

**Estimación de un
modelo de vectores
autorregresivos para
pronóstico de demanda.
Un estudio empírico**

**Bruno de Jesús Rahmer
Hernando Garzón Saenz
José Solana Garzón**

Estimación de un modelo de vectores autorregresivos para pronóstico de demanda. Un estudio empírico

Estimation of a vector autoregression model for demand forecasting. An empirical study

BRUNO DE JESÚS RAHMER, HERNANDO GARZÓN SAÉNZ, JOSÉ SOLANA GARZÓN

Fundación Universitaria Tecnológico Comfenalco (Colombia)

brunodejesus.2509@gmail.com, hando2001@gmail.com, ingjosemsolana@gmail.com

Resumen: El diseño e implantación de robustos sistemas de pronóstico constituye una necesidad latente en las unidades organizacionales que subyacen en una atmósfera difusa como lo es el entorno mercantil, escenario en el cual convergen consumidores y ofertantes que establecen sendas relaciones comerciales. Este artículo tiene por objetivo construir y validar un modelo vectorial autorregresivo para la estimación cuantitativa de las ventas en una empresa manufacturera localizada en el clúster petroquímico de la ciudad de Cartagena de Indias, previa verificación de los supuestos formales y distribucionales necesarios para una estimación eficiente vía mínimos cuadrados ordinarios. Los modelos de esta naturaleza comparecen como una útil alternativa de modelización econométrica que además posibilita la identificación de relaciones de causalidad contemporáneas entre las variables analizadas, así como también la generación de proyecciones en el mediano plazo. La modalidad de investigación es de tipo híbrido, pues en la fase inicial, el estudio se asimila a uno de tipo descriptivo retrospectivo y en la etapa conclusiva adquiere un carácter explicativo post-facto. Los resultados derivados de esta aplicación indican que la cuota de mercado y las estrategias de marketing contribuyen significativamente a la estimación de las ventas netas mensuales del producto ofertado y que el modelo propuesto, posee una eficiencia

Abstract: The design and implementation of robust forecasting systems constitutes a latent need in the organizational units that underlie a diffuse atmosphere such as the commercial environment, in which consumers and suppliers converge and commercial relationships are established. The purpose of this article is to construct and validate a vector autoregressive model for the quantitative estimation of sales in a manufacturing company located in the petrochemical cluster of the city of Cartagena de Indias, after verification of the formal and distributional assumptions necessary for an efficient estimate via ordinary least squares. Models of this nature appear as a useful alternative for econometric modeling that also enables the identification of contemporary causality relationships among the variables analyzed, as well as the generation of prospects in the medium term. The research modality is a hybrid type, because in the initial phase, the study is retrospective and descriptive type, in the conclusive stage is a post-facto explanatory study. The results derived from the application of the proposed methodology indicate that the market share and promotions contribute significantly to the estimation of the net monthly sales of the product offered and that the model has a non-negligible predictive efficiency, in contrast to classical models validated in this context, arguing therefore that such is susceptible to be extrapolated in other manufacturing contexts.

predictiva nada despreciable, en contraste con los modelos clásicos validados en este contexto, arguyéndose por consiguiente que tal es susceptible a ser extrapolado en otros contextos fabriles.

Palabras clave. Econometría, industria metalúrgica, series temporales

Código JEL: C3, L25

Key words. Econometrics, metalurgy industry, temporal series

JEL Code: C3, L25

RAHMER, BJ, GARZÓN H, SOLANA J (2019). Estimación de un modelo de vectores autorregresivos para pronóstico de demanda. Un estudio empírico. CEASGA-working papers, 2(2), 92-113.

ISSN: 2173-5859

info@ceasga.es

1. INTRODUCCIÓN

En clústeres industriales la adquisición de ventajas competitivas diferenciales y el alcance de mejoras sustantivas en el proceso de planificación de ventas y operaciones a través del análisis de previsión y estrategias de la inteligencia empresarial tal y como lo precisan Hani, Duen-Ren y Hai Hoang (2016) es una perentoriedad debido a la presión ejercida por el entorno mercantil. Por tanto, el ejercicio de pronosticar sistemáticamente del comportamiento futuro de variables críticas como el nivel de demanda de bienes y servicios en el entorno mercantil donde tales son intercambiados, es una labor determinante que permite erigir alternativas de planeación viables y prever la intensidad de uso de los factores productivos (Delgadillo-Ruiz, Ramírez-Moreno, Leos-Rodríguez, Salas González y Valdez-Cepeda, 2016) en procura de adaptar la estrategia mercadotécnica y de este modo, paliar la incertidumbre consustancial a este escenario. En consecuencia, la integración de técnicas de pronósticos y estrategias de modelización avanzada para una proyección verosímil de las demandas futuras, permiten estructurar sistemas de planeación flexibles y ajustar anticipativamente los componentes de la oferta en función de las demandas futuras.

Es consabido que la obtención de predicciones a través del tratamiento de series temporales multivariantes tiene sus bemoles y sostenidos, pues a pesar de ostentar un alto nivel de eficiencia, en contraste con aquellos modelos basados en series temporales escalares que son incapaces de extraer cúmulos de información provenientes de ciertos componentes que ejercen una influencia preponderante en los fenómenos analizados. En este sentido, el modelo vectorial autorregresivo (VAR), surge como una metodología de identificación de relaciones estructurales para series temporales multivariantes y contempla

los acoplamientos dinámicos entre las variables analizadas e incorporan información no estadística. Están compuestos por un conjunto de ecuaciones en forma reducida y carentes de restricciones, es decir, que no existe distinción apriorística entre variables endógenas y exógenas y el mismo conjunto de variables explicativas aparece en cada ecuación del sistema. De esta forma, cada variable es expresada como una función lineal de sus valores retardados y de los rezagos de las demás variables incluidas en la estructura funcional del modelo. Este planteamiento es plausible si se considera que dentro de un sistema complejo –cualquiera que éste sea– existe interdependencia en cualquier instante, entre sus elementos constitutivos y, por lo tanto, es prácticamente imposible establecer, ex ante, la condición de exogeneidad-endogeneidad partiendo tan sólo de postulados e hipótesis que se subsumen en algún marco teórico específico (Loria, 2014).

Profusamente se han estimado y validado modelos VAR que utilizan únicamente como input a variables de orden macroeconómico (Zhang, Miner, Geng y Jiang, 2017) (Trujillo Calagua, 2015) a fin de evaluar sus interacciones bajo una perspectiva largoplacista. El modelo propuesto en el presente *paper* mantiene grandes contactos con algunos estudios homólogos presentados en el contexto microeconómico dado que se establece como objetivo cardinal la estimación de la demanda en el corto plazo (Li, Li y Zhang, 2017).

Acótese que la presencia de tales modelos en escenario de análisis a nivel micro prácticamente marginal, por tanto, el modelo aquí desarrollado puede concebirse como una alternativa de pronóstico a corto plazo que proporciona un invaluable soporte para el proceso de toma de decisiones en el dominio de la planificación de las operaciones para empresas del sector manufacturero cartagenero. Queda ratificado que la utilidad de este modelo reside, no únicamente en su capacidad proyectiva, sino también en su aptitud para identificar las dinámicas de interrelación subyacentes en el sistema multicuacional y los posibles movimientos concomitantes entre las variables analizadas, a saber: ventas, promoción e inversión en capital. Se afirma, por consiguiente, que el abanico de posibilidades y métodos para el tratamiento de series temporales y la subsiguiente proyección del comportamiento futuro de la demanda en este espacio geográfico y más concretamente, en este clúster, es amplificado con la introducción de la metodología aquí propuesta.

2. METODOLOGÍA

El enfoque investigativo es de carácter mixto en virtud del enlazamiento de dos modalidades investigativas bien diferenciadas. En instancia inicial, el presente estudio adquiere un carácter netamente descriptivo y retrospectivo (Barreras Serrano, Sánchez López, Figueroa Saavedra, Olivas Valdez y Pérez Linares, 2014). En esta fase se seleccionan diversas submuestras de datos con frecuencia trimestral, correspondientes a las series de ventas netas de derivados de hidrocarburos, nivel de inversión en estrategias

de promoción y marketing de los bienes ofertados (en términos agregados) e inversión de capital de 8 empresas manufactureras localizadas en el clúster petroquímico plástico Cartagenero. La información está comprendida en el período 2012-2018.

En etapas posteriores la plataforma de investigación empírica trasciende la mera descripción de los fenómenos examinados y está dirigida dilucidar las relaciones de causalidad contemporáneas entre las variables seleccionadas a fin de construir y validar un modelo vectorial autorregresivo con el objetivo de prospectar los valores futuros de la demanda agregada. Se hipotetiza que los factores condicionantes de las fluctuaciones de la demanda son: la inversión en el despliegue estrategias de marketing y la inversión en capital. Ésta última variable designa los montos invertidos en activos enmarcados en tres categorías bien diferenciadas, como lo son: propiedad, maquinaria o equipamiento.

Se descarta del análisis, la variabilidad de los bienes sustitutivos y complementarios se parte de la asunción que los cambios en las preferencias son prácticamente inescrutables y que la renta media disponible de los consumidores finales, esto es, el flujo monetario percibido por los individuos que le permite adquirir los bienes ofertados es prácticamente uniforme en todos los periodos referenciales. La evolución de los precios obedece a factores externos globales y sectoriales y otros factores controlados por las unidades estratégicas de negocio.

En la fase inicial del proceso investigativo se realiza la estimación vía mínimos cuadrados ordinarios del modelo vectorial autorregresivo que incorpora las tres variables selectas. La selección de la longitud óptima del rezago se efectúa atendiendo a diversos criterios de información como el de Akaike o Schwarz (Suharsono, Aziza y Pramesti, 2017). Subsiguientemente se procede a determinar si las variables guardan entre sí relaciones de causalidad en el sentido Granger a través de la aplicación de un test estadístico que lleva la misma denominación.

En los apartes posteriores se reportan los resultados de la aplicación de una batería de tests estadísticos direccionados a verificar el cumplimiento de los supuestos clásicos tradicionales que garantizan la consistencia y eficiencia en la estimación del modelo vectorial autorregresivo vía mínimos cuadrados clásicos. Tales son: la prueba de autocorrelación de Portmanteau, el test de Heteroscedasticidad de White y la prueba de normalidad de Jarque-Bera. Una vez verificados los supuestos formales, se procede a generar las funciones impulso respuesta-con el objeto de identificar las interacciones dinámicas que caracterizan el sistema multiecuacional- y las predicciones de la demanda a partir de los datos históricos colectados a efectos de evaluar su eficiencia predictiva y contrastar los resultados con otros modelos de series temporales. En el cuadro 1 se reporta un resumen descriptivo de las variables seleccionadas para la construcción del modelo. La base de datos fue construida a través de un extensivo trabajo de campo.

Cuadro 1. Resumen de las variables seleccionadas

Variable	Unidad de medida	Media	Mediana	Desviación estándar	Tamaño muestral
Ventas	Miles de Pesos colombianos)	3235,4	1198,1	25,0982	132
Inversión en Marketing	Miles de Pesos colombianos)	266	244,03	12,0912	132
Inversión en capital	Miles de Pesos colombianos)	786,45	768,21	19,0983	129 ¹

¹ La presencia de valores perdidos en el estudio obedece a la imposibilidad para capturar la información de la inversión en capital en la fase inicial del estudio. Las variables han sido estandarizadas para suprimir el efecto de las escalas en la estimación del modelo.

Fuente: Elaboración propia

3. RESULTADOS EMPÍRICOS Y DISCUSIÓN

3.1. Estimación del modelo de vectores autorregresivos

Considérese entonces, el siguiente proceso estocástico vectorial VAR(P) generador de múltiples series temporales:

$$Y_t = \psi + \Pi_1 y_{t-1} + \dots + \Pi_n y_{t-n} + \beta x_t + \mu_t$$

Que en su forma matricial equivale a:

$$\begin{bmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-n+1} \end{bmatrix} = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_n \end{bmatrix} + \begin{pmatrix} \Pi_1 & \Pi_2 & \dots & \Pi_{n-1} & \Pi_n \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-n+1} \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-n+1} \end{bmatrix} + \begin{bmatrix} \mu_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Siendo $Y_t = (y_{1t}, y_{2t}, \dots, y_{mt})'$ el vector $k \times 1$ de las variables endógenas, $x_t = (x_{1t}, x_{2t}, \dots, x_{mt})'$ el vector $k \times 1$ de variables exógenas, $\Pi_t = (\Pi_{1t}, \Pi_{2t}, \dots, \Pi_{mt})'$ la matriz de coeficientes de las variables exógenas y μ_t el vector $k \times 1$ de innovaciones, con estructura de ruido blanco, es decir que su esperanza es $E(\mu_t) = 0 \forall t$ y $Var(\mu_t) = \Sigma$ que permanece invariable (Nicholson, Bien y Matteso, 2016).

Debido a que los supuestos de Gauss-Markov se verifican en cada ecuación del sistema, la estimación por mínimos cuadrados ordinarios (MCO) es insesgada, asintóticamente eficiente y consistente. Métodos de estimación basados en información completa como el de máxima verosimilitud no poseen ventaja comparativa alguna, pues en la matriz de coeficientes no existen restricciones y las variables explicativas son todas retardadas y no están correladas con los términos de error.

Reescribese entonces el proceso VAR(P) en una forma compactada:

$$Y = BZ + U$$

Considérese entonces el estimador insesgado $B \hat{=} YZ'(ZZ')^{-1}$ y el estimador de la matriz de covarianza $Y = \frac{1}{T}(Y - \hat{B}Z) \frac{1}{T}(Y - \hat{B}Z)'$ a sabiendas de que $B = (\phi_1, \phi_2, \dots, \phi_p, \beta)$ $Z = (Z_1, Z_1 \dots Z_T)$. La selección la longitud óptima del rezago del modelo vectorial autorregresivo se efectúa atendiendo simultáneamente a ciertas medidas de calidad relativa como los son los criterios de información convencionales dado por $AIC = -\frac{2L}{T} + \frac{2k}{T}$ y de Schwarz cuya fórmula es $AIC = -\frac{2L}{T} + \frac{k \log(T)}{T}$ y a la evaluación de la significancia individual de los coeficientes estimados. Los criterios reportados se derivan de una minimización insesgada del valor esperado del criterio de Kullback-Leibler (Zhang, Zou y Carroll, 2015). Se sigue entonces, que un modelo con orden óptimo será aquel que minimice los criterios evaluados en contraste con modelos alternativos de distintos órdenes.

Obsérvese que en el Cuadro 2 se computan una serie de estadísticos globales del modelo, entre ellos: El estadístico R^2 que da cuenta sobre grado de ajuste del modelo, Se refrenda la idoneidad del modelo que contiene 2 rezagos, pues la influencia marginal de retardos adicionales es desdeñable y, por tanto, su inclusión derivará en la pérdida de parsimonia y la reducción de grados de libertad. Adicionalmente, el modelo en cuestión produce un valor mínimo de los criterios de información reportados respecto a otros modelos tentativos. El determinante de la varianza residual admite un valor casi inescrutable, deduciéndose, por tanto, que la estimación del modelo es eficiente.

3.1.1 Contraste de auto-correlación

En este apartado se aplica el test de Portmanteau que evalúa la significancia conjunta de las autocorrelaciones residuales hasta el rezago especificado en la fase precedente. Así, la hipótesis nula es $H_0 = (\rho_1, \rho_2, \dots, \rho_h) = 0$ y el estadístico construido se traza como: $Q_h = T \sum_{i=1}^h tr(\hat{C}'_i, \hat{C}_0^{-1}, \hat{C}_i, \hat{C}_0^{-1})$, aproximándose asintóticamente a una distribución χ^2 con grados de libertad $k^2(h - p)$.

Cuadro 2. Estimación del modelo

	VENTAS (ω)	MARKETING (φ)	INVERSION EN CAPITAL (π)	MEDIDAS DE BONDAD DEL MODELO			
$\omega_{(-1)}$	-0.024	0.1035	0.076	R-Cuadrado	0.9045	0.8923	0.8355
	(0.077)	(0.064)	(0.125)	R-Cuadrado	0.8973	0.884132	0.8230
	[-0.312]	[1.601]	[0.605]	Suma de			
$\omega_{(-2)}$	-0.0092	0.1165	0.0822	Cuadrados	2.53E+08	1.74E+08	6.60E+08
	(0.078)	(0.065)	(0.126)	residuales			
	[-0.117]	[1.791]	[0.649]	Akaïke AIC	17.89439	17.52213	18.85346
$\varphi_{(-1)}$	0.2505	0.3217	0.3065	Schwarz SC	18.094	17.190	19.05324
	(0.128)	(0.107)	(0.208)	Media (Ventas)	9546.2	7612.9	16157.27
	[1.942]	[3.005]	[1.471]	SD (Ventas)	5582.5	4362.6	6869.958
$\varphi_{(-2)}$	0.2855	0.3039	0.242	Determinante Covarianza			
	(0.133)	(0.1105)	(0.215)	Residual (Ajustada GL)			
	[2.144]	[2.7491]	[1.127]	5.57E+19			
$\pi_{(-1)}$	-0.1132	0.0379	0.2643	Determinante Covarianza			
	(0.066)	(0.055)	(0.107)	Residual			
	[-1.698]	[0.685]	[2.453]	-2310.231			
$\pi_{(-2)}$	0.6258	0.0511	0.2679	Log-Verosimilitud			
	(0.0659)	(0.05471)	(0.106)	54.81400			
	[9.497]	[0.934]	[2.517]				
\mathbb{C}	-2074.84	-415.585	2262.69				
	(542.53)	(450.392)	(876.36)				
	[-3.824]	[-0.922]	[2.581]				

Fuente: Elaboración propia

En el Cuadro 3 se computa esta prueba de autocorrelación en los residuos de un modelo en la que examina si alguna autocorrelación, es significativamente distinto de cero. Para efectos de simplificación sólo figuran 9 rezagos. Dado que el p-valor asociado a las autocorrelaciones residuales en cada uno de los rezagos reportados (excluyendo los rezagos inferiores o iguales al orden del modelo VAR) es superior a 0,05 no se rechaza la hipótesis nula de inexistencia de autocorrelación y por consiguiente el modelo VAR

captura apropiadamente la información provista por las variables elegidas (Bose, Hravnak y Sereika, 2017).

Cuadro 3. Contraste de autocorrelación

Rezagos	Q-Stat	Prob.	Q-Stat ajustado	Prob.	Grados de Libertad
1	4.632355		4.686853		
2	11.99586		12.22568		
3	15.17828	0.086201	15.52312	0.077501	9
4	18.12747	0.447312	18.61618	0.415803	18
5	20.65673	0.802104	21.30157	0.772006	27
6	24.59332	0.925006	25.53340	0.902806	36
7	26.29909	0.988301	27.39032	0.982201	45
8	30.68506	0.995610	32.22613	0.991902	54
9	40.26116	0.988630	42.92151	0.975204	63

Fuente: Elaboración propia

3.1.2 Contraste de normalidad de los residuos

A fin de verificar el supuesto de normalidad residual se utiliza el test de Jarque Bera, disponiendo del método de factorización de Cholesky. El estadístico de la prueba viene dado por: $J - B = T \left[\frac{m_3^2}{6} + \frac{(m_4^2 - 3)}{24} \right]$ y que sigue una distribución χ^2 y m_3 y m_4 representan el sesgo (tercer momento) y la curtosis (cuarto momento), respectivamente, cuyos estimadores vienen dados por $m_3 = \sum_t v_t^3 / T$ y $m_4 = \sum_t v_t^4 / T$ y $v_t = P u_t \sim N(0, I_k)$ siendo P la matriz de factorización de $k \times k$, determinada a partir de la expresión $P = G D^{1/2} G'$, donde D es la matriz diagonal que conjunta los eigenvalores propios de la matriz residual de varianzas y covarianzas, mientras que G es la matriz cuyas columnas corresponden a los autovectores (Sánchez López Cruz y Marcelo, 2016).

A tenor de los resultados provistos por el test conjunto de Jarque Bera reportado en el Cuadro 4, que responde a la expresión funcional $\lambda = \lambda_3 + \lambda_3 \rightarrow \chi^2(2k)$ se corrobora el hecho de que los valores residuales se comportan según una distribución normal multivariada, pues en todas las pruebas registradas el p-valor es superior a 0,05.

Cuadro 4. Test de Jarque Bera

Componente	Asimetría	Chi-Square χ^2	Grados de Libertad	Prob.
Conjunto		4.5671	3	0.206383
Componente	Kurtosis	Chi-Square χ^2	Grados de Libertad	Prob.
Conjunto		3.9134	3	0.270968
Test de Jarque-Bera	Grados de Libertad	Prob.		
Conjunto	6	0.0000		

Fuente: Elaboración propia

El contraste de cointegración de Johansen es un método aplicable a series temporales para determinar cuál es el rango de cointegración de un vector de variables $I(1)$. En caso tal de hallar un orden de integración común entre las variables, existirá una ecuación de co-integración entre ellas, que permite inferir la existencia de una relación de largo plazo. (Córdova Olivera, 2014). En caso tal de que se acepte la hipótesis de contención de raíz unitaria en las series en niveles, se requiere de diferenciación de primer o n orden para lograr la condición de estacionariedad. Se llega a la conclusión de que las series están integradas en $I(0)$ en tanto el estadístico Dickey-Fuller Aumentado reporta valores superiores a cada uno de los valores críticos de Mackinnon para diferentes niveles de confianza estudiados, a saber, 99% y 95%.

La prueba de traza está expresada por:

$$\lambda_{traza} = -T * \left[\sum_{i=1}^s \ln(1 - \lambda_i) \right]$$

Siendo s la cantidad de vectores de co-integración y λ_i los valores propios de la matriz $Y = \left[\sum_{j=1}^k \Pi_j \right] - I_r$. A partir del resultado de la prueba de traza provisto en el Cuadro 5 se depende la no existencia de relaciones de co-integración al 5% de significancia, pues las variables son estacionarias.

Cuadro 5. Test de cointegración

Hipótesis	Traza		0.05	
Cantidad de vectores de Co-Integración)	Eigenvalor	Estadístico	Valor Crítico	Prob.
0	0.025906	11.36519	12.51798	0.0773

Fuente: Elaboración propia

3.1.3 Contraste de heterocedasticidad de White

El test de White –que no precisa de una definición concreta de heterocedasticidad– es aplicado para verificar si la varianza es constante en la nube de datos mediante la regresión de los productos cruzados de los residuales sobre los productos cruzados de las variables explicativas. Para evaluar la significancia conjunta de la regresión se computa el estadístico $LM = T_g R_m^2$ que se aproxima asintóticamente a una distribución χ^2 con grados de libertad $n \left[\frac{k(k-1)}{2} \right]$. Se percibe en el Cuadro 6 los valores p asociados a los componentes individuales, con 72 grados de libertad es mayor a 0.05, por consiguiente, no se rechaza la hipótesis nula de homocedasticidad.

Cuadro 6. Contraste de White

Componentes Dependiente	R^2	$F(12, 73)$	Prob.
$res_1 * res_1$	0.090136	0.602646	0.804204
$res_2 * res_2$	0.076553	0.504303	0.883900
$res_3 * res_3$	0.124133	0.862163	0.556901
$res_2 * res_1$	0.025660	5.111689	0.904312
$res_3 * res_1$	0.053795	7.082660	0.787531
$res_3 * res_2$	0.131992	0.925050	0.499113

Fuente: Elaboración propia

Se procede a realizar el cómputo de las funciones impulso-respuesta para la examinación de las interacciones dinámicas. La dinámica del modelo estimado dependerá

intrínsecamente de la estructura de correlaciones contemporáneas manifiesta en la matriz de innovaciones.

3.2. CONTRASTE DE CAUSALIDAD DE GRANGER

Una vez seleccionado el número óptimo de retardos a considerar se aplica una prueba que reviste singular interés en el dominio econométrico, denominada contraste de causalidad de Granger. Intuitivamente, puede aseverarse que una variable Y_t es causada por X_t si ésta contribuye a mejorar su estimación. En caso contrario se cumple que: $E(Y_t / Y_{t-1}, \dots, X_{t-1}, \dots) = E(Y_t / Y_{t-1}, \dots)$. Debe anotarse que la prueba de Granger utiliza el concepto ampliado de correlación para detectar relaciones causales, y por tanto no es un indicador certero de verdadera causalidad, solo señala la precedencia temporal entre las variables de estudio y es un indicador en la previsión del comportamiento futuro de las variables (Morrone, 2015). Consideremos la primera regresión del modelo VAR(P):

$$X_t = \sum_{t=1}^n \delta_t Y_{t-1} + \sum_{t=1}^n \vartheta_t X_{t-1} + \mu_{1t}$$

Se confirma que Y_t no será causa en el sentido Granger de X_t si $\vartheta_1, \vartheta_2, \vartheta_3, \dots, \vartheta_t = 0$. Para contrastar formalmente tal asunción se ha computado el estadístico χ^2 . En el Cuadro 7 se computan los resultados del test de Granger. Un p-valor superior a 0,05 impide invalidar la hipótesis nula, indicando pues, que las variables analizadas guardan relaciones causales contemporáneas. Ello puede ser verificado igualmente con el resultado del test de Wald que en el que se rechaza la hipótesis nula de que los rezagos de las variables deben ser excluidos de cada ecuación del sistema. Merece un comentario aparte el comportamiento de la variable “Ventas” dado que, según los resultados suministrados, su comportamiento puede ser explicado por el efecto conjunto de la inversión del capital y la inversión en marketing.

Cuadro 7. Contraste de causalidad de Granger

VENTAS			
	Chi-Square χ^2	Grados de Libertad	P-Valor
MARKETING	14.48839	2	0.00070
INVERSION EN CAPITAL	94.88669	2	0.00000
Bloque	152.5511	4	0.0000
MARKETING			
	Chi-Square χ^2	Grados de Libertad	P-Valor
VENTAS	6.06830	2	0.04810
INVERSION EN CAPITAL	12.50907	2	0.00190
Bloque	11.95169	4	0.01770
INVERSION EN CAPITAL			
	Chi-Square χ^2	Grados de Libertad	P-Valor
VENTAS	9.42910	2	0.00890
MARKETING	5.89263	2	0.05000
Bloque	15.56631	4	0.00370

Fuente: Elaboración propia

3.3. FUNCIÓN IMPULSO-RESPUESTA

La gráfica de la función impulso-respuesta muestra el shock de las variables endógenas sobre las demás variables de sistema multiecuacional ante cambios en las innovaciones, de modo que un shock en una variable cualquiera en el período i afectará directamente a la misma y se transmitirá al resto de variables endógenas mediante la estructura dinámica representada por el modelo de vectores autorregresivos. Habitualmente las innovaciones son contemporáneamente interdependientes y por ende, los impulsos deben ser ortogonalizados a través de la descomposición de la matriz residual de covarianzas $\Sigma_u = W\Sigma_e W'$ mediante la factorización de Cholesky que proporciona una matriz triangular inferior P tal que $\Sigma_u = PP'$ y admitiendo por tanto, cierto grado de arbitrariedad para la atribución de los efectos comunes. Se acota que el VAR satisface la

condición de estabilidad en tanto que la totalidad de raíces del sistema $|I - \Phi z - \Phi_1 z^2 - \dots - \Phi_1 z^p| = 0$ se hallan fuera del círculo unitario y en virtud del Teorema de descomposición de Wold, el modelo puede representarse como un proceso de media móvil de orden infinito $VMA(\infty)$ del tipo (Shahin, Ali y Ali, 2014):

$$Y_t = \sum_{s=0}^{\infty} \Psi_s \varepsilon_{t-s}$$

Siendo los impulsos-respuestas para los shocks ortogonales $\Psi_s = \frac{\partial Y_t}{\partial \varepsilon_{t-i}} = \frac{\partial Y_{t+s}}{\partial \varepsilon_t}$ y $Var(Y_t) = \sum_{n=0}^{\infty} \Psi_s \Omega \Psi_s'$. Así, por ejemplo, el coeficiente $\Psi_{ij}(0)$ representa el impacto instantáneo de un cambio unitario en ε_{nt} sobre y_{mt} . En el mismo sentido, el coeficiente $\Psi_{ij}(1)$ representa la respuesta, un periodo posterior frente al cambio unitario en ε_{nt-s} sobre y_{mt} .

En la práctica los coeficientes autorregresivos no son conocidos con certeza, sin embargo, es plausible su estimación vía regresión por mínimos cuadrados ordinarios. Cuando tales valores son utilizados para el cálculo de Ψ_n es imperativo reportar los errores estándares.

Sea entonces: $\sqrt{T}(\Psi_{s,T} - \Psi_s) \xrightarrow{L} N(0, G_s(\Omega \otimes Q^{-1})G_s')$, los coeficientes impulso-respuestas estarán dados por la raíz cuadrada del elemento diagonal asociado a $\hat{G}_{s,T}(\hat{\Omega}_T \otimes \hat{Q}_T^{-1})\hat{G}_{s,T}'$, siendo $\hat{G}_{s,T} = \frac{\partial \Psi_s(\pi)}{\partial \pi'} \Big|_{\pi=\hat{\pi}_T}$ y $\hat{Q}_T = 1/T \sum_{t=1}^T X_t X_t'$. (Hamilton, 1994)

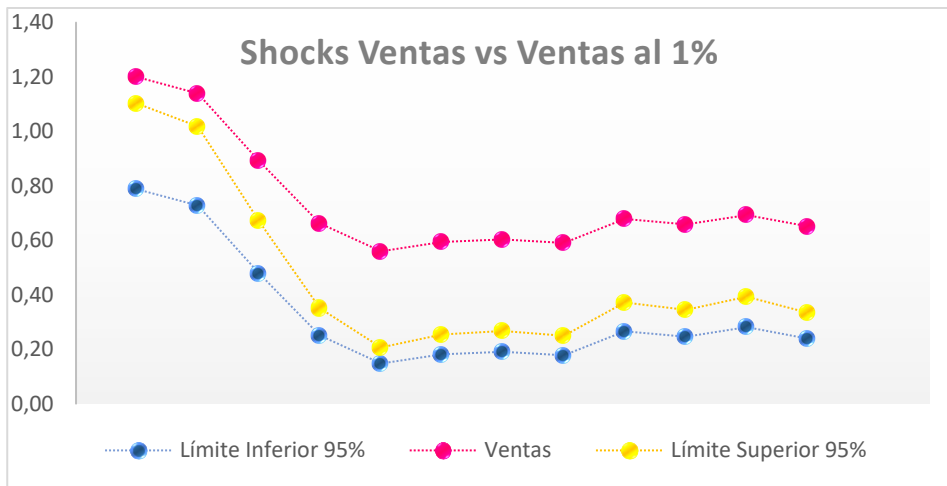
Las gráficas impulso-respuesta pueden entenderse como representaciones de medias móviles del sistema de ecuaciones. Se observa así un resumen de los efectos generados por los impulsos inducido sobre cada una de las variables localizadas en las columnas y que por homogeneidad se equipara al valor de su desviación típica del error. Los intervalos de confianza de las gráficas reportadas son construidos asumiendo un 95% de confianza. La lectura de estas representaciones es levemente simple: Los gráficos corresponden a las respuestas de la variable ventas ante los shocks de la misma y ante los shocks de las otras, esto es, marketing e inversión de capital.

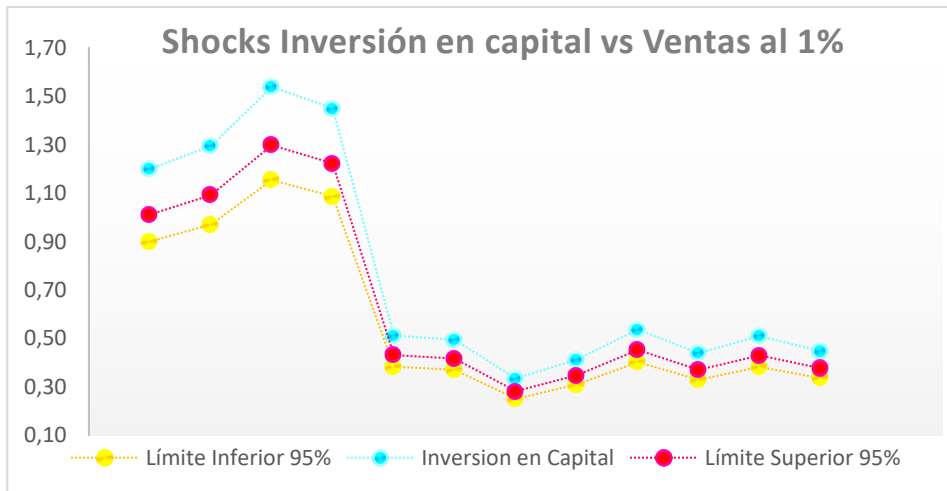
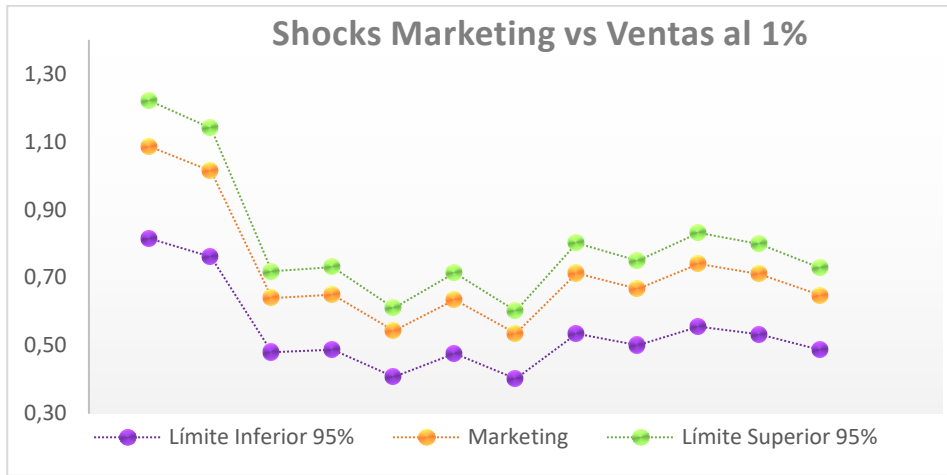
Nótese que tanto la demanda, como la inversión en estrategias de promoción y la inversión en capital están determinadas conjuntamente por un modelo multiecuacional de vectores autorregresivos con constante y 2 retardos en las variables endógenas. Una modificación en ε_i tiene efectos instantáneos en el valor presente de la demanda transmitidos a los valores futuros de la inversión en promoción y la variable inversión en capital. Así, por ejemplo, un shock positivo en la variable ventas de 1.2 puntos

porcentuales generará los siguientes efectos: La perpetuación del efecto en la misma variable que genera un aumento de 0,94 en $t + 1$, y de 1,14 en $t + 2$. Tal shock tenderá a amortiguarse y diluirse desde el tercer periodo, este mismo *shock* ocasiona un incremento de la inversión en marketing en 1,09 puntos para $t + 1$ y 1,01 para $t + 2$.

Igualmente se produce un aumento sobre la inversión en capital de 1,20 en el primer periodo. Tal incremento que se prolonga hasta el cuarto periodo y luego se diluye en la progresión temporal. El perfil de las funciones impulso respuesta (FIR) exhibe un superlativo nivel de estabilidad en el modelo, pues son perceptibles efectos transitorios y no abruptos ante variaciones de las innovaciones de cada una de las variables. Esta situación es esperable en un VAR estable, donde las FIR suelen converger a cero en el corto plazo, lo que equivale a aducir en otros términos que los choques sobre las variables no tienen efectos permanentes sobre otras variables de interés

Gráfico 1. Funciones impulso-respuestas





Fuente: Elaboración propia

Debido a que los shocks estructurales son estandarizados a shock a un punto porcentual, el eje vertical indica la variación de la variable que responde ante un cambio unitario en la variable de impulso, después de t periodos, cuantificada en términos porcentuales. Así, existirá un efecto estadísticamente significativo entre variables cuando los intervalos de la variable i -ésima que responde al impulso, se localicen en el mismo cuadrante del gráfico en acotación temporal específica; en consecuencia, si la variable se ubica sobre el eje las abscisas existirá una relación positiva entre las variables, en caso contrario, si se ubican debajo de este eje.

Tal y como se visibiliza, los efectos de las variables sobre la demanda evidencian una tendencia prácticamente estable en los diez periodos evaluados, debido a que los shocks causados sobre la variable ventas, salvo que existe un shock creciente de la variable ventas sobre la misma en los periodos iniciales y un shock creciente del marketing sobre las ventas en el periodo inicial hasta que converge a un valor próximo a cero. La variable inversión en capital tiene efectos prácticamente inescrutables sobre la variable dependiente.

3.4. PRONÓSTICOS

El propósito ulterior de la especificación del modelo vectorial autorregresivo no es otro sino el de generar predicciones de la demanda en un marco temporal intermedio. Bajo la asunción de que los vectores $[Y_1, Y_2, \dots, Y_n]$ son observados y se tiene conocimiento exhaustivo de las matrices de coeficientes autorregresivos denotada por Φ y la matriz de varianza-covarianza residual Ω se efectúan proyecciones intramuestrales.

La predicción de Y_{t+1} con la información disponible en el periodo t será la esperanza matemática de Y_{t+1} condicional de la totalidad de información disponible en el periodo t . En este sentido, el predictor lineal óptimo minimizará el error cuadrático medio, es decir, minimiza esta función de pérdida en cada componente y está dado por:

$$\hat{Y}_{T+h|T} = \hat{\Pi}_1 \hat{Y}_{T+h-1|T} + \dots + \hat{\Pi}_p \hat{Y}_{T+h-p|T}$$

Por otro lado, el error de pronóstico a h pasos posteriores será:

$$Y_{T+h} - \hat{Y}_{T+h|T} = \sum_{s=0}^{h-1} \Psi_s \varepsilon_{t+h-s} + (Y_{T+h} - \hat{Y}_{T+h|T})$$

En el Cuadro 8 se relacionan algunas medidas de bondad de ajuste predictivo cuyos valores refrendan la superlativa calidad del mismo (Nugroho, Hartati, Subanar y Mustofa, 2014).

Así por ejemplo cuando el índice de desigualdad de Theil es próximo a la nulidad, circunstancia ideal en la cual $\hat{Y}_{T+h} = Y_{T+h}$ se confirma que las proyecciones son “perfectas”. Las predicciones generadas a partir del modelo se reportan en el gráfico 2.

Nótese que las proyecciones generadas a 10 pasos adelante indican un comportamiento poco turbulento entre las tres variables analizadas.

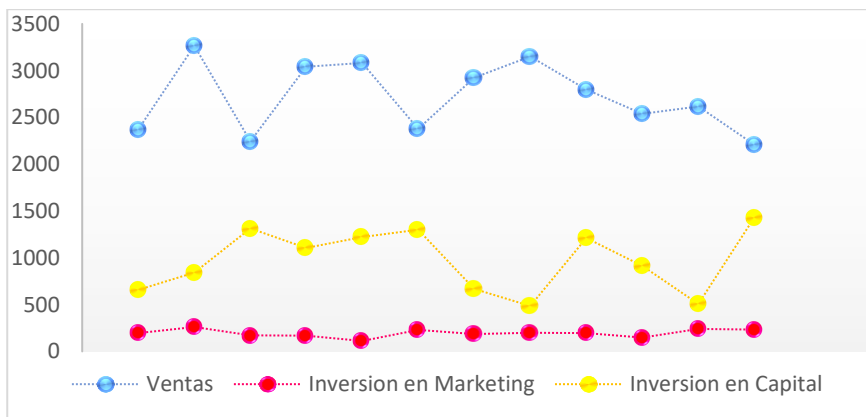
Cuadro 8. Medidas de bondad de ajuste del modelo

MEDIDAS	Error Cuadrado Medio	Error Porcentual Absoluto Medio	U-Theil
FÓRMULA	$\frac{1}{p} \sum_{t=1}^p (y_{T+h} - \hat{y}_{T+h})^2$	$\sum_{t=1}^p \frac{((y_{T+h} - \hat{y}_{T+h})^2)}{(y_{T+h})} \cdot 100\%$	$U = \sqrt{\frac{1/p \sum_{t=1}^p (y_{T+h} - \hat{y}_{T+h})^2}{\sum_{t=1}^p y_{T+h}^2}}$
VALOR	3.71049	11,098%	0.135016

Fuente: Elaboración propia

Una somera inspección visual permite concluir que la inversión en capital experimenta cambios bruscos en los periodos donde ocurren decrecimientos o aumentos en las ventas, no se registran movimientos en la variable inversión en marketing, la cual exhibe una conducta relativamente estable en la progresión temporal. Se detecta además que existen movimientos concomitantes entre las variables inversión en capital y ventas. Nótese que cuando decae el nivel de demanda, la inversión de capital alcanza máximos. En el caso opuesto, es decir, cuando suceden crecimientos abruptos de la demanda, la inversión decrece intempestivamente

Gráfico 2. Pronósticos generados.



Fuente: Elaboración propia

Se percibe, adicionalmente que el valor del error cuadrático medio y el error porcentual absoluto también es relativamente bajo, en contraste con los modelos competentes que se contemplan en el Cuadro 9, motivo por el cual se sustenta resueltamente que el modelo estimado es superior en potencial predictivo.

Cuadro 9. Comparativa con modelos alternativos

MODELOS ALTERNATIVOS DE PRONÓSTICO	Error Cuadrado Medio	Error Porcentual Absoluto Medio
Caminata Aleatoria	4,806	15,492
Caminata Aleatoria Con Drift	4,797	15,151
Tendencia Lineal	4,614	14,449
Tendencia Cuadrática	3,838	14,171
Tendencia Exponencial	3,971	14,149
Promedio Móvil Simple	4,861	14,301
Suavización Exponencial Simple	3,898	14,326
Suavización Exponencial De Brown	4,456	13,834
Suavización Exponencial De Holt	3,252	13,213
Suavización Exponencial De Winter	3,102	13,350
ARIMA(1, 1, 1)X(1, 1, 2)12	4,898	13,605
ARIMA(1, 0, 1)	3,029	13,640
ARIMA(2, 1, 1)	4,993	12,860

Fuente: Elaboración propia

4. CONSIDERACIONES FINALES

El desconcertante ejercicio de pronóstico de ventas es indisputablemente el punto céntrico del proceso de planeación de ventas y operaciones en el contexto de las empresas prestadoras de servicios y manufactureras y por tanto es una labor que no admite posposición. Esta fase es el punto inicial para la vertebración de sistemas decisionales flexibles y la consecuente reproducción de dinámicas direccionadas por las unidades organizacionales con arreglo a ajustar anticipativamente los componentes de la oferta, en

procura de obtener ventajas adaptativas en el azaroso entorno económico donde residen, así como también la consecución de otros fines en el orden táctico y estratégico.

En este *paper* se construyó y validó un modelo de Vectores Autorregresivos para efectos de estimación cuantitativa de las ventas en el contexto del clúster petroquímico plástico cartagenero. El uso de un modelo de esta naturaleza se justifica en tanto que explota adecuadamente la dinámica exhibida por múltiples variables en contraste con otros modelos predictivos elementales. Los resultados provistos indican que las variables selectas, a saber, inversión en capital de los consumidores y precio unitario contribuyen significativamente a la estimación de las ventas mensuales y que entre estas variables existen relaciones de causalidad en el sentido Granger. Adicionalmente, se detecta que los supuestos formales sobre los cuales se asienta la metodología diseñada son verificados íntegramente. Asimismo, se corrobora mediante el cómputo de ciertas medidas de bondad de predicción que el modelo de pronóstico diseñado revela una alta eficiencia predictiva respecto a los otros que se traen a colación.

Se insta a diseñar casos de estudio en el que se realicen análisis comparativos donde se desarrollen otras técnicas de proyección distintas a las ya conocidas (la regresión lineal, el suavizamiento, e inclusive los modelos autorregresivos univariados) que integren múltiples variables. De igual modo, se exhorta a diseñar casos empíricos en los cuales se refrende la validez de otros enfoques de modelado econométrico en el contexto microempresarial, donde existe una enigmática subvaloración del abanico de técnicas circunscritas en este campo, y evitar que queden relegadas a tal punto que hallen fuera de la usanza de la planificación de las ventas y operaciones en sectores productivos que se hallan en pleno crecimiento.

5. BIBLIOGRAFÍA

- Barreras Serrano, A., Sánchez López, E., Figueroa Saavedra, F., Olivas Valdez, J. y Pérez Linares, C. (2014). Uso de un modelo univariado de series de tiempo para la predicción, en el corto plazo, del comportamiento de la producción de carne de bovino en Baja California, México. *Veterinaria México*, 45, 1-9.
- Bose, E., Hravnak, M. y Sereika, S. (2017). Vector Autoregressive Models and Granger Causality in Time Series Analysis in Nursing Research: Dynamic Changes Among Vital Signs Prior to Cardiorespiratory Instability Events as an Example. *Nursing Research*, 66(1), 12–19. doi:10.1097/NNR.000000000000193
- Córdova Olivera, P. (2014). Sistema de pensiones y profundidad financiera: evidencia empírica deointegración para el caso boliviano. *Investigación & Desarrollo*, 1(14), 22-43.

- Delgadillo-Ruiz, O., Ramírez-Moreno, P., Leos-Rodríguez, J., Salas González, J. y Valdez-Cepeda, R. (2016). Pronósticos y series de tiempo de rendimientos de granos básicos en México. *Acta universitaria*, 26(3), 23-32. doi:10.15174/au.2016.882.
- Hamilton, J. (1994). *Time Series Analysis*. New Jersey: Princeton University Press.
- Hani, O., Duen-Ren, L. y Hai Hoang, V. (2016). A Hybrid Neural Network Model for Sales Forecasting Based on ARIMA and Search Popularity of Article Titles. *Computational Intelligence and Neuroscience*, 2016, 9. doi:10.1155/2016/9656453
- Li, T., Li, X. y Zhang, X. (2017). The Design and Implementation of Vector Autoregressive Model and Structural Vector Autoregressive Model Based on Spark. *Big Data Computing and Communications (BIGCOM), 2017 3rd International Conference on*. Chengdu: IEEE. doi:10.1109/BIGCOM.2017.46
- Loria, E. (2014). *Econometría con aplicaciones*. México D.F: Pearson Education.
- Morrone, H. (2015). ¿La demanda y la rentabilidad estimulan la acumulación de capital? Un análisis para el Brasil. *Revista CEPAL*, 163-175.
- Nicholson, W., Bien, J. y Matteso, D. (2016). Hierarchical Vector Autoregression. *Cornell University*, 1-26.
- Nugroho, A., Hartati, S., Subanar y Mustofa, K. (2014). Vector Autoregression (Var) Model for Rainfall Forecast and Isohyet Mapping in Semarang - Central Java-Indonesia. *International Journal of Advanced Computer Science and Applications*, 5(11), 44-49.
- Sánchez López Cruz , F. y Marcelo, J. (2016). Determinantes económicos de los flujos de viajeros a México. *Revista de análisis económico*, 31(2), 3-36. doi:10.4067/S0718-88702016000200001.
- Shahin, M., Ali, M. y Ali, A. (2014). Vector Autoregression (VAR) Modeling. En T. Islam, P. Srivastava, M. Gupta, X. Zhu, & S. Mukherjee, *Computational Intelligence Techniques in Earth and Environmental Sciences* (pág. 266). New York: Springer.
- Suharsono, A., Aziza, A. y Pramesti, W. (2017). Comparison of vector autoregressive (VAR) and vector error correction models (VECM) for index of ASEAN stock price. *ICWOMAA 2017, 1913*, págs. 1-9. doi:10.1063/1.5016666
- Trujillo Calagua, G. (2015). La Metodología del Vector Autorregresivo: Presentación y Algunas Aplicaciones. *UCV-SCIENLIA*, 2(2), 103-108.
- Zhang, X., Zou, G. y Carroll, R. (2015). Model averaging based on Kullback-Leibler. *Statistica Sinica*, 25, 1583-1598. doi:10.5705/ss.2013.326

Zhang, Y., Miner, Z., Geng, N. y Jiang, Y. (2017). Forecasting electric vehicles sales with univariate and multivariate time series models: The case of China. *PLOS ONE*. doi:10.1371/journal.pone.0176729