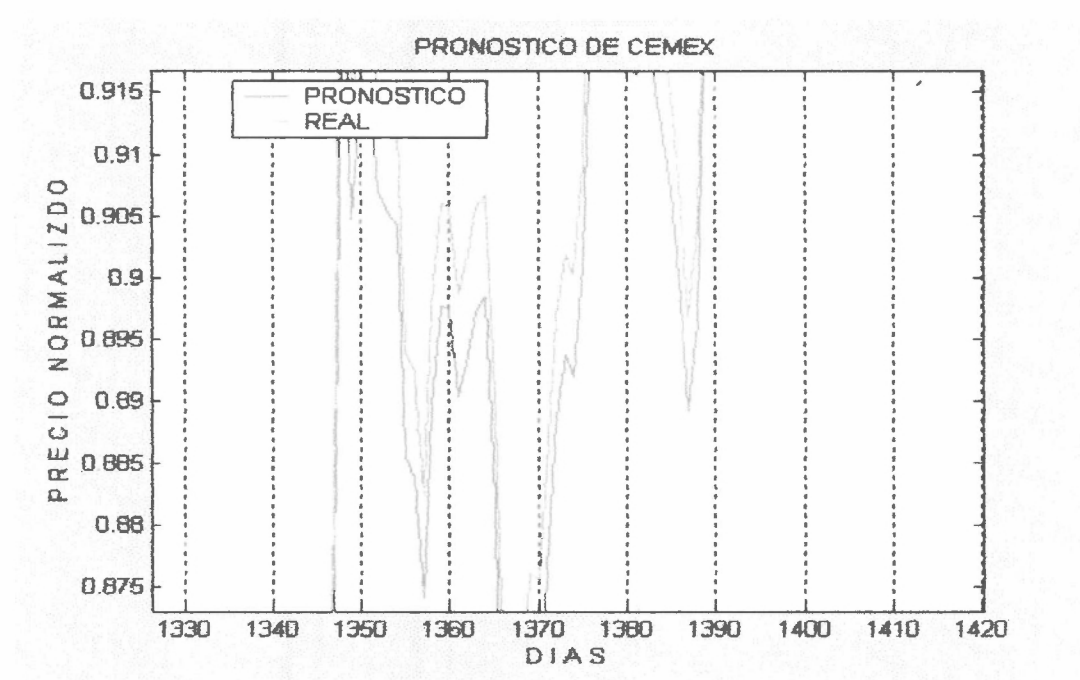


Figura 4.14 Zoom de Cemex a 4 días por ACP

Errores de CEMEX a 4 días por modelo ACP

Acción	EMC (1300)	EMC (100)	EMC TOTAL (1400)	Max Error absoluto	PME 1300	PME 100
CEMEX	8.6562e- 006	4.1038e- 006	1.3282e- 005	0.00354	-0.002	.000053

Tabla 4.14 Errores de aproximación de Cemex a 4 días por método ACP

Observamos que el modelo que mejor ajusta al pronóstico de TELMEX por el método de lógica difusa según las tablas 4.8 y 4.10 (errores) es el modelo ACP; ya que tiene un EMC en los 100 puntos a comprobar mas bajo al igual que su PME, comparado con el modelo correlación es mas bajo, y cabe mencionar que el modelo correlación en todos los ejemplos anteriores se comportó mejor que el cualquier método tradicional de pronóstico; dentro de los 1300 puntos iniciales para su aprendizaje se contempla un error bajo, pero no puede llegar a cero debido a la generalización de las reglas para todo el espectro.

Observamos que el modelo que mejor ajusta al pronóstico de CEMEX por el método de lógica difusa según las tablas 4.12 y 4.14 (errores) es el modelo ACP ; ya que tiene un EMC en los 100 puntos a comprobar mas bajo; su PME de los 100 últimos puntos está ligeramente subestimado; dentro de los 1300 puntos iniciales para su aprendizaje se contempla un error bajo, pero al igual que para TELMEX no puede llegar a cero debido a la generalización de las reglas para todo el espectro. De la misma manera que TELMEX, el modelo correlación en todos los ejemplos anteriores se comportó mejor que el cualquier método tradicional de pronóstico.

El error en estos métodos difusos en cualquiera de sus modelos puede seguir disminuyendo según lo realizado en la figura 3.12 (encontrando funciones de membresía adecuadas o adecuando manualmente las reglas)

4.3 Pronóstico Tradicional Vs Difuso

En necesario poder comprobar que el pronóstico por el método difuso pronostica con menor grado de error según nuestros resultados anteriores, y para ello se realizará el pronóstico a 12 días. Se utilizará el modelo difuso ACP ya que ha pronosticado mejor mientras que para

seleccionar el método de pronóstico tradicional el que mejor se comportó para este sistema a 12 días fue la regresión múltiple.

4.3.1 Pronóstico Telmex

La figura 4.15 es el pronóstico por el método de la lógica difusa utilizando el modelo ACP para TELMEX; se tiene: señal original (verde), la salida del pronóstico por el método de lógica difusa ACP (amarillo) y la regresión múltiple (roja) lo que se puede obtener en estas gráficas son los puntos colapsados, es decir en donde existen puntos amarillos, verdes ó rojos encimados, son en estos puntos donde se está separando el pronóstico. Podemos observar en estas gráficas que en todos los puntos máximos o mínimos responde el sistema de lógica difusa tratando de suavizar la curva, es decir no exagerando estos puntos. Así mismo se observa que existen tramos donde el pronóstico por métodos tradicionales se despega del real.

La figura 4.16 es la ampliación en la zona del pronóstico de la acción de TELMEX por el método de ACP vs tradicional vs real, este tipo de gráfica nos sirve para poder observar con mas detalle como trata de suavizar cada método con respecto a la señal original, y se puede observar que la regresión múltiple se despega un poco del real o del difuso

Pronóstico de TELMEX a 12 días.

Método Usado	Acción a pronosticar	Datos Históricos	Pronóstico	Intervalo	Variables Utilizadas	Coefficientes Regresión
Regresión Múltiple (ROJO)	TELMEX Figura 4.15	800	100	12	TELMEX KOFL VITRO HYLSAMX (Tabla 3.3)	A0= 0.0843 A1= 0.0203 A2= -0.1880 A3= -0.0351 A4= 1.0696
Difuso ACP (AMARILLO)	TELMEX Figura 4.16	1300	100	12	3 componentes Principales TELMEX (Tabla 3.6)	Campana 4,3,3,4 (M.F)

Tabla 4.15 Pronóstico Telmex a 12 días

De la tabla 4.16 se puede observar los errores generados por el sistema difuso en el modelo ACP, en todos los casos son menores por la regresión múltiple.

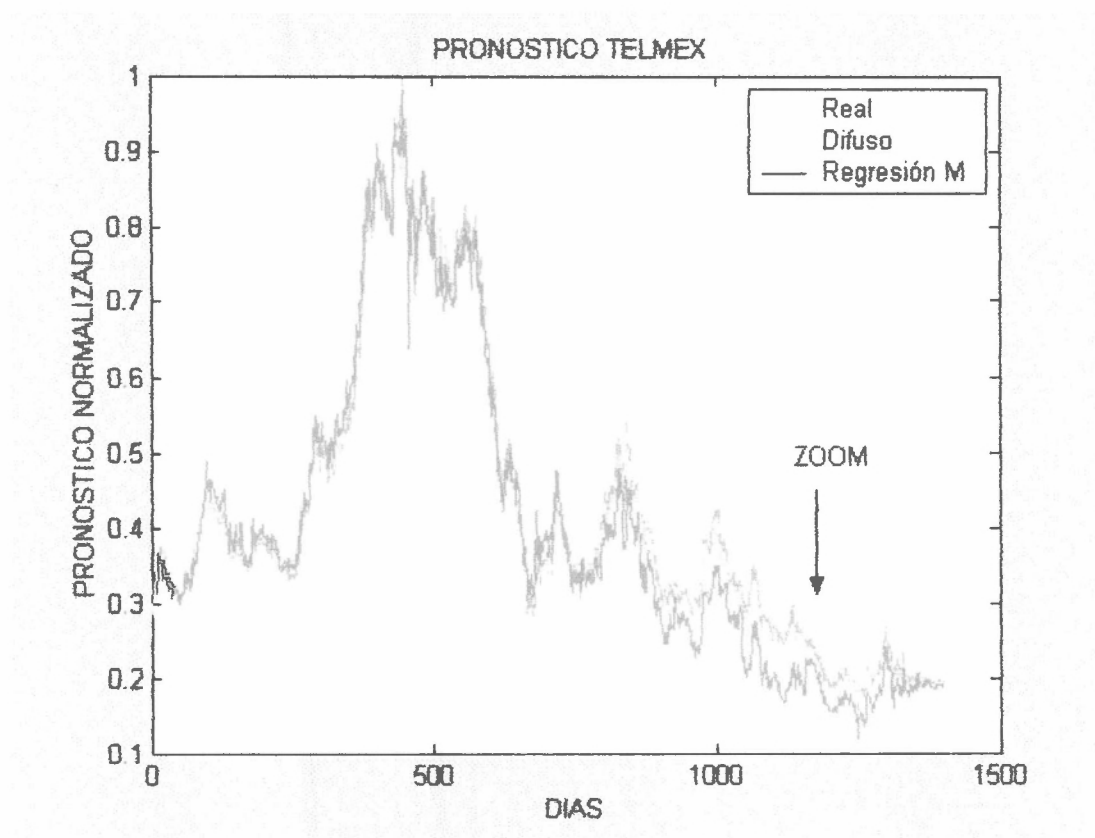


Figura 4.15 Pronóstico de Telmex a 12 días

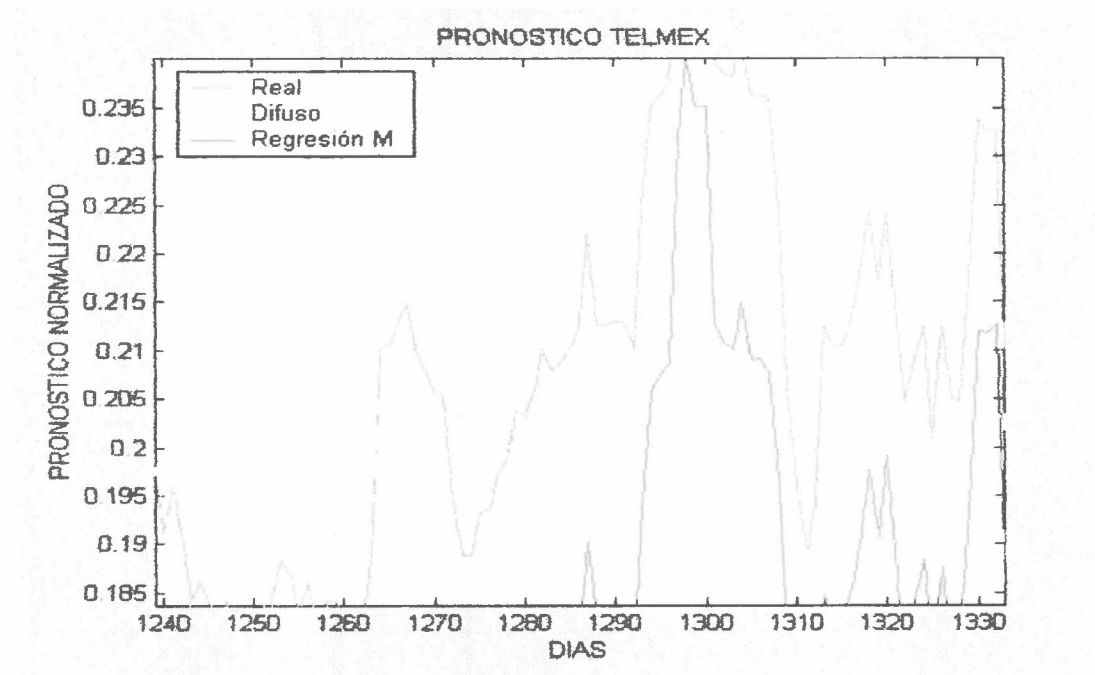


Figura 4.16 Zoom Pronóstico de Telmex a 12 días

Errores de TELMEX a 12 días

Método	EMC (1300) Aprendizaje	EMC (100) Generalización	EMC TOTAL (1400)	Max Error absoluto	PME 1300	PME 100
Regresión Múltiple	3.1210e-004	0.0013	0.0013	0.0982	0.0606	0.0712
Difuso ACP	4.025e-005	2.1365e-005	3.2546e- 004	.0023	.00058	.00784

Tabla 4.16 Errores de aproximación de Telmex a 12 días

4.3.2 Pronóstico Cemex

La figura 4.17 es el pronóstico por el método de la lógica difusa utilizando el modelo ACP para CEMEX; se tiene: señal original (verde), la salida del pronóstico por el método de lógica difusa ACP (amarillo) y la regresión múltiple (roja) lo que se puede obtener en estas gráficas son los puntos colapsados, es decir en donde existen puntos amarillos, verdes ó rojos encimados, son en estos puntos donde se está separando el pronóstico. Podemos observar en estas gráficas que en todos los puntos máximos o mínimos responde el sistema de lógica difusa ACP, tratando de suavizar la curva. Así mismo se observa que existen tramos donde el pronóstico por métodos tradicionales se despega del real.

La figura 4.18 es la ampliación en la zona del pronóstico de la acción de CEMEX por el método de ACP vs tradicional, y se puede observar que la regresión múltiple se despega un poco del real o del difuso

Pronóstico de CEMEX a 12 días

Método Usado	Acción a pronosticar	Datos Históricos	Pronóstico	Intervalo	Variables Utilizadas	Coefficientes Regresión
Regresión Múltiple (ROJO)	CEMEX Figura 4.17	800	100	12	CEMEX KOFL VITRO HYLSAMX (Tabla 3.3)	A0=0.0820 A1= 0.0185 A2= 0.0161 A3= -0.0082 A4= 0.8279
Difuso ACP (AMARILLO)	CEMEX Figura 4.18	1300	100	12	3 componentes Principales TELMEX (Tabla 3.6)	Campana 4,4,3,4 (M.F)

Tabla 4.17 Pronóstico Cemex a 12 días.

Errores de CEMEX a 12 días

Método	EMC (1300) Aprendizaje	EMC (100) Generalización	EMC TOTAL (1400)	Max Error absoluto	PME 1300	PME 100
Regresión Múltiple	0.0010	0.0032	0.0012	0.0793	0.0175	0.0646
Difuso ACP	8.9725e-005	0.0017	1.0336e-004	0.0651	-0.0021	0.0083

Tabla 4.18 Errores de aproximación de Cemex a 12 días

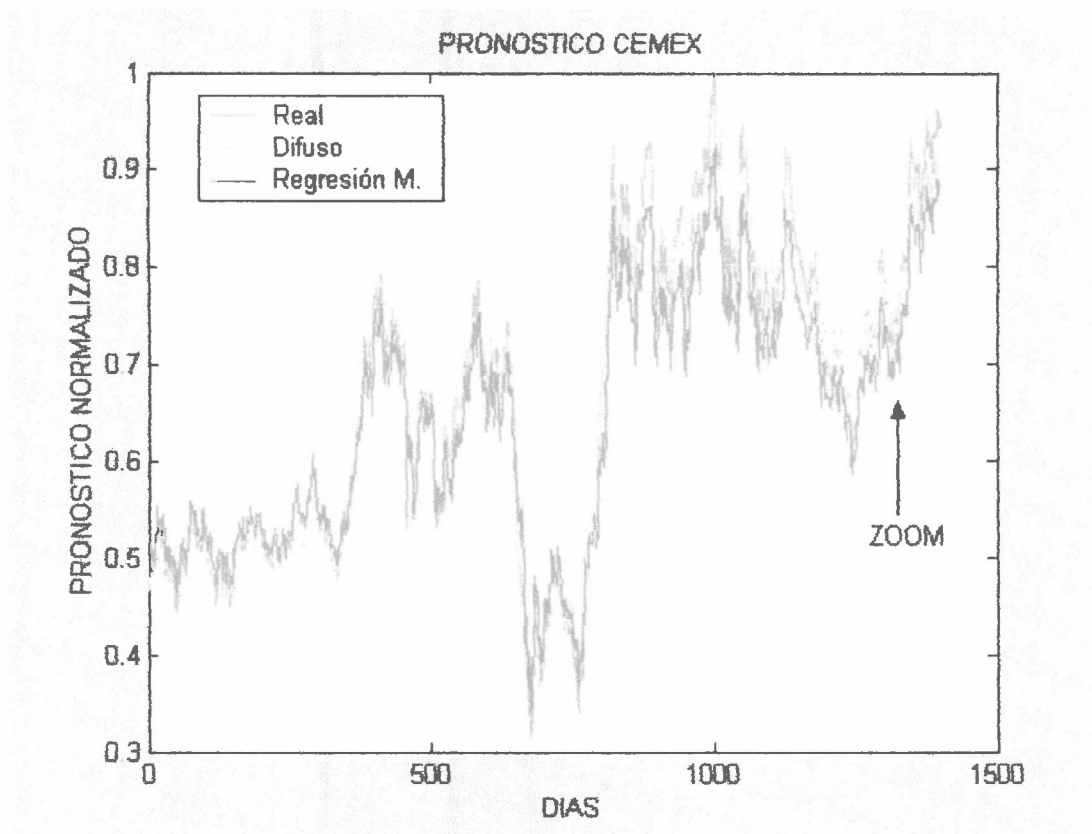


Figura 4.17 Pronóstico de Cemex a 12 días

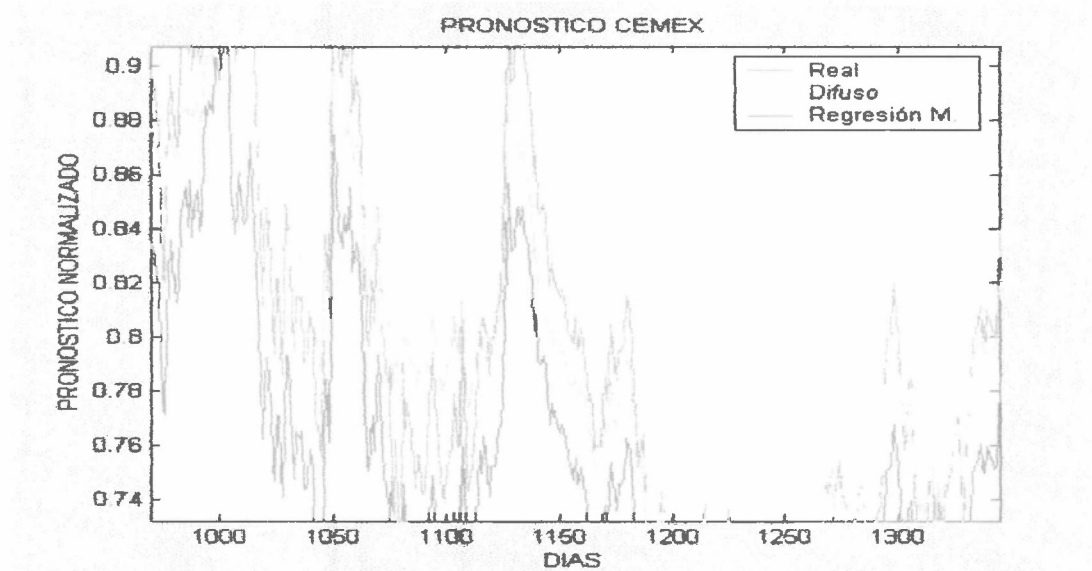


Figura 4.18 Zoom Pronóstico de Cemex a 12 días

De los resultados anteriores llegamos a concluir que tiene lo siguiente:

Para el caso de Telmex el mejor método tradicional de pronóstico fue la regresión múltiple mientras que en el pronóstico difuso a 4 días y a 12 días el mejor modelo fue el ACP. El modelo difuso ACP comparado con el tradicional, resultó ser mejor en todos sus errores, es decir siempre el sistema difuso se pegó mas al real; el EMC y el PME en la generalización (últimos 100 puntos) ha sido bastante significativa en comparación con el método tradicional.

Para el caso de Cemex el mejor método tradicional de pronóstico a 4 días fue la regresión lineal y para 12 días fue la regresión múltiple mientras que en el pronóstico difuso a 4 días y a 12 días el mejor modelo fue el ACP. El modelo difuso ACP comparado con el tradicional, resultó ser mejor en todos sus errores, es decir siempre el sistema difuso se pegó mas al real, al igual que TELMEX ; ya que tiene un EMC en los 100 puntos a comprobar mas bajo al igual que su PME; dentro de los 1300 puntos iniciales para su aprendizaje se contempla un error bajo, pero no puede llegar a cero debido a la generalización de las reglas para todo el espectro.

Par cualquier tipo de modelo difuso usado el error es muy bajo en el aprendizaje mientras que en la generalización el error aumenta.

Cabe mencionar que en el pronóstico difuso por el modelo correlación, respondió mejor que cualquier método tradicional para TELMEX y CEMEX, y un poco mas elevado que el modelo ACP. Este modelo de correlación resultó ser muy fácil de manipular su actualización, puesto que no se necesita de todo el vector de componentes principales.

El error en estos métodos puede seguir disminuyendo según lo realizado en la figura 3.12 (encontrando funciones de membresía adecuadas o adecuando manualmente las reglas)

5 CONCLUSIONES.

Tomando como base el desarrollo teórico y práctico realizado a lo largo del presente trabajo de tesis, y considerando los resultados obtenidos, se llega a los siguientes puntos, que se dan como las conclusiones más importantes de la tesis:

- Los sistemas neurodifusos emplean las cualidades propias de los sistemas difusos para incorporar el conocimiento de expertos en forma de reglas del tipo si...entonces, junto con la propiedad de las redes neuronales, que los hace ser adaptables y, de esta forma, pueden ajustar los valores de los parámetros de sus premisas o de sus consecuentes basados en un conjunto de pares de datos de entrada/salida. Estas características los hacen muy adecuados para tratar con el problema que se plantea en este trabajo, correspondiente al pronóstico.
- Se consideró a la arquitectura ANFIS como la parte fundamental de la metodología en la cual se basaría el desarrollo de este trabajo, debido, principalmente, al algoritmo de aprendizaje que emplea para el ajuste de sus parámetros adaptables. Este algoritmo es muy eficiente para minimizar la medida del error de salida del sistema cuando se ajustan los parámetros de sus funciones de membresía.
- La aplicación que se propuso en este trabajo es un problema del área financiera, pero la generalidad de los métodos construidos permite que fácilmente puedan ocuparse en otros campos de aplicación diferentes.
- Los sistemas expertos basados en lógica difusa van minimizando su error en base al aprendizaje, selección de información, reglas basadas en la experiencia.
- Se realizaron los modelos de la selección de información: ACP y Correlación, en donde mediante tecnología ANFIS se pueden hacer un pronóstico muy preciso, depende del experto el usar el método adecuado. Este tipo de diseño es novedoso y aún es tema abierto de investigación.
- Como ejemplo de aplicación, se utilizaron acciones de CEMEX y TELMEX, ya que estas acciones son bastante bursátiles y teóricamente se rigen sobre el mercado de valores

es decir sin manipulaciones, con lo que el sistema se vuelve altamente no lineal e ideal para utilizar la lógica difusa.

- La obtención y manipulación de bases de datos requiere de una veracidad. Los resultados obtenidos con los dos métodos nos dan como conclusión que en que en la parte del aprendizaje (los primeros 1300 puntos) los dos métodos (ACP y Correlación) tienen un error cuadrático medio bajo mientras que en la parte de la corrida (los 100 últimos puntos); con mas exactitud es el método ACP, esto es debido a que al usar los componentes principales realmente se extrae un porcentaje alto de la información, y es con esto que se manipula la información casi exacta.
- El método de Correlación es mas fácil utilizarlo que el método del ACP, ya que el método del ACP requiere conocer la entrada del vector de componentes principales (todos los valores del espectro) , en el método de correlación solo usamos las cabezas de variables, pero también se tiene que estar actualizando este sistema de cabezas de variables por lo menos cada 100 puntos.
- Los resultados indican que el EMC y el PME del sistema que desarrollamos es menor que los métodos tradicionales a un método tradicional usado en la plataforma de Matlab , y lo mejor de todo es que este error cuadrático medio no es estático, es decir mediante la afinación del sistema podemos disminuirlo.
- El sistema neurodifuso pueden ir disminuyendo su error de salida encontrando variables representativas, es decir variables que de alguna manera describan el comportamiento de la acción, pueden ser variables económicas, financieras, políticas, estadísticas, etc.
- El método de pronóstico moderno basado en la lógica difusa respondió con mayor precisión que los métodos tradicionales, en nuestros dos casos particulares de pronóstico de acciones: TELMEX y CEMEX.

5.1 Trabajo Futuro

Los resultados obtenidos en el presente trabajo, pueden servir como motivación para realizar trabajos posteriores que busquen elaborar un sistema ANFIS que incluya algunas características adicionales a las aquí desarrolladas. Entre ellas se pueden mencionar las siguientes:

- Buscar nuevas variables que puedan representar el espectro de información.
- Diseñar la interfase de simulación y mantenimiento en tiempo real.
- Para el método de correlación no es necesariamente que el variable que mas correlacionada este con el vector ortogonal tenga que ser grupo de cabeza, probablemente sea un trabajo de prueba y error.
- Lo más incomodo de trabajar con el método ACP es que se tiene que tener la información en línea de todas las acciones para obtener nuestro vector ACP, por lo que pudiera ser conveniente tener un programa que baje los datos de internet y así como los ordene para introducirlos al ACP.
- Exportar esta misma lógica para un sistema en tiempo real, quizá lo mas importante sea que los métodos de aprendizaje se requieren una o dos veces por semana para que el pronóstico tenga buena resolución.
- Ajustar el sistema difuso en base a estar comparando la salida, en algunas situaciones se puede ajustar con mas aprendizaje o a la mejor utilizando funciones de membresía híbridas.
- En la parte del aprendizaje podría servir utilizar un sistema de Mandani.
- Automatización del mantenimiento del sistema difuso ver figura 4.12

Apéndice A

Análisis de Componentes Principales[18].

Dado un conjunto de M vectores reales de media cero, aleatorios, $x^K \in R^n, 1 \leq K \leq M$, la matriz

de correlación R_x se determina como:

$$R_x = \frac{1}{M} \sum_{K=1}^M (x^K) (x^K)^T \quad (\text{A.1})$$

Sea $\Psi = [\psi_1, \psi_2, \dots, \psi_n]$ la matriz modal $n \times n$ de los vectores propios asociados de R_x , es decir se tiene que

$$R_x \Psi = \Psi \Lambda \quad (\text{A.2})$$

donde Λ corresponde a una matriz diagonal que contiene los valores propios de R_x ordenados en forma decreciente. Existen n proyecciones del vector de datos x.

Específicamente,

$$v_j = \psi_j^T x = x^T \psi_j, \quad j=1, \dots, n \quad (\text{A.3})$$

donde los v_j son las proyecciones de x sobre las n direcciones principales representadas por los vectores unitarios ψ_j . Los v_j se llaman componentes principales.

La data original puede ser reconstruida en forma exacta a partir de las proyecciones sobre las direcciones principales, v_j , $v = [v_1, v_2, \dots, v_m]^T = [x^T \psi_1, x^T \psi_2, \dots, x^T \psi_m]^T = \Psi^T x$

$$(\text{A.4})$$

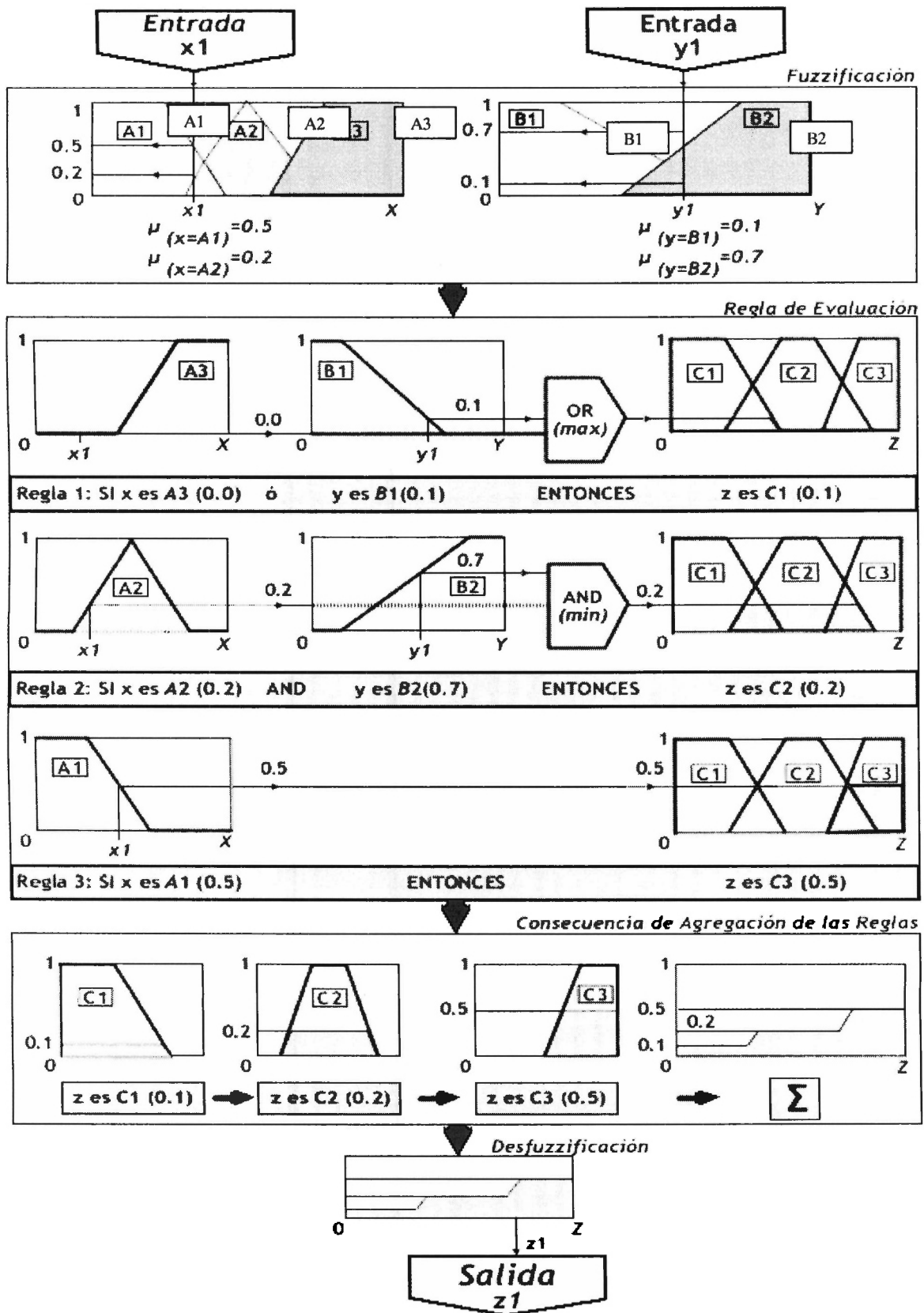
Y multiplicando a ambos lados por Ψ , se obtiene:

$$x = \Psi v = \sum_{j=1}^n v_j \psi_j \quad (\text{A.5})$$

notar que Ψ es ortonormal con respecto a sus columnas: $\Psi \times \Psi^T = I$. El vector v, de largo n, esta constituido por las proyecciones del vector de entrada x sobre los vectores propios unitarios contenidos en las columnas de Ψ . Se cumple que los vectores propios de la matriz de correlación R_x representan las direcciones principales a lo largo de las cuales las varianzas de las proyecciones del vector de entrada x tienen sus valores extremos. Los valores propios asociados definen los valores extremos de tales varianzas.

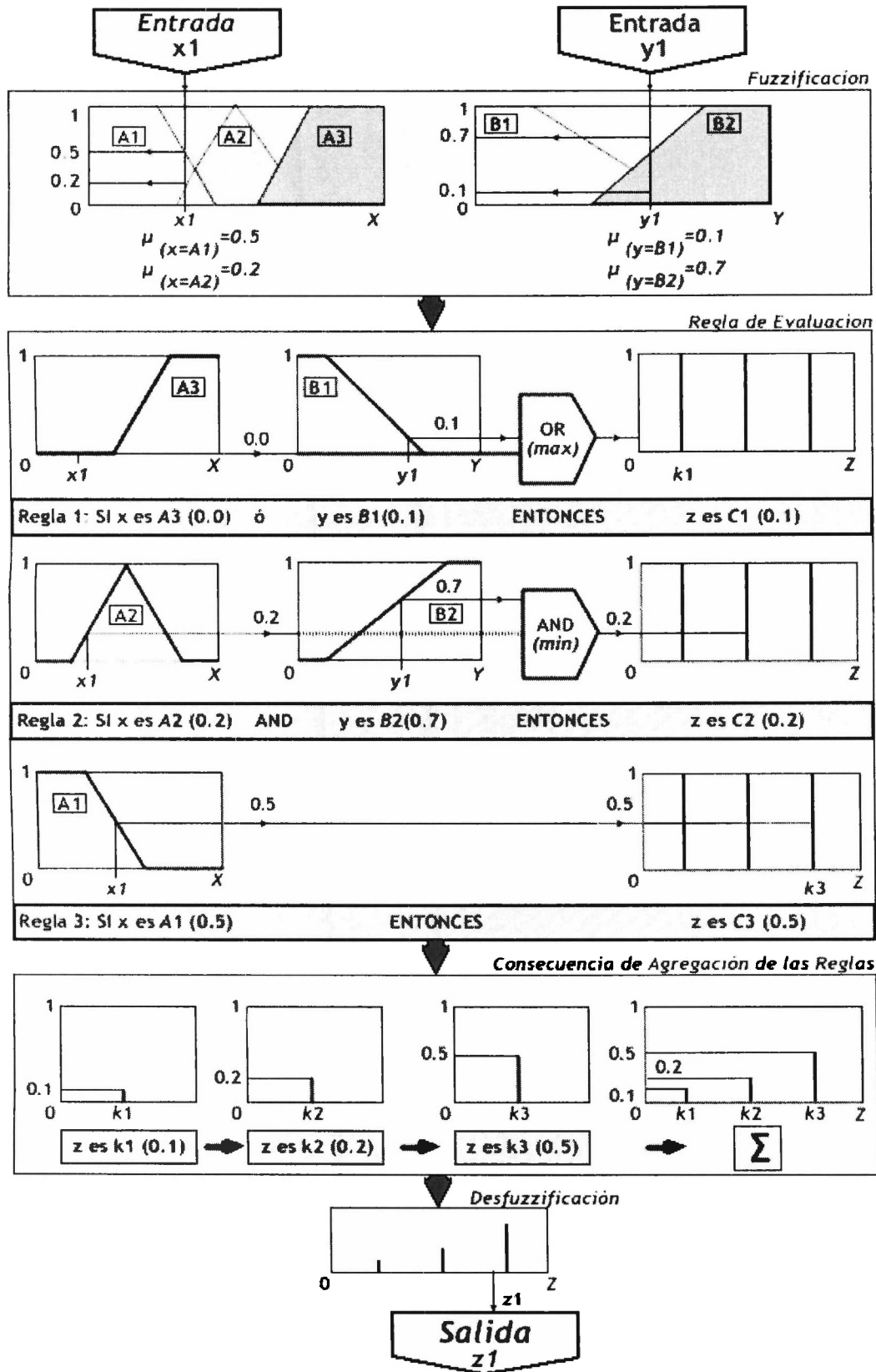
Apéndice B

Esquema General del Controlador Mandani[10]



Apéndice C

Esquema General del Controlador Sugeno[10]



Anexo 1

Programas Usados

Archivo	Descripción
Basesinnames3.txt	Base de datos 31 x 1400 (acciones)
Basededatos.xls	Base de datos 31 x 1400 (acciones) fechadas
Acp.txt	Componentes principales de nuestra base de datos 31 x 1400
ACP.matlab	Programa para generar los componentes principales, factores de correlación, vectores propios, valores medios, acumulativos
Metodo2.matlab	Programa para obtener los grupos de cabezas
ACPnew.matlab	Programa para generar los componentes principales en base a % de información
Generadorcito2.matlab	Programa para generar las matrices, en el orden respectivo y con su salida para aprender
Generadorcito10.matlab	Programa para usar método tradicionales
ANFIS	Programa en de Neuro-Fuzzy
Presentación.matlab	Simulador del sistema Difuso para términos de presentación

Variables Matriciales Generadas

Variable	Descripción
Cmp	Componentes principales
Tip=#	Componentes deseados para el % informativo
Vpr	Vectores propios
Corrfac	Correlación entre X y cmp

BIBLIOGRAFIA.

- [1]. ADAMS, FRED. *The Business Forecasting Revolution*. Prentice hall. 2002.USA. 768p
- [2]. ALLEN, SMITH. *Análisis Numérico*. Prentice Hall, (3 ed.)1997. México. 670 p
- [3]. BAALBERG. *Fuzzy Logic*, Prentice may. 2001. USA. 445 p
- [4]. BABUSKA, ROBERT. *Fuzzy Logic Control Advances in Applications Wold Scientific*. 1999. Estados Unidos. 579 p.
- [5]. COUTO, EDUARDO. *Invierta con éxito en la Bolsa de Valores*. Prentice Hall. 2000. México. 280 p.
- [6]. DRAPER,N.R. *Applied Regression Analysis*. Prentice Hall. 1978. USA. 256 p.
- [7]. DROSSMAN, BANKANHEN. *Advanced Forecast*. Prentice Hall. 2001. Estados Unidos. 380 p.
- [8]. HANKE, JOHN. *Pronósticos en los Negocios*. Prentice Hall.1999. México. 599 p
- [9]. JANG, ROGER. *Fuzzy Logic Toolbox*. Math Works. 2001. Estados Unidos. 459 p.
- [10]. KASABOV, NIKOLA *Neuro-Fuzzy Techniques for Intelligent Information System*. .Physica-Verlag. 1999. Alemania 460 p.
- [11]. KIUPEL, HEN. *ANFIS evaluation concept*, Prep. IEEE In Conf on Systems. 767-776. Vancouver (1999)
- [12]. KNIGHT, RICH. *Artificial Intelligence*. Mc Graw Hill (2 ed). 1995. Estados Unidos. 755 p.
- [13]. LAVENBACH, CLEARY. *The Profesional Forecaster*. Lifetime Learning Publications. 1982. USA. 179 p
- [14]. LUTHE, SCHUTZ. *Métodos numéricos*. Limusa. 1986. México.860 p.
- [15]. NAUCK,D. AND F. KLOWONN. *Neuro-Fuzzy Classification Initialized by Fuzzy Clustering*, Japan/USA Symposium on Flexible Automatization. Vol 2. 1994
- [16]. NEWTON, JOSEPH. *Time Series Análisis*. Prentice Hall. 2001. USA. 476 p.
- [17]. OGATA, KATSUHIKO *Ingenieria de Control Moderna*. Prentice Hall. (1985).México. 590p
- [18]. WILLOUGHBY, STEPHEN. *Probabilidad y Estadística*. Publicaciones cultural. (12 ed)1988. España. 850 p.
- [19]. ROJAS, I., J.L. BERNIER, M. RODRIGUEZ-ALVAREZ, A. PRIETO (University of Granada), *What are the Main Functional Block Involved in the Design of Adaptive Neuro-Fuzzy Inference Systems?*, Proceedings of the IEEE-INNS-ENNS, International Joint Conference on Neural Networks (IJCNN'00). Oct. 30 2000.
- [20]. WINTERS, P.R. *Forecasting Sales by Exponentially Weighted Moving Average*. Prentice Hall. USA. (1985). 390 p.