

ESTÚDIO COMPARATIVO DE MODELOS DE PRONÓSTICO DE VENTAS

Guillermo Corres¹

Lucia Isabel Passoni²

Claudia Zárate³

Alejandra Esteban⁴

RESUMEN: La mayoría de las decisiones que deben tomarse en las operaciones, se hacen a partir de los pronósticos y el error o incertidumbre de la demanda asociados a los mismos. El propósito de este trabajo fue comparar el comportamiento de métodos de pronóstico de distintos orígenes sobre las ventas históricas de productos individuales con marcado componente estacional. El análisis del comportamiento se realizó a través de los indicadores MAPE y de su robustez, considerando su desviación estándar. Se calcularon pronósticos sobre 5 productos de la misma familia utilizando métodos de diversos orígenes: determinísticos (suavizado exponencial en sus versiones simples y complejas), ARIMA (Método Autorregresivo Integrado de Medias Móviles) y ANFIS (Sistema Adaptativo de Inferencia Neuro Difuso). A fin de disminuir el valor de dichos errores se introdujeron en los modelos de ARIMA y ANFIS, series de variables independientes que podrían afectar el comportamiento de la serie analizada. Los resultados obtenidos demuestran que cuando los datos fueron analizados en forma trimestral arrojan menores valores de error y este resultado se asocia a la menor variabilidad de la serie. Además, es posible utilizar métodos sencillos en su aplicación e interpretación. No obstante, es importante destacar que si bien ANFIS y ARIMA no siempre resultaron los mejores modelos, ambos métodos poseen la ventaja de poder mejorarse, ya sea modificando las variables independientes que se incorporen al modelo o los parámetros del modelo aunque requiera mayor tiempo computacional.

Palabras clave: Pronósticos. Suavizado exponencial. ARIMA. ANFIS.

¹ Ingeniero Industrial, Universidad Nacional Del Centro de La Provincia de Buenos Aires, Facultad de Ciencias Económicas, Argentina. E-mail: corres@rec.unicen.edu.ar.

² Dra. Ingeniería Electrónica, Universidad Nacional de Mar del Plata, Facultad de Ingeniería, Argentina. E-mail: lpassoni@fi.mdp.edu.ar.

³ Ingeniero Mecánico, Universidad Nacional de Mar del Plata, Facultad de Ingeniería, Argentina. E-mail: cnzarate@fi.mdp.edu.ar.

⁴ Ingeniero Química, Universidad Nacional de Mar del Plata, Facultad de Ingeniería, Argentina. E-mail: aesteban@fi.mdp.edu.ar.

1 INTRODUCCIÓN

Todas las organizaciones se desempeñan en una atmósfera de incertidumbre. Los métodos de pronósticos se utilizan para predecir dichas operaciones de naturaleza incierta. Pronosticar consiste en la estimación y el análisis de la demanda futura para un producto en particular, un componente o un servicio, a través de diferentes técnicas de previsión. El pronóstico de la demanda futura es central en cualquier actividad de planificación y de operaciones, en particular en actividades relacionadas a la logística y a la cadena de suministro.

A nivel organizacional, el pronóstico de las ventas se considera una entrada esencial para el resto de las funciones que deben desarrollarse, tales como marketing, producción, compras, finanzas, contaduría, etc.

Es de fundamental importancia considerar el error que inevitablemente viene asociado a toda predicción. Numerosos autores como Zhang (2003) han reconocido la importancia del mismo en el momento de la toma de decisiones. Considerar el pronóstico sin tener en cuenta el error del mismo puede producir impactos no deseados en todas las áreas de la empresa, desde los aprovisionamientos, pasando por la producción y en la comercialización de los productos.

Las ventas minoristas de productos finales poseen en general un comportamiento marcadamente estacional. Este hecho ha generado importantes desarrollos desde la investigación con objeto de obtener modelos que puedan predecir el comportamiento de series históricas con estas características.

Los modelos cuantitativos que tradicionalmente se han utilizado incluyen método descomposición en series de tiempo, suavizado exponencial, como así también los métodos de regresión, autorregresivos y promedio móvil integrado (ARIMA). Todos estos métodos mencionados poseen un fundamento estadístico formal y de todos ellos, quizás el que mejores resultados ha tenido en términos de mejor desempeño en el modelado ha sido ARIMA.

Una de las principales limitaciones de estos métodos es que son esencialmente lineales. Estos modelos lineales tienen una gran ventaja de ser fáciles de implementar y de explicar. No obstante, pueden ser totalmente inadecuados cuando el mecanismo que genera la serie es no lineal. No parece ser muy razonable asumir a priori que una serie de tiempo, en particular, es generada por un proceso lineal, cuando en realidad, en el mundo real, los sistemas son frecuentemente no lineales (ZHANG; PATUWO; HU, 1997).

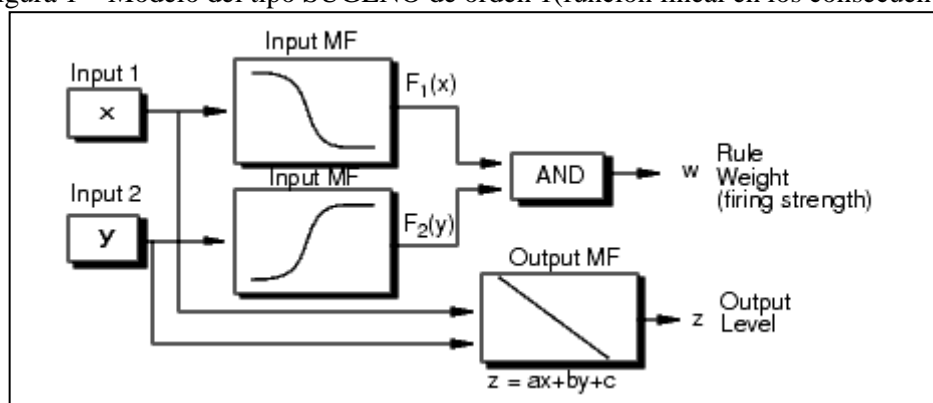
Diversas técnicas del área del “*Softcomputing*” tales como Redes Neuronales Artificiales (RNA), Sistemas de Inferencia Difusa (FIS), y algoritmos evolutivos, han demostrado tener un amplio campo de aplicación en el área de los pronósticos.

Las Redes Neuronales Artificiales, (RNA) son técnicas inspiradas en el funcionamiento de sistemas biológicos y se han aplicado en numerosos trabajos demostrado ser una herramienta muy atractiva en el área de pronósticos. Aunque los resultados de muchos autores como Chu et al. (2003), Thomassey et al. (2004), entre otros, pareciera subordinados a determinados contextos, su capacidad de modelar relaciones no lineales, su entrenamiento y facilidad son potencialmente atractivos para aplicaciones en el área de pronósticos.

Los Sistemas de Inferencia Difusa (FIS) también han sido aplicados a sistemas de pronósticos. Un sistema de Inferencia Difusa, *Fuzzy Inference System* (FIS) se compone de una base de reglas (base de conocimiento), del tipo si-entonces, donde el antecedente está formado por un conjunto de variables lingüísticas (conjuntos difusos) cuyo valor se obtiene utilizando operadores de lógica difusa. Se pueden implementar dos tipos de modelos distintos: en el caso del modelo tipo Mamdani el consecuente de cada regla es un conjunto difuso, mientras que en el tipo Sugeno (Figura 1), el consecuente es una función de las variables del antecedente, esta función puede ser de tipo 0, 1, 2, de acuerdo al orden de la misma.

Para la generación de un modelo de tipo Sugeno a partir de datos, las reglas del FIS se pueden obtener, de varias formas: una ellas es particionando el espacio de los datos en grillas equiespaciadas o bien utilizando algoritmos de agrupamiento difusos (*subtractive clustering, fuzzy c-means*).

Figura 1 – Modelo del tipo SUGENO de orden 1(función lineal en los consecuentes)



Fuente: Elaboración propia

El ANFIS (*Adaptive Neuro-Fuzzy Inference System*) adapta parámetros del FIS (de las funciones de pertenencia difusas, de las funciones de salida) usando tanto algoritmos de

retropropagación del error, solos o combinados con métodos de mínimos cuadrados. Este ajuste permite que los sistemas FIS aprendan de los datos que están modelando, generando automáticamente la base de reglas que mejor se adapta al conjunto de datos de entrenamiento.

2 MARCO TEÓRICO

En este apartado se fundamentará la aplicación de distintos modelos aplicados al cálculo de pronóstico. Se describen también las distintas medidas que se utilizan a efectos de evaluar el comportamiento de los modelos.

2.1 Técnicas de análisis de series de tiempo

Son numerosas las técnicas desarrolladas utilizando los datos históricos. Dentro de éstas, tal vez la metodología más útil para el pronóstico a corto plazo sea el ajuste exponencial (BALLOU, 2004). En esta técnica el pronóstico de demanda para el siguiente período estará dado por (Ecuación 1):

Pronóstico nuevo = α * (demanda real) + (1- α) (pronóstico anterior)

$$F_{(i+1)} = \alpha * A_i + (1 - \alpha) * F_i \quad (1)$$

Siendo

$F_{(i+1)}$ = Pronóstico para el período “ $i+1$ ”

α un factor de ponderación o constante de ajuste exponencial

A_i = Las ventas reales en período “ i ”

F_i = el pronóstico realizado para el período “ i ”

Derivado de este método, se hallan los modelos de ajuste exponencial corregidos por tendencia (*Holt*), por tendencia y estacionalidad (*Holt – Winters'*).

- Otro método que suele utilizarse por su facilidad de comprensión y aplicación es promedio móvil simple y ponderado.
- Modelo ARIMA, es quizás el método más completo para pronosticar series de tiempos, ya sea que se trate de series estacionales o no. Su gran amplitud en términos de aplicaciones es debido al trabajo de Box y Jenkins quienes desarrollaron un método sistemático y práctico para construir el modelo. En publicaciones previas (CORRES et al., 2008) se ha hecho una revisión rigurosa acerca de esta metodología.

2.1 Errores en los Pronósticos: elección de un modelo adecuado

El error o residuo se define como la diferencia entre el valor de lo que realmente ha ocurrido y el valor de la proyección (Ecuación 2).

$$Error = e_t = A_t - F_t \quad (2)$$

Siendo:

A_t = ventas reales en el período “t”

F_t = pronóstico para el período “t”

La literatura provee diversas formas para medir el error del pronóstico, como una herramienta utilizada al momento de la elección de un método. Algunos de los métodos más empleados son: Desviación Media Absoluta (MAD), Porcentaje del Error Medio Absoluto (MAPE), Error Cuadrático Medio (RMSE), Suma de los Errores BIAS, entre otros.

Según Chopra y Meindl (2008), el error de pronóstico contiene información valiosa que debe analizarse por dos razones: por un lado, permite detectar si el método predice con precisión o ya es tiempo de cambiar el mismo; por otro lado todos los planes de contingencia deben considerar el error de pronóstico.

Cuando se debe seleccionar un procedimiento de pronóstico el objetivo final será el de mantener los costos lo más bajo posible. Ellos incluyen el de obtención de pronósticos como así también el del error.

Si se analiza el costo de obtención del pronóstico, no sólo ha de considerarse el precio de hardware, software y personal especializado, sino también el tiempo que insume ya sea de procesamiento como así también de recolección de datos, monitoreo del proceso e interpretación de los resultados. Dichos costos deberán balancearse con los beneficios recibidos.

En cambio, evaluar el impacto sobre los costos del error de pronóstico es una tarea difícil. Muchos autores han realizado esfuerzos en esa dirección, Xie et al. (2004), Sanders y Graman (2006), Kerkkänen et al. (2009), entre otros. Ellos demuestran que el impacto del error no es constante y depende mucho del tipo de organización de la cual se trate.

En el área de operaciones, el error de pronóstico puede impactar ya sea en la planificación de la producción, en la planificación de la capacidad o en los almacenes. En estos últimos, es necesario considerar que a efectos de la optimización de los recursos, no

significa lo mismo el error en exceso o sobrepronosticar que origina mayores costos de posesión-, que pronosticar en defecto o subpronosticar, que generará rupturas de inventario.

De cualquier forma, la consideración primordial en la elección de un método de pronóstico es que los resultados deben facilitar el proceso de la toma de decisiones de los administradores de la organización. Hanke y Wichern (2006) explican, que rara vez, un solo método funciona para todos los casos. No obstante es más eficiente aquel pronosticador capaz de formular una mezcla habilidosa de técnicas cuantitativas de pronósticos con un buen juicio personal y que evita los extremos donde se depende completamente de cualquiera de las dos.

Existen distintas formas de medir el error. Entre las más usadas se encuentran:

- La desviación media absoluta MAD. Es una forma sencilla y útil de medir el error. Es el promedio de todos los errores de la proyección, independientemente de que los errores sean positivos o negativos. La MAD es más útil cuando el analista quiere medir el error de pronóstico en las mismas unidades que la serie original, Hancke et al. (2006).

Se calcula como (Ecuación 3):

$$MAD = \frac{\sum_{t=1}^{t=n} |A_t - F_t|}{n} \quad (3)$$

Siendo “n” la cantidad de períodos considerados

- La Raíz de la Media del Error Cuadrático Medio RMSE es otra forma de medir el error en el pronóstico y se obtiene de promediar los residuos al cuadrado (Ecuación 4).

$$RMSE = \sqrt{\frac{\sum_{t=1}^{t=n} (A_t - F_t)^2}{n}} \quad (4)$$

La RMSE puede ser un mejor indicador que la MAD (FOGARTY; BLACKSTONE; HOFFMANN, 1999; BERENSON; LEVINE; KREHBIEL, 2001). Hay quienes consideran que los errores muy grandes en el pronóstico de ventas tienen un impacto más negativo que el efecto positivo generado por un error pequeño. El RMSE asigna más ponderación a los

errores grandes y es por esto que se la considera al momento de seleccionar el mejor método de pronóstico.

- A veces es más útil calcular los errores de pronóstico en porcentajes. El error porcentual absoluto MAPE, proporciona una indicación de cuán grandes son los errores de pronóstico en comparación con los valores reales de la serie (Ecuación 5):

$$MAPE = \frac{1}{n} * \sum_{t=1}^{t=n} \frac{|A_t - F_t|}{A_t} \quad (5)$$

MAPE normaliza los cálculos de error, y permite comparar los errores en el pronóstico para distintas series de tiempo. Considerando la forma de calcular este estimador, se convierte en una medición muy significativa para aplicar cuando se intenta determinar la influencia del error del pronóstico en las áreas operativas de la organización (LUXHOJ; RIIS, STENSBALLE, 1996).

El objetivo de este trabajo es aplicar distintos métodos para obtener los pronósticos de ventas de cinco productos pertenecientes a una misma familia, que poseen comportamiento marcadamente estacional. Los resultados se evaluarán a través de la MAPE, de la robustez del indicador y por las características técnicas y de aplicación de los distintos métodos.

3 MATERIALES Y MÉTODOS

Para el desarrollo de este trabajo se parte de los datos reales de ventas durante 10 años de productos pertenecientes a una misma empresa de conservas de la ciudad de Mar del Plata, con características de estacionalidad e insumos similares entre sí.

Se analizan los datos de cinco productos: Atún en Aceite (AA), Atún al Natural (AN), Caballa en Aceite (CA), Caballa al Natural (CN) y Sardina en Aceite (SA). Los datos se consideran en forma mensual y trimestral.

Los modelos se obtienen utilizando los datos de los primeros 9 nueve años. Dichos modelos se aplican luego para el último año y se comparan los resultados con las ventas reales correspondientes a ese año. A partir de esta diferencia, evalúa el error del modelo.

Los métodos aplicados son provenientes de distintos orígenes, a saber:

- Métodos determinísticos (Promedio móvil, suavizado exponencial, *Holt – Winters*’ y sus variantes)
- Método estocástico (Método ARIMA)
- Método derivado de Inteligencia Computacional ANFIS

Una potencialidad interesante que presentan ARIMA y ANFIS es la posibilidad de incorporar al conjunto de datos que alimentan al sistema, un conjunto de variables, denominadas “variables exógenas” que podrían contribuir a mejorar la salida.

Si se consideran los productos cuyas ventas se están pronosticando, los mismos poseen características similares dado que, como se mencionó corresponden a una misma familia. En consecuencia, para mejorar la salida del pronóstico de un producto, se alimentó al sistema de matrices de datos, obtenidas de las ventas de los otros productos.

A los resultados generados por estos modelos modificados se los identifica como ARIMA EXO y ANFIS EXO.

En la secciones 3.1 a 3.3 se presentan, a modo de ejemplo, algunos de los resultados de la aplicación de los métodos a las series de ventas reales.

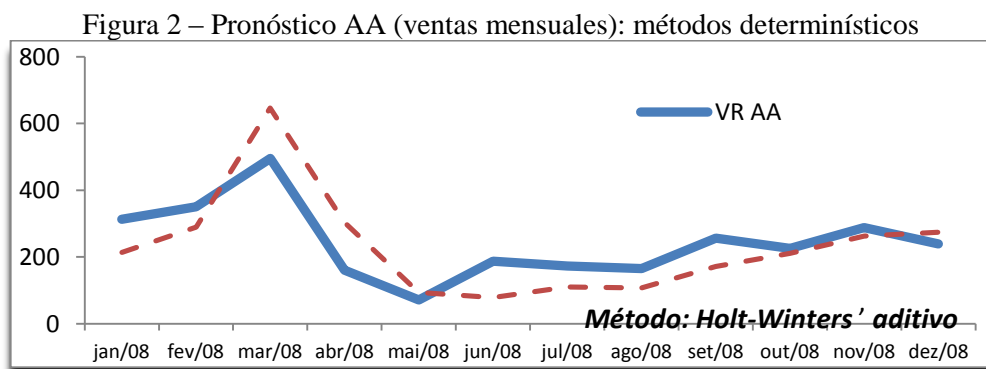
3.1 Obtención del pronóstico aplicando métodos determinísticos

Se utiliza el *software Crystal Ball*, del cual se obtienen los pronósticos solicitados, indicándole al sistema que haga la selección del mejor método analizando el valor de MAPE.

Los métodos que aplica este software son:

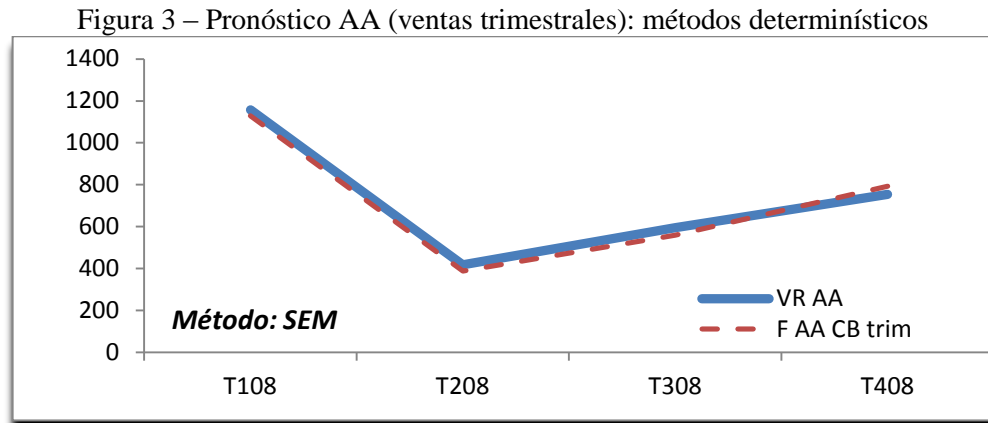
- Suavizado exponencial simple y doble (SES, SED)
- Promedio móvil simple y doble (SMA, DMA)
- Suavizado exponencial corregido por tendencia, aditivo y multiplicativo (STA STM)
- Suavizado exponencial corregido por estacionalidad, aditivo y multiplicativo (SEA SEM)
- Holt Winters’ aditivo y multiplicativo. (HWA, HWM)

En las Figuras 2 y 3 se presentan los valores de pronóstico para el caso de AA en forma mensual y trimestral.



Fuente: Elaboración propia

Para este conjunto de datos mensuales se advierte que el mejor método – *Holt Winters'* aditivo –, sobrepronostica en la época de pico máximo de demanda (Marzo/Abril) mientras que en el resto de los meses subpronostica.



Fuente: Elaboración propia

En cambio, en esta serie de datos de ventas trimestrales, se puede observar que para AA, el mejor método fue el de Suavizado Exponencial corregido por Estacionalidad Multiplicativo. En este caso es notorio el buen ajuste del pronóstico a las ventas reales.

3.2 Obtención del pronóstico utilizando ARIMA

Para aplicación de la metodología Box Jenkins se utilizó el software SPSS® (Statistical Package for the Social Sciences).

Los modelos Arima se expresan en función de (p, d, q) regular y $(P, D, Q)_4$ estacional. Estos parámetros son obtenidos analizando la serie a través de sus correlogramas (gráficas de autocorrelación), y corroborados por un test de significación de *t* de *Student*.

En Tabla 1 se presentan los coeficientes del modelo, para la obtención de pronósticos del último año para AA.

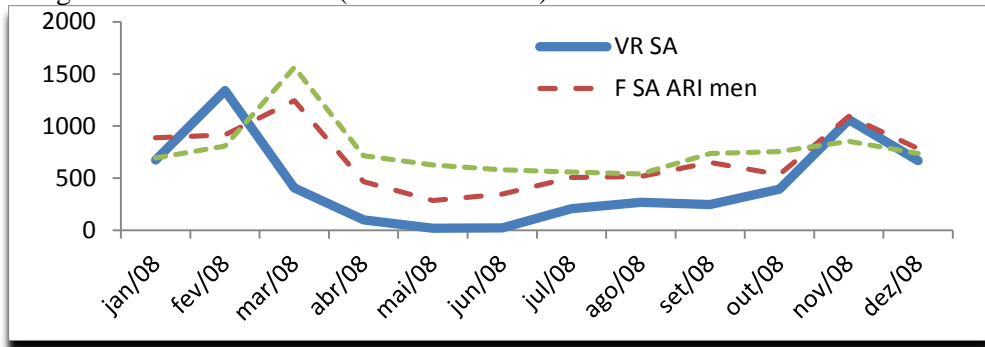
Tabla 1 – Coeficientes modelo ARIMA, producto SA

Coeficientes del Modelo		
ARIMA mensual	$(p,d,q) (1,1,1)$	$(P,D,Q)(1,0,0)_{12}$
ARIMA trimestral	$(p,d,q) (0,1,0)$	$(P,D,Q)(1,0,1)_4$

Fuente: Elaboración propia

Figuras 4 y 5 presentan los pronósticos resultantes de aplicar ARIMA y ARIMA–EXO a la serie de SA, datos mensuales y trimestrales respectivamente.

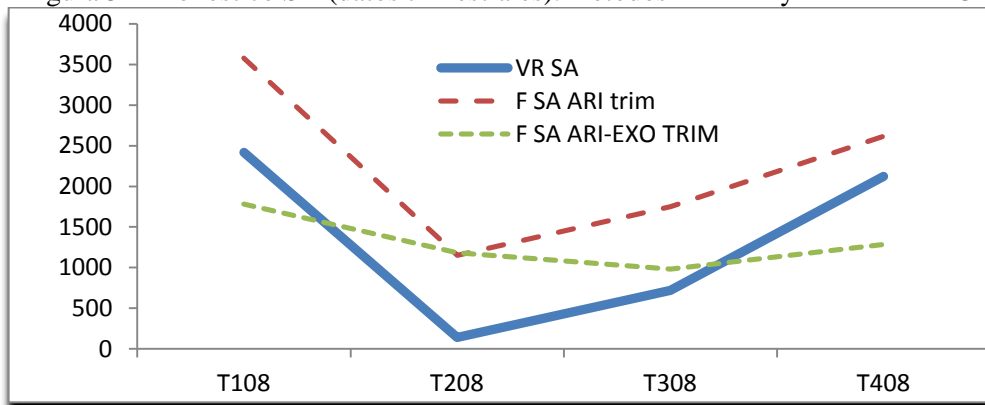
Figura 4 – Pronóstico SA (datos mensuales): método ARIMA Y ARIMA-EXO



Fuente: Elaboración propia

Es interesante observar que para los datos mensuales de SA, ambos métodos proyectan el primer pico de demanda atrasado en un período, mientras que el último pico se predice mejor. El retraso en la predicción del primer pico se podría atribuir al aprendizaje de los datos históricos que presentaron en la mayoría de los casos, dicho pico en el mes de marzo. La serie del último año lo presenta, atípicamente, en el mes de febrero.

Figura 5 – Pronóstico SA (datos trimestrales): métodos ARIMA y ARIMA-EXO



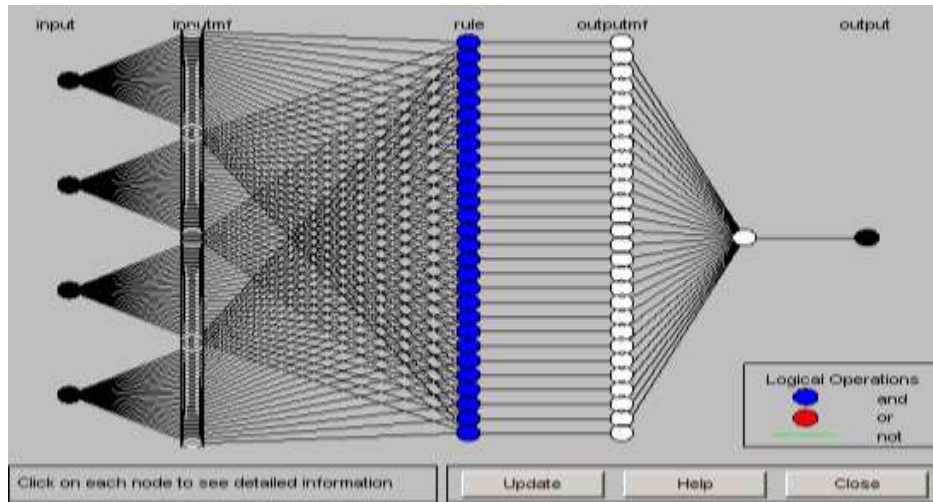
Fuente: Elaboración propia

Para los datos trimestrales del mismo producto, en cambio, se observa que la metodología ARIMA sigue la tendencia de las ventas reales de manera más cercana que en el caso de los datos mensuales (Figura 4). No obstante, no se observa en general para el caso de SA un buen ajuste de ambos métodos a las ventas reales. ARIMA sobrepronostica en casi todos los períodos en el caso de los datos mensuales y en todos los períodos trimestrales. La incorporación de las variables exógenas al método podría significar una mejora al comportamiento para los datos en forma trimestral pero no en el caso mensual.

3.3 Obtención del pronóstico aplicando ANFIS

Para la obtención de pronósticos aplicando esta herramienta se utilizó el software MATLAB®, *Fuzzy Toolbox*. La estructura del modelo ANFIS obtenida para los datos trimestrales se muestra en Figura 6.

Figura 6 – Estructura ANFIS, AA trimestral

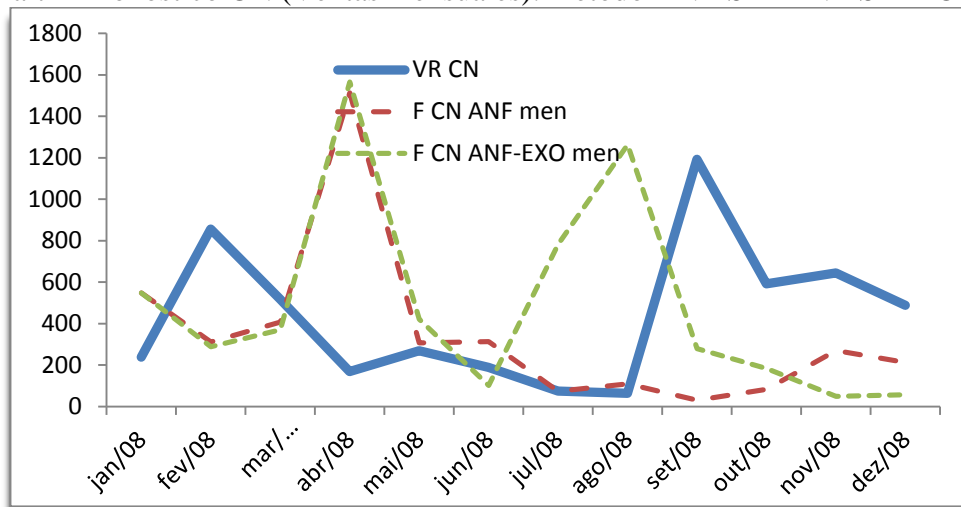


Fuente: Elaboración propia

En la misma se observa el vector de 4 entradas al sistema, la conversión de las mismas a través de 28 funciones de membresía y una salida. Los parámetros que se utilizaron en el modelo variaron según se trató de datos mensuales o trimestrales. En el primer caso se utilizó un *radii* (radio de influencia)=0.25 - 0.3 en el caso de datos trimestrales un *radii*=0.25.

En Figuras 7 y 8 se muestra la aplicación de esta metodología a la predicción de ventas de CN, datos mensuales y trimestrales respectivamente.

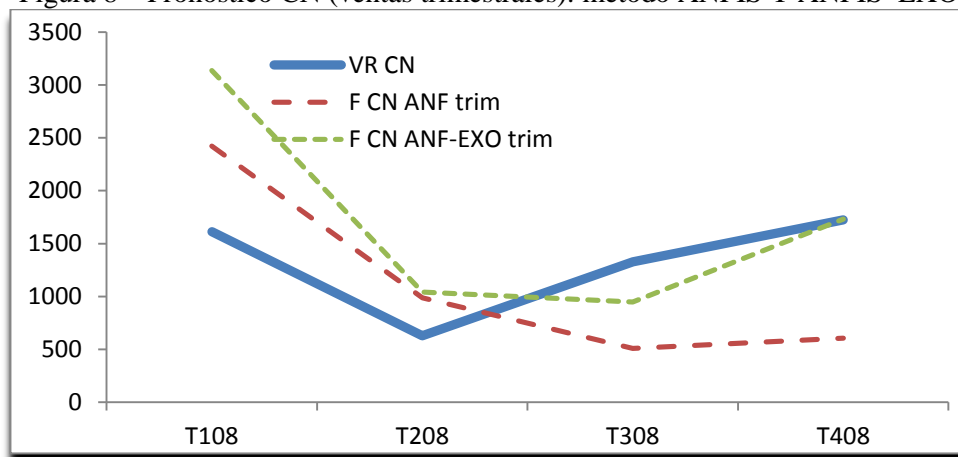
Figura 7 – Pronóstico CN (Ventas mensuales): método ANFIS Y ANFIS-EXO



Fuente: Elaboración propia

La aplicación del modelo ANFIS y ANFIS EXO sobre estos datos de ventas mensuales de CN muestra la incapacidad del mismo, en este horizonte de predicción, para anticipar el comportamiento del mercado en el año que en particular se está analizando.

Figura 8 – Pronóstico CN (ventas trimestrales): método ANFIS Y ANFIS-EXO



Fuente: Elaboración propia

Cuando se analizaron las ventas trimestrales también se observó que el modelo no es bueno, aunque se puede observar que sigue la tendencia. De acuerdo a lo mostrado en ambas figuras, el método ANFIS no tiene un buen comportamiento en el modelado de esta serie y no mejora cuando se incorporan las variables exógenas.

4 RESULTADOS Y DISCUSIÓN

En esta sección, se presentan los resultados de aplicar los 3 modelos con su variante de utilizar variables exógenas, a las ventas de los 5 productos. Dichos resultados se cuantifican a partir del cálculo de MAPE.

4.1 Análisis integral de las series mensuales y trimestrales

En Tablas 2 y 3 se presentan las MAPE promedio anual trimestral y mensual, arrojada por los distintos métodos sobre las series de cada uno de los productos.

Tabla 2 – MAPE Promedio Anual (datos trimestrales)

	AA	AN	CA	CN	SA	Prom. Trim.
CB	0.0449	0.1446	0.4061	0.4570	0.4686	0.3042
ARI	0.5691	0.5117	0.5308	0.5370	0.6624	0.5622
ARI-EXO	0.7146	0.6769	0.6535	0.6425	0.7147	0.6804
ANFIS	0.4351	0.4957	0.5761	0.6895	0.9727	0.6338
ANFIS-EXO	0.7849	0.7855	0.8207	0.8212	0.8859	0.8196
MAPE PROM trim	0.5240	0.5229	0.5974	0.6294	0.7409	0.6001

Fuente: Elaboración propia

Los valores presentados indicarían que, en general, utilizando datos trimestrales los Métodos Determinísticos se obtiene el mejor ajuste, es decir que su aplicación deriva en obtener menor MAPE. Con el mismo criterio se observa el mejor comportamiento de los productos AA y AN respecto del resto de los productos.

Tabla 3 – MAPE Promedio Anual (datos mensuales)

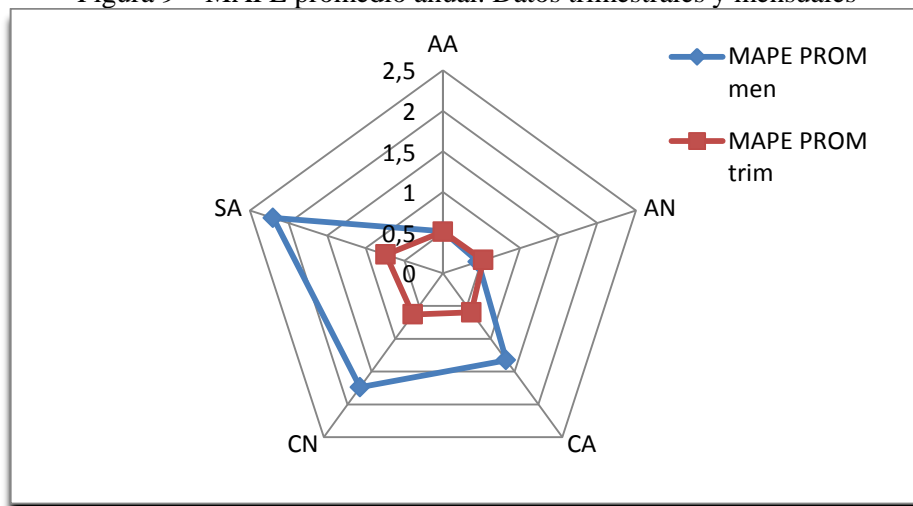
	AA	AN	CA	CN	SA	Prom. Men.
CB	0.3654	0.3384	0.8558	1.2167	1.5875	0.8727
ARI	0.2585	0.2409	0.7699	0.8828	1.1599	0.6624
ARI-EXO	1.0218	0.8272	1.0105	1.0608	1.6662	1.1173
ANFIS	0.4823	0.4807	0.7130	0.8520	1.1936	0.7443
ANFIS EXO	0.4416	0.3823	3.2771	4.6804	5.3952	2.8353
MAPE PROM men	0.5139	0.4539	1.3253	1.7385	2.2005	1.2464

Fuente: Elaboración propia

El ajuste de los modelos sobre los datos mensuales resulta en un mejor comportamiento integral del método ARIMA. De todas formas los productos AA y AN siguen produciendo los mejores resultados.

En Figura 9 se presentan los valores de MAPE promedio Anual para cada uno de los productos, analizando los datos trimestrales y mensuales.

Figura 9 – MAPE promedio anual. Datos trimestrales y mensuales



Fuente: Elaboración propia

Es notable la marcada diferencia entre el comportamiento de los modelos cuando se trata de datos trimestrales y mensuales. En todos los productos, los valores de MAPE cuando se analizaron los datos en forma trimestral dieron menores.

En Tablas 4 y 5 se presentan los datos estadísticos de las series de ventas históricas mensuales y trimestrales, respectivamente, para cada uno de los productos.

Tabla 4 – Datos estadísticos ventas reales mensuales

	AA MEN	AN MEN	CA MEN	CN MEN	SA MEN
MIN	10	30	9	4	2
MAX	1312	962	4368	2219	2036
MEDIA	235	224	656	482	512
STD	210.098	144.197	724.485	423.057	399.700
CV	0.894	0.643	1.105	0.878	0.781

Fuente: Elaboración propia

Las características de la serie de datos mensuales muestran que AN es la serie es que presenta menor Coeficiente de Variación (CV), seguida de CN y AA, mientras que CA presenta el mayor valor. Este coeficiente indica las características de variabilidad de las series y estaría indicando entonces que AN es una serie mas “suave” y por el contrario que CA sería la serie mas turbulenta. También resulta del análisis que las ventas de CA cuadriplican en valor a las de AN.

Tabla 5 – Datos estadísticos ventas reales trimestrales

	AA TRIM	AN TRIM	CA TRIM	CN TRIM	SA TRIM
MIN	174	156	109	120	72
MAX	1820	1815	7254	3527	3957
MEDIA	705	672	1967	1445	1535
STD	429.384	326.329	1523.762	848.270	843.261
CV	0.609	0.485	0.775	0.587	0.549

Fuente: Elaboración propia

La variabilidad en las series históricas, expresadas a través del Coeficiente de Variación, disminuye alrededor de un 30% cuando se analizan los datos en forma trimestral.

Esto contribuiría a que los modelos que se aplican para predecir el comportamiento puedan ajustar mejor en estos casos y dar como resultado la menor MAPE como se observa en figura 10. De hecho la MAPE promedio trimestral es alrededor de un 50% menor que la MAPE promedio mensual (ver Tablas 2 y 3).

4.2 Análisis de los resultados utilizando los datos de cada producto en forma Mensual

En Tabla 6 se presenta a título de ejemplo, los pronósticos obtenidos para el último año, con el modelo de HWA, en el caso de AA mensual, con la MAPE asociada a cada uno de los pronósticos. A fin de considerar la robustez del parámetro, se calcula la desviación estándar del mismo.

Tabla 6 – Salida de aplicación de método **HWA** a las ventas mensuales de AA

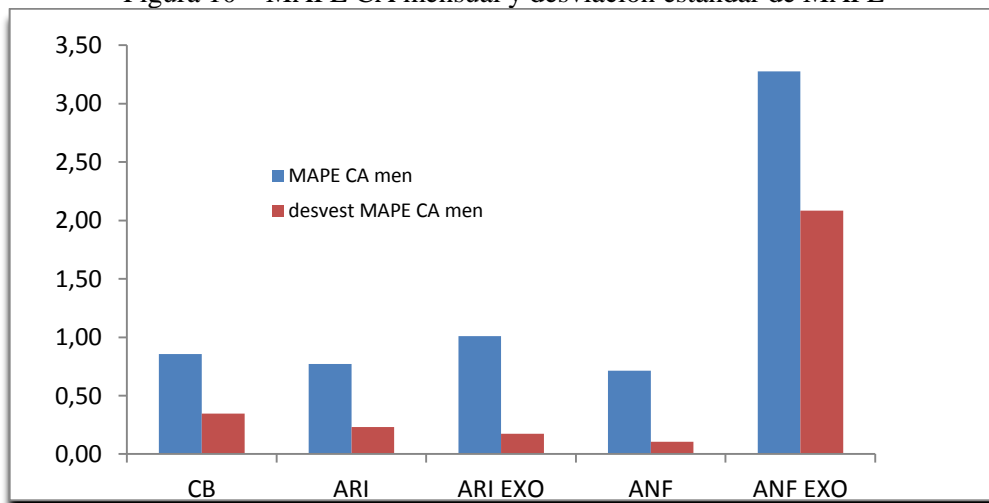
	VR AA	F AA CB men	MAPE
--	-------	-------------	------

VR AA	F AA CB men	MAPE
313	213	0.31894
350	289	0.24597
495	646	0.26588
160	304	0.42459
71	94	0.40476
187	78	0.43438
173	110	0.42428
165	107	0.41488
256	172	0.40540
225	211	0.37113
288	262	0.34559
239	274	0.32873
	Promedio	0.36538
	Desvest	0.06426064

Fuente: Elaboración propia

La Figura 10 presenta las MAPE obtenidas para el producto AA según los distintos métodos aplicados y su desviación estándar.

Figura 10 – MAPE CA mensual y desviación estándar de MAPE



Fuente: Elaboración propia

Para el producto CA los métodos ANFIS y ARIMA se comportan en forma adecuada, dado que ambos métodos arrojan la MAPE más pequeña y su desviación estándar también lo es. Esto significa que cualquiera de los dos métodos podría aplicarse al cálculo del pronóstico de este producto considerando los datos mensuales.

En Tabla 7 se presentan los valores de MAPE y su desviación estándar asociada para cada uno de los productos, según el método utilizado.

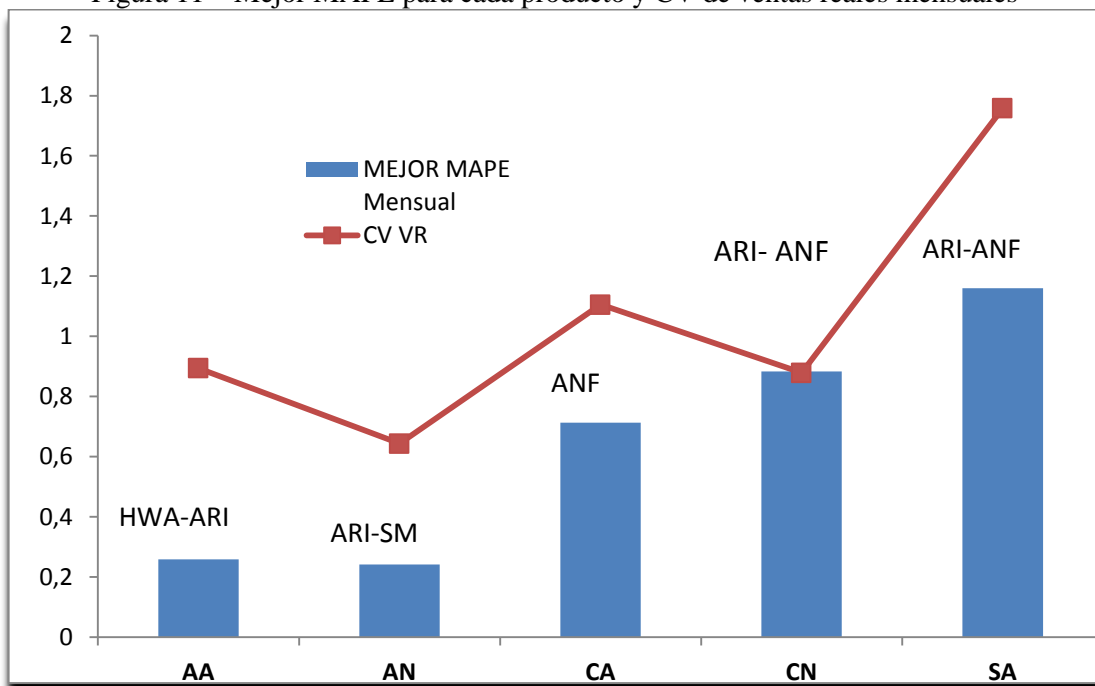
Tabla 7 – MAPE y Desviación Estándar para cada producto, según método aplicado datos mensuales

	CB	ARI	ARI EXO	ANF	ANF EXO
MAPE AA	0.3654	0.2585	1.0218	0.4823	0.4416
Desv. Std MAPE AA	0.0643	0.0804	0.4358	0.0888	0.0922
MAPE AN	0.3384	0.2409	0.8272	0.4807	0.3823
Desv. Std MAPE AN	0.0163	0.0155	0.1043	0.0441	0.0098
MAPE CA	0.8558	0.7699	1.0105	0.7130	3.2771
Desv. Std MAPE CA	0.3477	0.2310	0.1728	0.1061	2.0828
MAPE CN	1.2167	0.8828	1.0608	0.8520	4.6804
Desv. Std MAPE CN	0.1004	0.0334	0.0452	0.0619	0.1283
MAPE SA	1.5875	1.1599	1.6662	1.1936	5.3952
Desv. Std MAPE SA	0.2484	0.2519	0.5069	0.2360	0.7535

Fuente: Elaboración propia

En la Figura 11 se presentan los valores menores de MAPE hallados para cada producto, indicando específicamente el mejor método. También se muestra el Coeficiente de Variación de la serie de las ventas reales correspondiente a cada uno de los productos.

Figura 11 – Mejor MAPE para cada producto y CV de ventas reales mensuales



Fuente: Elaboración propia

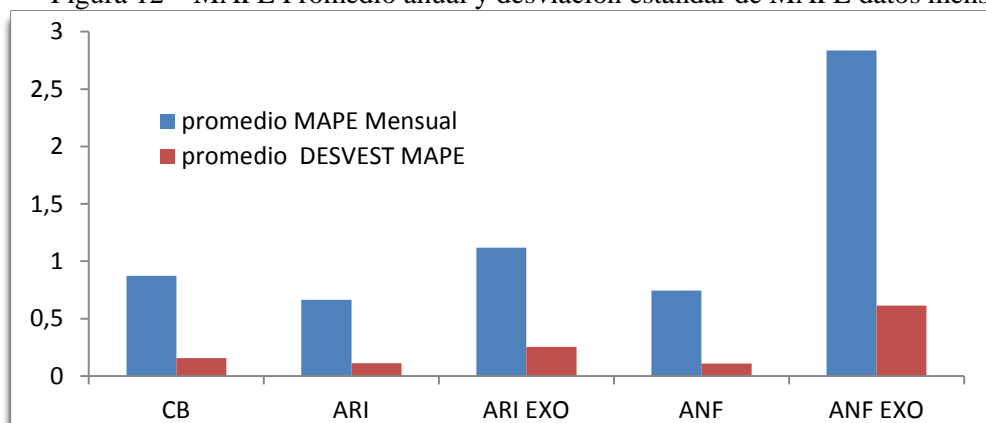
Por un lado se observa que en el caso de analizar los datos en forma mensual, el método de ARIMA se presenta como uno de los modelos que mejor ajusta, considerando en forma paralela su robustez. En el caso de CA, el mejor modelo resultó ser ANFIS.

No obstante, el método ANFIS también tuvo un buen comportamiento en general para el resto de los productos, aunque un poco mayor que ARIMA.

Se puede encontrar correspondencia directa entre el Coeficiente de Variación de la serie de ventas históricas y la mejor MAPE hallada, para todos los productos excepto para CA.

En la Figura 12 se presentan los promedios de todos los productos de MAPE para el año completo y su correspondiente valor de desviación estándar asociado.

Figura 12 – MAPE Promedio anual y desviación estándar de MAPE datos mensuales



Fuente: Elaboración propia

Se puede advertir de la misma Figura 12 que los métodos ARIMA EXO y ANFIS EXO, no han mostrado un buen comportamiento, indicando este resultado que las variables seleccionadas no contribuyen en forma adecuada a mejorar el ajuste del modelo.

4.3 Análisis de los resultados utilizando los datos de cada producto en forma Trimestral

El tratamiento de los datos en forma trimestral se realizó en forma similar al efectuado a los datos mensuales. En Tabla 8 se presentan los valores de MAPE y su desviación estándar asociada para cada uno de los productos, según el método utilizado.

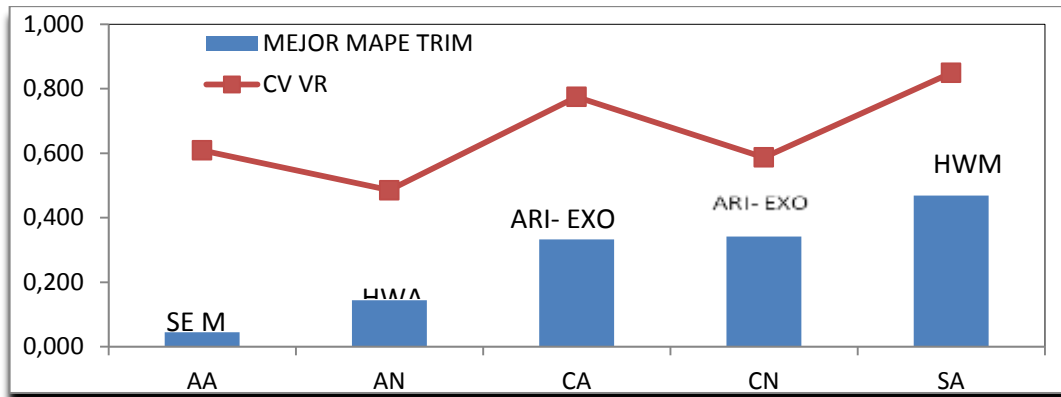
Tabla 8 – MAPE y desviación estándar para cada producto, según método aplicado datos trimestrales

	CB	ARI	ARI EXO	ANF	ANF EXO
MAPE AA	0,045	0,643	0,257	0,435	0,397
Desv. Std MAPE AA	0,013	0,044	0,142	0,074	0,020
MAPE AN	0,145	0,531	0,301	0,496	0,627
Desv. Std MAPE AN	0,047	0,072	0,017	0,043	0,162
MAPE CA	1,161	0,570	0,333	0,576	0,997
Desv. Std MAPE CA	0,189	0,116	0,046	0,100	0,128
MAPE CN	0,466	0,578	0,382	0,690	0,959
Desv. Std MAPE CN	0,074	0,022	0,028	0,007	0,058
MAPE SA	0,469	0,826	0,655	0,973	1,262
Desv. Std MAPE SA	0,039	0,189	0,187	0,211	0,286

Fuente: Elaboración propia

La Figura 13 presenta los valores menores de MAPE hallados para cada producto, indicando específicamente el mejor método y se relaciona con Coeficiente de Variación de la serie de las ventas reales correspondiente a cada uno de los productos.

Figura 13 – Mejor MAPE para cada producto y CV de ventas reales trimestrales



Fuente: Elaboración propia

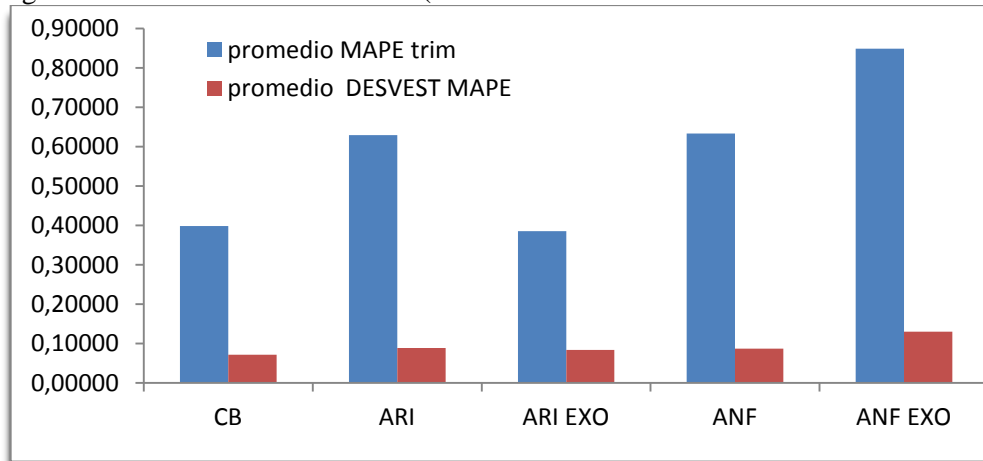
A diferencia de los datos mensuales, en el caso de los datos trimestrales, los métodos determinísticos ajustaron mejor a los atunes y a las sardinias, mientras que las caballas fueron mejor modeladas por ARIMA EXO.

Probablemente, el hecho de haber “suavizado” los datos, pasando de datos mensuales a trimestrales, permite que los métodos más simples ajusten mejor que los más complejos.

Por otro lado, la relación directa que se encuentra entre el Coeficiente de Variación y la MAPE para los datos mensuales, no es tan ajustada en el caso de los datos trimestrales, aunque se puede observar cierta tendencia de correlación entre ambos parámetros.

En la Figura 14 se presentan los promedios de MAPE para el año completo de todos los productos y su correspondiente valor de desviación estándar asociado.

Figura 14 – MAPE Promedio anual (desviación estándar de MAPE datos trimestrales)



Fuente: Elaboración propia

En este caso, los métodos determinísticos y el ARIMA EXO son los que en general mostraron el mejor comportamiento.

Por otra parte, es posible apreciar en la figura, el aumento de la robustez de los parámetros, observándose para todos los métodos que la desviación estándar de la MAPE es casi constante y sensiblemente menor que en para el caso de los datos mensuales.

5 CONCLUSIONES

Se aplicaron métodos de pronósticos de distintos orígenes a las ventas de cinco productos, pertenecientes a una familia de conservas de pescado, utilizándose los datos en forma mensual y trimestral.

Los métodos que se aplicaron fueron: métodos determinísticos, Método ARIMA, método ARIMA corregido por variables exógenas, método ANFIS y método ANFIS corregido por variables exógenas. Los resultados, en términos de precisión, variaron según el producto y la forma de agrupar los datos.

Cuando los datos se consideraron en forma mensual (108 datos utilizados para pronosticar 12), ARIMA fue el método que en general, arrojó menor MAPE y dichos valores fueron los más robustos. El método ANFIS también logró un buen desempeño en términos de valores de MAPE y robustez.

Cuando los datos se consideraron en forma trimestral (36 datos para pronosticar los siguientes 4) se obtuvo que los métodos más simples agrupados en este trabajo bajo el término Métodos Determinísticos, lograron un buen ajuste, siempre considerando los menores valores de MAPE y su robustez asociada. En este caso el método ARIMA EXO, logró también un buen ajuste con las caballas (productos CA y CN).

El hecho de incorporar las variables exógenas a los modelos de ANFIS y ARIMA, no generó la mejora esperada. Esto es claramente indicativo de que dichas variables incorporadas no contribuyen a la mejora del modelo y es necesario encontrar otras. Un estudio de componentes principales podría contribuir a perfeccionar estas entradas.

A modo de conclusión general, se puede indicar que:

- No se encontró un único método para todos los productos.
- La menor variabilidad de las series históricas de ventas permitirían obtener mejores resultados en términos del error del pronóstico.
- La buena respuesta del método ARIMA en el caso de los datos mensuales podría atribuirse a que se trabaja con mayor cantidad de datos. No obstante cuando se redujo

la cantidad de datos (datos trimestrales) ARIMA siguió demostrando ser un buen método cuando se trata de pronosticar series con marcada estacionalidad.

- Los métodos determinísticos también lograron una buena respuesta, fundamentalmente con los datos trimestrales, es decir, con series que poseen menor variabilidad intrínseca. Estos métodos son mucho más sencillos en su aplicación y en la interpretación de los resultados, de modo que respecto de los costos, resultan de mayor accesibilidad.
- Finalmente, la ventaja de ANFIS es que puede mejorarse su respuesta variando los parámetros del modelo. Esto puede ser motivo de mayor consumo de tiempo de cálculo computacional. La decisión acerca de esto dependerá en gran medida de los costos que el error del pronóstico genera en la organización.
- La cuantificación del error de los modelos a través de la MAPE o de cualquiera de los indicadores presentados en este trabajo, no consideran el efecto del signo del error para evaluar la bondad del modelo. No obstante, si se considera la administración de un almacén, por ejemplo, un error positivo del pronóstico implicaría el aumento de los costos de posesión y un error negativo, el costo de ruptura de las existencias. Estos costos, aunque proporcionales a la cantidad de unidades contenidas en el error, no poseen la misma relevancia, cualquiera sea la organización analizada. A partir de los resultados precedentes, y considerando los conceptos ya mencionados en la introducción, relacionados con el signo del error, se propone a futuro, el análisis de los efectos del error del pronóstico sobre la cadena productiva de la organización.

A COMPARATIVE STUDY OF MODELS FOR SALES FORECASTING

ABSTRACT: Most of the decisions to be made in operations, are based on forecasts and error or uncertainty demand associated with them. The aim of this work is to compare the accuracy of various models for forecasting sales. The analysis was performed using indicators such as MAPE and its standard deviation. Five forecasts were calculated from the same family products using methods from various sources: deterministic (exponential smoothing in simple and complex versions), ARIMA (Autoregressive Integrated Method of Moving Averages) and ANFIS (Adaptative Neuro-Fuzzy Inference System). In order to reduce the error methods, we analyzed other alternatives such as independent variables introduction in ARIMA and ANFIS models, that could affect the behavior of the series analyzed. The results

show that when the data were analyzed on a quarterly yield, they threw lower error values and this result is associated with less variability series. Also, simple methods can be used. However, it is important to note that while ANFIS and ARIMA models proved not to be the lowest MAPE in several cases, they have the advantage of be improved, either by changing the independent variables or by incorporated other parameters into the model, although this requires more expensive computational time.

Keywords: Forecasting. Exponential smoothing. ARIMA. ANFIS.

REFERENCIAS

BERENSON M.; LEVINE D.; KREHBIEL, T. **Estadística para la administración**, capítulo 11, 2º. Ed., Prentice Hall, p. 593-6602001.

CHOPRA, S.; MEINDL, P. **Administración de la cadena de suministro**, Estrategia, Planeación y Operación. Prentice Hall, 2008.

CHU, C. ZHANG, G. A comparative study of linear and nonlinear models for aggregate retail sales forecasting. **International Journal of Production Economics**, v. 86, p. 217-231, 2003.

CORRES, G.; ESTEBAN, A.; GARCÍA, J.; ZÁRATE, C. Metodología Box – Jenkins (ARIMA) con SPSS. Parte I. **Revista de la EPIO**, v. 29, p. 161-176, 2008.

FOGARTY, D.; BLACKSTONE, J.; HOFFMANN, T. **Administración de la producción e inventarios**. 2º. Ed., CECSA, p. 91-141, 235-281, 1999.

HANKE, J.; WHICHERN, D. **Pronósticos en los negocios**. Pearson Educación, 2006.

KERKKÄNEN, A.; KORPELA, J.; HUIKONEN, J. Demand forecasting errors in industrial context: measurement and impacts. **International Journal of Production Economics**, v. 118, p. 43-48, 2009.

LUXHOJ, J; RIIS, J; STENSBAALLE, B. A hybrid econometric – neural network modeling approach for sales forecasting. **International Journal of Production Economics**, v. 43, p. 175-192, 1996.

THOMASSEY S., HAPPIETTE M., CASTELAIN J. A global forecasting support system adapted to textile distribution. **International Journal of Production Economics**, v. 96, p. 81-95, 2005.

SANDERS, N., GRAMAN, G. Quantifying costs of forecast errors: a case of study of the warehouse environment. **OMEGA, International Journal of Management Science**, v. 37, p. 116-125, 2009.

XIE, J., LEE, T.S., ZHAO, X. Impact of forecasting error on the performance of capacitated multi-item production systems. **Computers and Industrial Engineering**, v. 6, n. 224, p. 200-219, 2004.

ZHANG G., PATUWO B. E., HU M. Forecasting with artificial neural networks: the state of the art. **International Journal of Forecasting**, v. 14, p. 35-62, 1997.

ZHANG, P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, p. 159-175, 2003.

Originals recebidos em: 06/12/2013

Aceito para publicação em: 13/08/2014