



# EL COLEGIO DE MÉXICO

## CENTRO DE ESTUDIOS ECONÓMICOS

### **MAESTRÍA EN ECONOMÍA**

TRABAJO DE INVESTIGACIÓN PARA OBTENER EL GRADO DE  
MAESTRO EN ECONOMÍA

**PREDICCIÓN DE LA DEMANDA  
DE GASOLINA MEDIANTE  
ANÁLISIS DE SERIES DE TIEMPO**

**SAMUEL HERNÁNDEZ STANFORD**

PROMOCIÓN 2016-2018

ASESOR:

DR. ENEAS A. CALDIÑO GARCÍA

JUNIO 2018

# Agradecimientos

Aprovecho estas líneas para mostrar mi más profundo agradecimiento a todas aquellas personas que me brindaron su apoyo en la realización de este trabajo y a lo largo de estos dos años, que han enriquecido notablemente mi formación profesional y personal.

Un especial reconocimiento merece el cariño, aliento y paciencia de la valiente mujer que decidió compartir la vida conmigo; Elizabeth Ruiz, este logro sin duda es de ambos.

A mis padres y mis hermanos les recuerdo que les debo todo lo que soy, este trabajo no tendría por qué ser la excepción. Muchas gracias por su comprensión e interés constantes.

Una mención especial amerita el doctor Sergio Galina, quien, por alguna razón que aún no me queda del todo clara, puso todo su empeño desinteresado en ayudarme a cumplir esta meta y no ha dejado de tener para conmigo más que atenciones y valiosos consejos. Unos de tantos me llevaron a las oficinas de Edze Kornelis y Gabriel Aguilar, quienes se tomaron el tiempo de proveerme de ideas e insumos indispensables para mi investigación. A ellos, también, aprovecho para hacerlos partícipes de mi gratitud.

También quiero dar las gracias a los profesores de El Colegio de México y en especial a mi asesor y coordinador del Centro de Estudios Económicos, el doctor Eneas Caldiño, cuya experiencia y profesionalismo fueron claves para el desarrollo de esta tesis.

Finalmente, agradezco a mis compañeros, Ivette Agúndez, Marco Gómez y Jorge Sosa, por sus atinadas aportaciones y observaciones, pero sobre todo, por su amistad.



# Resumen

Esta tesis bosqueja los primeros efectos de la implementación de la Reforma Energética en el mercado mexicano de petrolíferos. Mientras las típicas fuerzas de mercado comienzan a buscar su equilibrio, la apremiante necesidad de fortalecer la infraestructura de logística y almacenamiento de combustibles brinda a los jugadores privados atractivas oportunidades de inversión, en un entorno de precios del crudo todavía deprimido. Este trabajo analiza las ventas, de enero de 2002 a octubre de 2017, de las Terminales de Almacenamiento y Reparto usando metodologías basadas en series de tiempo, para desarrollar modelos predictivos de demanda usando la entidad federativa como agrupador e incorporando variables exógenas, toda vez que éstas aporten poder explicativo al modelo.



# Índice general

Introducción	1
1. Reforma Energética	3
2. Demanda de combustibles en México	6
3. Revisión de literatura	14
4. Motivación del proyecto	18
5. Datos	19
6. Modelo para la demanda nacional	21
7. Modelos para la predicción de la demanda a las TAR	28
8. Conclusiones	39
Anexos	41
A. Coeficientes estimados para cada estado y medidas de la precisión en la estimación	43
B. Función impulso respuesta (IRF) para el estado de Nuevo León	47
Bibliografía	49
Índice de figuras y tablas	53



# Introducción

En este documento se desarrollan modelos para la predicción de la demanda mensual de gasolinas mediante metodologías basadas en series de tiempo. Las unidades de análisis de interés para el pronóstico de la demanda son las 78 Terminales de Almacenamiento y Reparto propiedad de Petróleos Mexicanos. La exactitud de los valores pronosticados es evaluada con base en el error cuadrado medio calculado a partir de una porción de los datos históricos separada para tal fin.

Un primer análisis a la base de datos nos permitió percatarnos de que las terminales únicamente abastecen a la entidad federativa en la que están localizadas<sup>1</sup>. En razón de lo anterior, el agrupador natural de las terminales es la entidad, y por tal motivo, la presente investigación contempla 31 modelos<sup>2</sup> de regresión LS dinámicos, 23 de ellos multivariados y los 8 restantes univariados, en procesos estacionales autorregresivos.

El horizonte temporal que cubre la base de datos incluye el periodo en que entró en vigor la Reforma Energética, misma que representa un cambio de paradigma en los procesos logísticos y comerciales de los petrolíferos. Evidentemente, este contexto no es el más oportuno para la calibración de los modelos predictivos; no obstante, los pronósticos estimados son razonablemente aproximados.

La apertura del mercado de combustibles desencadenará un nuevo equilibrio en el que las fuerzas de mercado adoptarán un papel preponderante, lo cual resulta inédito para el caso mexicano.

El presente documento está organizado en siete capítulos: Los primeros dos contextualizan al lector con respecto al mercado mexicano de petrolíferos; en el tercero se hace una breve revisión de la vasta literatura relacionada con los modelos de demanda de combustibles; el cuarto resalta la importancia de desarrollar un modelo predictivo para México; en el quinto se describen los datos usados para el análisis; el sexto y séptimo detallan la metodología usada, los modelos obtenidos y los resultados de éstos; finalmente, en el octavo capítulo se

---

<sup>1</sup>Con excepción de los estados de Tlaxcala y Quintana Roo, quienes no cuentan con TAR y son provistas por las terminales ubicadas en Puebla y Yucatán respectivamente.

<sup>2</sup>Las variables exógenas consideradas en los modelos de Puebla y Yucatán, engloban la información de esos estados y la del estado a la que suministran combustible.

resumen las conclusiones.

# Capítulo 1

## Reforma Energética

La entrada en vigor de la Reforma Energética en 2013 ha propiciado una transformación del modelo de mercado energético hacia la competencia y apertura, lo cual ha modificado estructuralmente la forma de hacer las cosas a lo largo de la cadena productiva de los petrolíferos. La nueva ley posibilita la incorporación de empresas privadas en actividades que anteriormente eran exclusivas del Estado, desde la exploración y extracción de crudo, hasta la refinación, almacenamiento, transporte, comercialización, importación y exportación de hidrocarburos.

La reforma estructural del sector, específicamente el artículo 70 de la Ley de Hidrocarburos, faculta a Petróleos Mexicanos, a través de su subsidiaria Pemex Logística, a celebrar procedimientos denominados como “Temporadas Abiertas”. Este tipo de iniciativas permite a terceros interesados la contratación de capacidad disponible en la infraestructura de almacenamiento y transporte por ductos de petrolíferos de Pemex Logística por un plazo determinado.

Estas modificaciones al marco regulatorio persiguen el objetivo de revertir el rezago significativo que enfrenta la infraestructura logística de petrolíferos en México, resultado de la subinversión que históricamente ha caracterizado al sector. En promedio, México cuenta con un inventario de combustibles equivalente a tres días de abastecimiento, lo que representa 10 % de la capacidad recomendada internacionalmente, mientras existen países como Estados Unidos que cuentan con 45 días de inventario. PEMEX cuenta con 77 Terminales de Almacenamiento y Reparto (TAR) en operación, de las cuales tres tienen una capacidad menor a dos días de almacenamiento y 43 % de las restantes no alcanza los cinco días de autoabastecimiento.

Las proyecciones de la Comisión Reguladora de Energía (CRE) indican que la capacidad de almacenamiento disponible se quintuplicará hacia 2030; buena parte de este crecimiento en infraestructura vendrá de la liberación de capacidad por 10.3 millones de barriles que hará la Comisión Federal de Electricidad (CFE) en ese mismo lapso. El planteamiento de este

proyecto considera que la CFE conserve la propiedad de las instalaciones, pero financie la conversión y adecuación de las mismas mediante la celebración de alianzas. La ampliación restante en los niveles de inventario requerirá de una importante inyección de capital privado, motivo por el cual, los cambios a la regulación debieron ser acompañados por la liberalización de los precios de las gasolinas y diésel.

El modelo de precios vigente hasta 2016 provocó que PEMEX perdiera recursos al no reconocer el costo logístico dentro de la estructura de precios finales al consumidor, la cual tampoco consideraba las diferencias regionales, y cuya rigidez provocaba que las variaciones en los precios internacionales se reflejaran con un considerable rezago, además de incluir un oneroso subsidio de 200 mil millones de pesos anuales (Apertura gradual y ordenada del mercado de gasolinas y diésel en México, 2017, SENER).

En cuanto al transporte de petrolíferos la situación no es menos apremiante. La mayor parte de la red de poliductos está saturada, es decir, presenta un porcentaje de utilización mayor al 90% de su capacidad nominal. En consecuencia, para la alimentación de las TAR desde las refinerías y los puntos de importación ha sido indispensable hacer uso de medios de transporte menos eficientes.

Tabla 1.1: Participación y costo de los distintos medios de transporte a TAR en 2016

Medio de transporte	Participación en el transporte a TAR (%)	Costo (centavos por kilómetro)
Ducto	67	1
Buque-tanque	12	2
Carro-tanque	5	6
Auto-tanque	16	14

El último eslabón de la cadena de suministro de combustibles hacia el usuario final no está exento de carencias. 40% de los municipios no cuentan con gasolineras debido a que la cantidad de habitantes no genera los estímulos necesarios para la inversión privada. Subsiguientemente, México cuenta con un menor número de estaciones de servicio en relación con otros países. Mientras Estados Unidos y Brasil promedian 2,677 y 5,461 habitantes por gasolinera respectivamente, la entidad mexicana con el mejor cociente es Tamaulipas con 6,421 y la media nacional es de 10,560 habitantes por estación. Esta situación ya se ha comenzado a subsanar. Entre 2014 y 2017 se registró un incremento de 7% en el número de gasolineras en el país, pasando de 10,830 a 11,586, y se espera que el ritmo de crecimiento se acelere en los próximos años.

Los efectos de la apertura del mercado se han comenzado a sentir en la parte final de la cadena de transformación de petróleo. Encontramos que al cierre del tercer trimestre de 2017 se contabilizan 652 permisos vigentes para la importación de petrolíferos, 12% de la

importación de diésel es efectuada por jugadores privados; se ha anunciado la aprobación de 38 nuevos proyectos para la ampliación de la infraestructura de almacenamiento, 6 enfocados al transporte por ducto; y se han posicionado 26 nuevas marcas para el expendio de combustibles al usuario final (Diagnóstico de la Industria de los Petrolíferos 2017).

## Capítulo 2

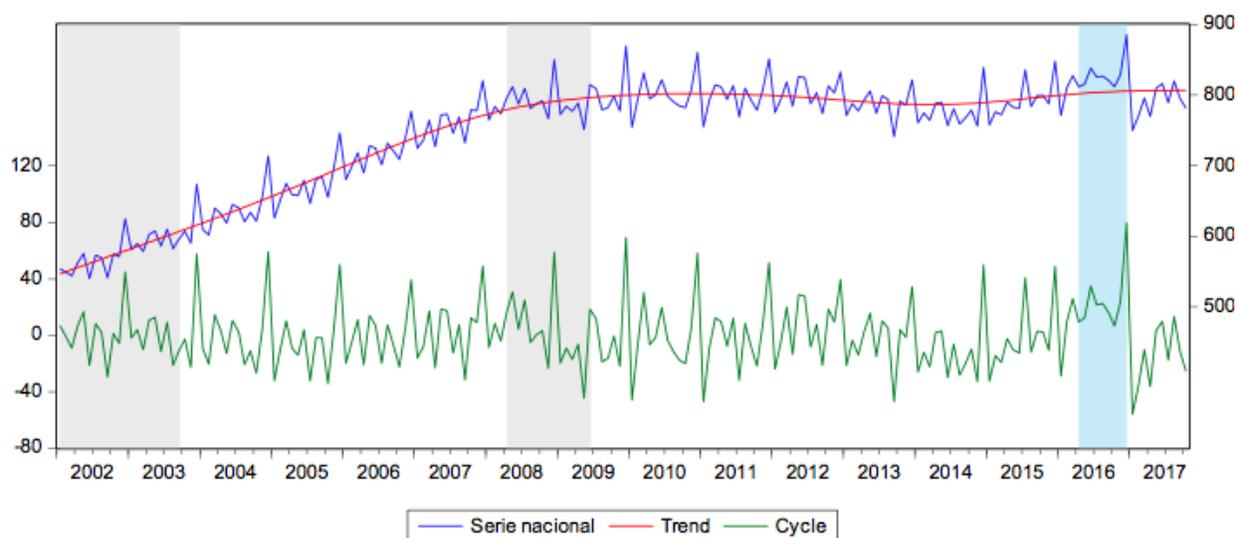
# Demanda de combustibles en México

Históricamente, la demanda nacional de combustibles ha presentado una importante correlación con el desarrollo económico del país, principalmente el de los sectores de transporte e industrial. Sin embargo, en los últimos años la tendencia en el comportamiento de la demanda de gasolinas ha cambiado, en buena medida por los avances tecnológicos encaminados hacia una mayor eficiencia energética y a reducir el impacto de las actividades productivas en el medio ambiente y la saturación de algunos mercados específicos.

Prueba de lo anterior es que entre 2003 y 2008 la demanda de gasolinas crecía a una tasa promedio anual que oscilaba alrededor del 6 %, mientras que en el periodo 2009-2016 el proceso parece haberse estabilizado, mostrando una tasa media anual apenas positiva (0.16 %). El siguiente gráfico muestra el proceso que describe la demanda nacional de gasolinas (parte superior en azul) y su descomposición Hodrick-Prescott en tendencia (parte superior en rojo) y ciclo (parte inferior). Las áreas sombreadas en gris indican el periodo entre un pico y un valle en el indicador coincidente del Sistema de Indicadores Cíclicos del Instituto Nacional de Estadística y Geografía (INEGI). El área azul señala el intervalo de implementación de la estrategia de flexibilización del mercado de combustibles, el cual inició el 30 de marzo de 2016 en Baja California y Sonora, en donde se ubican 1,097 estaciones de servicio. El 15 de junio se continuó con la segunda etapa con 2,225 estaciones ubicadas en los estados de Chihuahua, Coahuila, Nuevo León, Tamaulipas y el municipio de Gómez Palacio en Durango; se extendió en su tercera etapa el 30 de octubre en Baja California Sur, Sinaloa y el resto de Durango donde se localizan 802 gasolineras; y concluyó el 30 de noviembre con los estados del centro y sur del país, mismos que contabilizan 7,650 gasolineras.

Es indiscutible el hecho de que el ciclo evidencia una fuerte estacionalidad en el proceso, con grandes picos en diciembre de cada año y fuertes caídas hacia enero del siguiente y cuya varianza aumenta ligeramente al paso de los años.

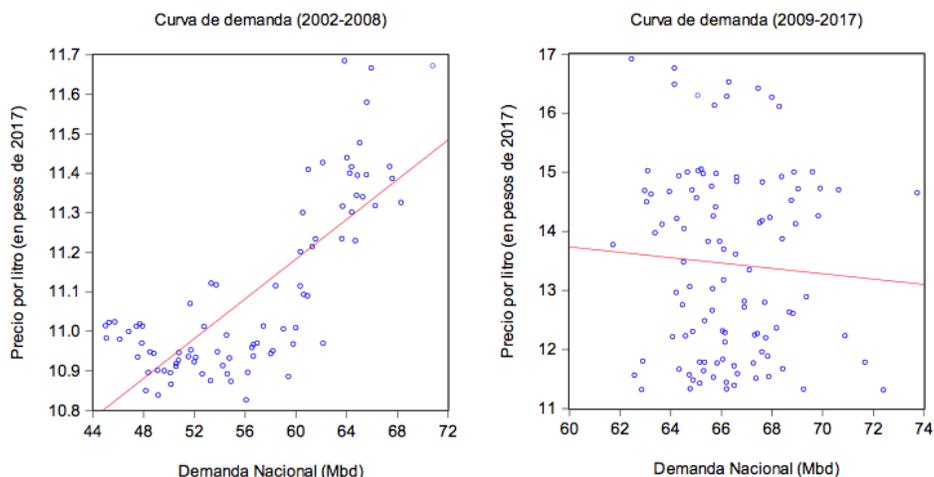
Figura 2.1: Demanda nacional de gasolinas HP (Mbd)



Dentro del horizonte temporal analizado existe un cambio estructural en el mercado mexicano de combustibles, teniendo como causas más probables la entrada en vigor de la Reforma Energética y específicamente la liberación de los precios de las gasolinas. Como evidencia de lo anterior tenemos que la pendiente de la curva de demanda nacional de gasolinas cambia de signo dependiendo del subconjunto de datos que usemos para construirla.

Si graficamos la demanda nacional contra el precio real en pesos a enero de 2017 por litro con los datos de 2002 a 2008, se obtiene una curva de demanda con pendiente positiva, lo que corrobora una de las anomalías descritas por Fullerton, Ibarra y Elizalde (2015). Al construir la curva con la información comprendida entre 2009 y 2017, la pendiente se vuelve negativa y crece en valor absoluto al restringir más la muestra hacia los datos más actuales.

Figura 2.2: Curvas de demanda (2002-2017)



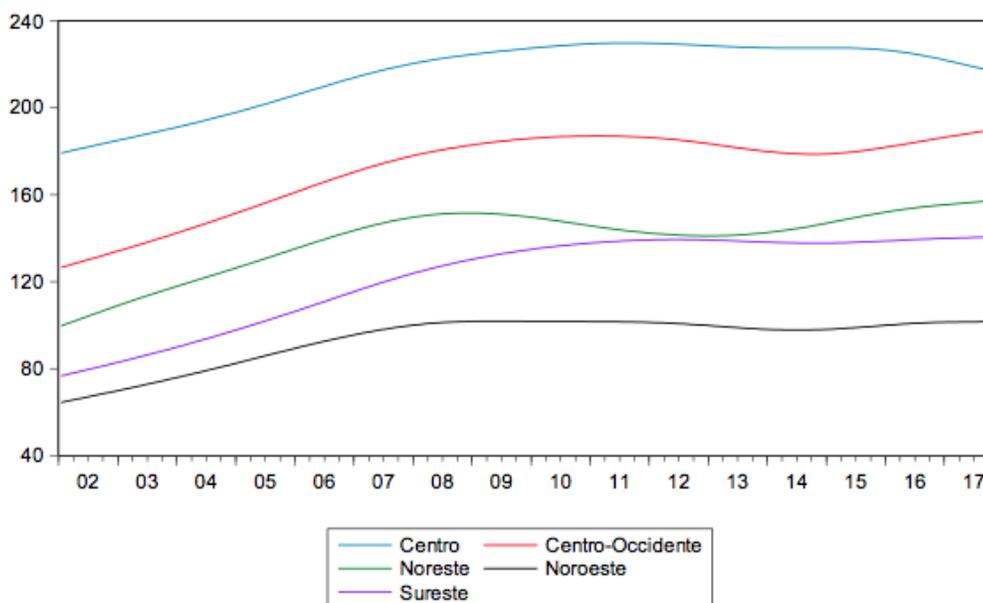
Una curva de demanda con pendiente positiva nos sugiere que con respecto al consumo de gasolina, el efecto ingreso sobrepasa al efecto sustitución, lo cual tiene sentido si observamos que el precio de las gasolinas aumentó apenas 11.3% entre 2002 y 2012, mientras que entre 2012 y 2017 asciende en 32.7%.

Haciendo el mismo análisis a nivel regional se replica el fenómeno. No obstante, el cambio de signo en la pendiente de sus curvas de demanda requiere de distintas filtraciones; es decir, el dominio de las curvas que tienen pendiente igual a cero es diferente para cada región.

El consumo regional de gasolinas describe un comportamiento heterogéneo en el intervalo analizado. La región centro presenta un perfil de demanda muy similar al nacional pero con una visible caída hacia los últimos dos años, situación opuesta a la región centro-occidente, la cual cierra con una pendiente positiva.

La demanda en la región noreste del país refleja una posible influencia de la crisis financiera de 2008, a partir de la cual comienza un periodo de decrecimiento que se detiene hasta 2012, en donde exhibe un mínimo local que casi iguala la demanda de la región sureste, recuperando sus niveles precrisis hasta inicios de 2016.

Figura 2.3: Tendencia en la demanda regional de gasolinas (Mbd)

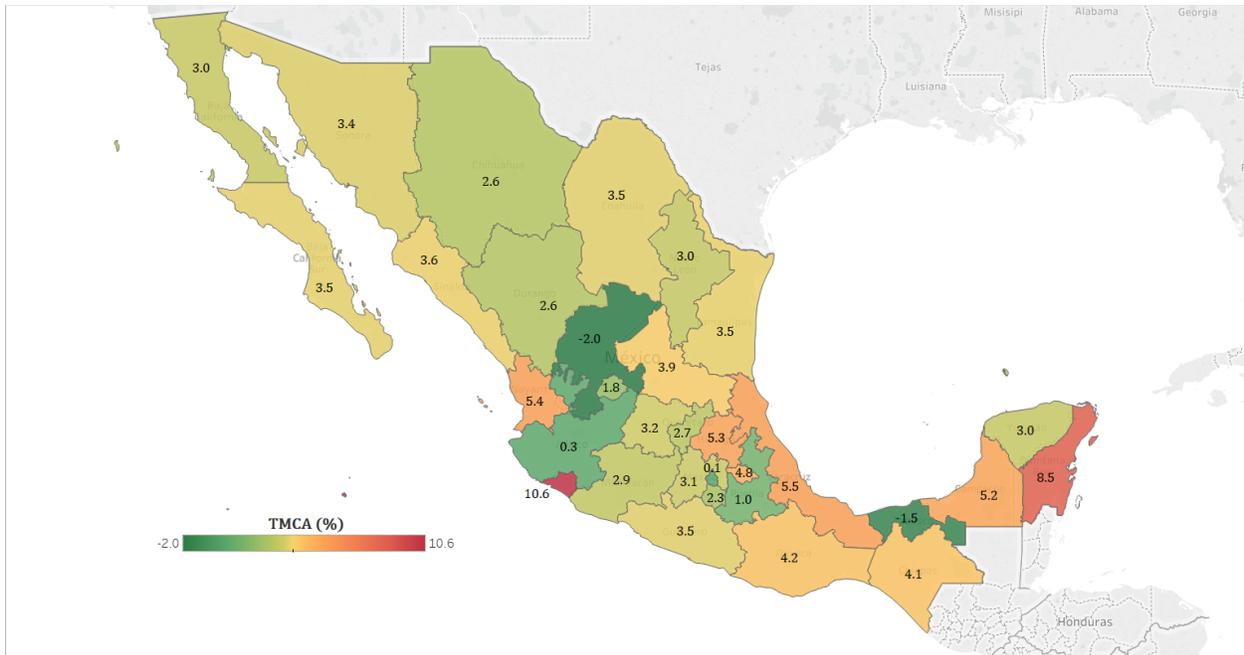


A nivel estatal también observamos un comportamiento muy heterogéneo en la demanda. Mientras algunas entidades como Colima y Quintana Roo presentan tasas de crecimiento anuales promedio de 10.6% y 8.5% respectivamente, otras como Zacatecas y Tabasco presentan decrecimientos de 2.0% y 1.5% respectivamente. Los estados que cuentan con las ciudades más grandes del país muestran crecimientos promedio cercanos al cero, como lo son Ciudad de México, Jalisco y Puebla. La demanda de Nuevo León crece en promedio

3% a lo largo del periodo analizado, lo cual puede deberse a la recuperación de la crisis, misma que afectó a los estados fronterizos en mayor proporción debido a la predominante participación del sector manufacturero en su composición sectorial (López y Peláez, 2015). En general podemos visualizar una tendencia a la convergencia debido a la saturación del mercado automotriz en las principales ciudades y una creciente participación de motores más eficientes o incluso eléctricos.

Ahondando en la saturación del mercado automotriz, se visualiza una gran heterogeneidad a nivel estatal. La Ciudad de México alcanza una relación de 621 autos registrados por cada mil habitantes hacia el final del periodo analizado, niveles comparables con países como Japón, Noruega y Francia, mientras que estados como Chiapas y Oaxaca registran apenas 143 y 124 autos por cada mil habitantes respectivamente, niveles semejantes a los de países como Zimbabwe, Ecuador y El Salvador. A nivel nacional, al tercer trimestre de 2017, en México existen 331 autos por cada millar de habitantes.

Figura 2.4: Crecimiento medio anual del consumo de gasolinas 2002-2016 (%)



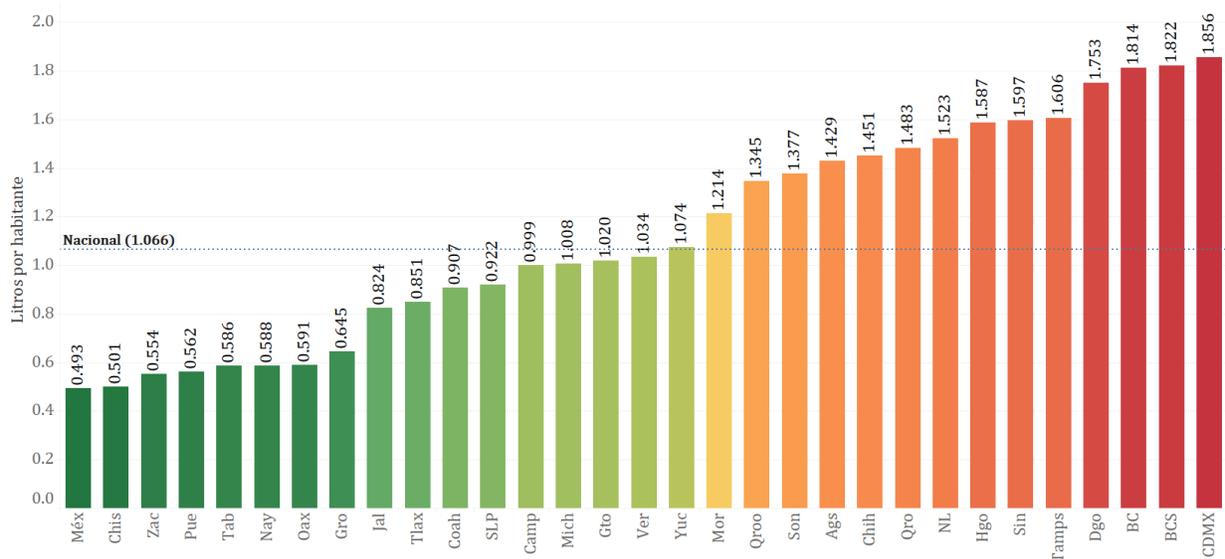
Cabe mencionar que las estadísticas anteriores están basadas en autos registrados en cada entidad federativa, por lo que es posible que el índice calculado para la Ciudad de México corresponda a una cota inferior ya que debido a la implementación de ciertas regulaciones fiscales y de tránsito, se han generado incentivos para que los residentes de la capital tramiten placas para sus automóviles de estados aledaños como Morelos o el Estado de México.

Con respecto al consumo de gasolina per cápita, sorprende el caso de Colima, el cual alcanzó un pico de 6.4 litros de gasolina diarios en diciembre de 2016. Contextualizando este nivel

de consumo, Estados Unidos, el mayor consumidor per cápita del mundo, promedió 4.4 litros diarios por habitante en 2015. Esta elevada demanda de combustible claramente está relacionada con el hecho de que el puerto de Manzanillo ocupa el primer lugar a nivel nacional en volumen de comercio internacional, movilizandando más de 2.5 billones de toneladas de mercancías en 2016.

El resto de las entidades que integran el último quintil de consumo per cápita son: Durango, la Ciudad de México, Baja California Sur, Baja California, Tamaulipas y Aguascalientes, las cuales superaron los 1.46 litros diarios por habitante en octubre de 2017. El primer 20 % de la distribución se caracteriza por un consumo inferior a los 0.63 litros diarios, en donde se ubican los estados de Oaxaca, Guerrero, Puebla, Tabasco, Estado de México, Nayarit y Chiapas.

Figura 2.5: Consumo per cápita de gasolinas en 2016



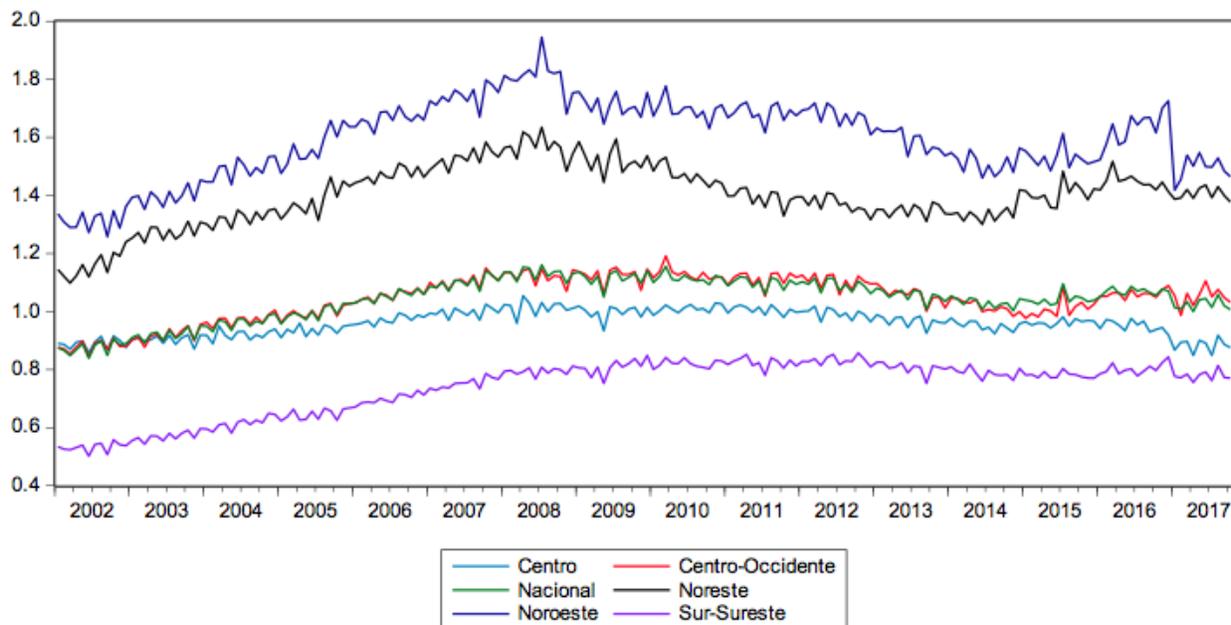
Nota: Por motivos de visualización, el gráfico excluye al estado de Colima.

A nivel regional se observa que el mayor consumo per cápita se concentra en los estados del norte del país, con una media de 1.60 litros para la región noroeste y 1.40 para la región noreste, seguidas por centro-occidente con 1.04, la cual prácticamente emula el consumo nacional, centro con 0.96 y finalmente sur-sureste con 0.74 litros por habitante a lo largo del periodo de análisis.

El siguiente gráfico presenta las series desestacionalizadas para el consumo de gasolina per cápita en las cinco regiones y a nivel nacional. Estas series en su conjunto arrojan estadísticos que permiten rechazar la hipótesis nula de raíz unitaria común e individual a un nivel de significancia de 1 %, por lo que pueden considerarse como procesos estacionarios. Haciendo

el mismo análisis con las series originales tenemos que a un nivel de significancia del 10 % podemos rechazar la hipótesis nula para la prueba de Dickey-Fuller aumentada y a 1 % para la de PP-Fisher, por lo que podemos inferir que los procesos tienen una media y varianza constantes.

Figura 2.6: Consumo per cápita de gasolinas (litros)



Nota: Series ajustadas por estacionalidad (Census X-12)

Resulta relevante señalar el marcado declive en el consumo sufrido en enero de 2017, en donde las series con ajuste estacional para las regiones noroeste y sur-sureste muestran sus caídas más importantes dentro del intervalo que abarca de enero de 2002 a octubre 2017, las cuales alcanzaron 17.8 % y 8.0 % respectivamente en relación al mes previo. Lo anterior coincide con el mayor shock en el precio del combustible a nivel nacional, el cual se incrementó en 14.2 % en el noroeste y en 11.6 % en el sur-sureste en términos reales.

Con respecto al impacto de las variables explicativas más recurrentes en la literatura relacionada con el consumo de combustibles, se puede concluir que para el caso mexicano el efecto de las mismas en la demanda a nivel entidad federativa es muy heterogéneo. Como muestra de la anterior aseveración se presenta el siguiente gráfico en donde se visualiza el crecimiento en el PIB per cápita y el crecimiento en el número de vehículos por habitante, además de la participación de cada estado en el consumo nacional de gasolinas en el periodo comprendido entre 2003 y 2016.

Es importante mencionar que el gráfico no incluye el incremento en los precios porque ese cambio sí fue muy semejante para todos los estados, con un rango que varía entre el 36 %

de aumento para Colima y 41 % para Hidalgo en términos reales dentro de los 14 años que abarca el horizonte temporal del gráfico.

Figura 2.7: Evolución 2003-2016



Nota: Por razones de visualización, la dispersión no incluye al estado de Campeche que presenta un desplome de 53 % en su producto per cápita, ni a Colima, que casi triplica su consumo, pasando de 1.6 a 4.7 litros de gasolina por habitante.

Para ilustrar el párrafo anterior podemos centrar nuestra atención en los estados de Tlaxcala y la Ciudad de México. Aunque pertenecen a la misma región, el primero presenta un decrecimiento en su PIB per cápita de 4 %, mientras que para el segundo este dato fue positivo y asciende a 40 %. Sin embargo, Tlaxcala presenta un incremento de 56 % en su demanda por habitante, mientras que la capital del país únicamente consumió un 7 % más, lo cual es aún más peculiar si nos fijamos en el hecho de que el número de vehículos por individuo subió 155 %. Esta intensificación en la posesión de vehículos está en línea con el aumento en el ingreso, mismo que puede asociarse con una rápida modernización del parque vehicular y, por lo tanto, un repunte sostenido en el rendimiento de los motores para finalmente explicar el crecimiento marginal del consumo. Este mecanismo puede ser común a mercados automotrices con un nivel de saturación considerable.

Otro contraste importante lo muestran Tabasco y Guerrero, cuyas economías crecieron de forma muy similar, sin embargo, el primero exhibe una importante caída en su demanda aún con un aumento en el parque per cápita de 118 %. Este comportamiento anómalo en el proceso de demanda de Tabasco se concentra en 2014, año en el que promedia una tasa mensual de crecimiento de -3.7 %.

Otras entidades con decrecimiento en el consumo son: Zacatecas, Jalisco, Puebla y Baja California Sur, entre los que se destaca el primero como el cuarto estado con mayor crecimiento económico entre 2003 y 2016, alcanzando un 38 %, aunque con un ascenso moderado en su parque per cápita de 9 %. No obstante, Jalisco ostenta un incremento en el número de vehículos por habitante de 59 %, una elevación de su producto de 23 % y un descenso en su demanda de 15 %.

Quintana Roo y el Estado de México lideran la estadística de aumento de vehículos por individuo con 278 % y 286 % respectivamente. Empero, Quintana Roo se posiciona como la segunda entidad con mayor aumento en su consumo de gasolinas, mientras que el Estado de México se ubica por debajo del promedio nacional con apenas 9 %.

A partir de este breve análisis descriptivo se advierte la conveniencia de establecer como unidad de análisis la entidad federativa, examinando el poder explicativo de las variables exógenas que se han mencionado, pero enfocándose principalmente en la dinámica de las series de tiempo.

# Capítulo 3

## Revisión de literatura

La investigación de la demanda de energía capturó el interés de los hacedores de políticas públicas a partir de 1973, año en que tuvo lugar la primera crisis petrolera, cuando los países que integraban la OPEP decidieron suspender las exportaciones de petróleo hacia los países que habían apoyado a Israel durante la guerra de Yom Kipur, la cual lo enfrentaba con Siria y Egipto. Esta medida afectó tanto a Estados Unidos como a sus aliados europeos, quienes vivieron un importante efecto inflacionario provocado por un aumento en el precio del barril de crudo de aproximadamente 367 % en términos reales, hacia finales de 1974. En respuesta a este shock, Estados Unidos instrumentó algunas políticas orientadas a la disminución de la demanda de energía como el establecimiento de límites de velocidad, la imposición del horario de verano y la adopción de ciertas directrices para reducir el tamaño y peso de los automóviles nuevos (Blinder, 1979).

El objetivo de enriquecer el entendimiento de las fuerzas detrás de la demanda energética y su respuesta ante shocks exógenos, dio origen a un intenso debate entre ingenieros y economistas que redundó en importantes desarrollos metodológicos que pusieron a disposición una gran variedad de modelos para analizar y predecir la demanda. Surgió una nueva tendencia de publicaciones relacionadas con la estimación de las elasticidades precio-demanda con una perspectiva de corto y largo plazos para diferentes países mediante la aplicación de diversos modelos, datos y técnicas (Wirl y Szirucsek, 1990).

Hacia principios de este siglo, el enfoque antes descrito evolucionó hacia la identificación y jerarquización de los principales determinantes del consumo de combustibles fósiles como el ingreso (Graham y Glaister, 2002), PIB per cápita, volumen de pasajeros en transporte público, parque vehicular (Chai, Wang, Wang y Guo 2012), crecimiento poblacional, densidad demográfica (Karathodorou y Graham 2010) y, más recientemente, algunos factores tecnológicos como las mejoras en la eficiencia energética, el uso de combustibles alternos como el gas natural y los biocombustibles, la creciente participación de los motores eléctricos en el mercado automotriz, y la instrumentación de regulaciones ambientales más exigentes

(IEA 2016), así como la conducta de los consumidores.

Las variables explicativas de la variación en la demanda de combustibles que han sido investigadas en las últimas cuatro décadas tienen diferente impacto en las distintas economías en donde se realice el análisis. Factores como la idiosincrasia y preferencias de los consumidores, madurez de la economía, saturación de los mercados domésticos, infraestructura de transporte y de refinación, y logística, entre otros, provocan una sensibilización del consumo de gasolinas del sector transporte hacia uno u otro componente. Varios de estos factores suelen englobarse para dar un tratamiento específico a economías en desarrollo y desarrolladas.

Las principales características que diferencian a los países en vías de desarrollo incluyen: pobre desempeño del sector energético, lenta transición de energías convencionales a modernas, deficiencias estructurales en la economía, sociedad y sistemas energéticos que resultan en una marcada división entre localidades urbanas y rurales, decisiones de inversión inadecuadas y subsidios incorrectamente dirigidos (Urban, Benders y Moll 2007).

En los últimos 10 años se ha publicado vasta literatura relacionada con la predicción de la demanda de combustibles tanto a corto como a largo plazo. En los siguientes párrafos se bosquejarán algunas de las investigaciones más relevantes y sus principales resultados.

Bhattacharyya y Timilsina (2009) evalúan modelos para pronosticar la demanda de energía en países en desarrollo, encontrando que los modelos econométricos generalmente arrojan estimaciones más aproximadas a la realidad en comparación con modelos teóricos de uso final. Sin embargo también suelen ser más intensivos en microdatos y por lo tanto son más sensibles ante información de baja calidad.

Li, Rose y Hensher (2009) hacen una revisión más profunda de varios modelos empíricos con distintos grados de complejidad, separando una fracción de su muestra para valorar la precisión en el pronóstico, encontrando que el modelo de tendencia cuadrática muestra el mejor desempeño a partir de datos anuales de la economía australiana. No obstante, refuerzan la idea de que no hay un modelo ideal para cualquier aplicación, sino que es indispensable no perder de vista cuáles son los objetivos del estudio y las características de las series. Finalmente, cierran sus recomendaciones afirmando que si el objetivo primario de la modelación es la predicción, los métodos basados en series de tiempo, como ARIMA (autorregresive integrated and moving average), son los más adecuados.

Suganthi y Samuel (2012) y Hong y Wei-Chiang (2013) revisitan varios de los modelos antes mencionados, incluyendo algunas técnicas como lógica difusa, algoritmos genéticos y redes neuronales, las cuales han sido aplicadas recientemente para la proyección de la demanda energética incorporando variables macroeconómicas y han alcanzado, junto con ARIMA, pronósticos que pueden ser usados como benchmark.

Derivado de la gran variabilidad entre las estimaciones empíricas de la elasticidad precio

en el consumo de gasolinas, algunos otros autores como Espey (1998), Hanly et al. (2002), Graham y Glaister (2002a), Brons et al. (2008), Havranek et al. (2012), Brons et al. (2006), entre otros, emplearon un modelo de meta-regresiones en donde la elasticidad es calculada como la variable dependiente sobre una serie de variables moderadoras. Esta aproximación conocida como meta-análisis se basa en el supuesto de que las estimaciones obtenidas bajo distintos métodos pueden ser combinadas haciendo uso de las relaciones lineales de las elasticidades puntuales. Esta metodología, además, permite descomponer la elasticidad precio de la demanda en estimaciones de elasticidad precio del rendimiento de combustible, del kilometraje por auto y de la cantidad de autos por habitante. Brons et al. (2006) encuentran una baja sensibilidad del precio con respecto a la demanda total de gasolina en las economías de Estados Unidos, Canadá y Australia, debido a una baja sensibilidad del precio con respecto al número de autos poseídos por habitante. Esto sugiere una alta dependencia al transporte automotriz, por lo que una política orientada a incrementar la carga impositiva a la gasolina será más efectiva si es acompañada de cargos por registro y cuotas fijas por la compra de vehículos. Las investigaciones meta-analíticas relevan que los estudios basados en datos de corte transversal suelen reportar elasticidades con mayor valor absoluto que aquellas que usan series de tiempo.

Para el caso mexicano, Crôtte, Noland y Graham (2010) estimaron las elasticidades de la demanda de gasolina por vehículo con respecto al precio, ingreso, parque vehicular y las tarifas del metro usando dos metodologías, un modelo cointegrado de series de tiempo, cuya estimación está basada en datos del periodo comprendido entre 1980 y 2006, y un modelo basado en el Método Generalizado de Momentos (GMM) que recoge datos de 1993 a 2006. Los resultados obtenidos mediante sus distintas especificaciones arrojan elasticidades con los mismos signos pero difieren considerablemente en magnitud, lo cual atribuyen en buena medida a la calidad de los datos. Con base en sus estimaciones de la elasticidad del ingreso concluyen que para mantener el nivel de consumo de gasolina el precio tendría que aumentar a un mayor ritmo que el PIB.

Fullerton, Ibarra y Elizalde (2015) realizan un análisis independiente para los dos grados de gasolina disponibles en México usando datos de consumo previos a la crisis de 2008, mediante técnicas de cointegración y corrección de errores. Sus resultados refieren que el efecto ingreso supera al efecto sustitución, lo que provoca que las curvas de demanda de ambos tipos de gasolina tengan pendiente positiva y, por tal motivo, ambos productos sean tratados como bienes normales. Los investigadores atribuyen estos hallazgos a que el régimen de precios vigente mantenía un precio artificialmente bajo que impedía su equilibrio a través de la acción de las fuerzas de mercado.

La predicción del consumo de combustibles ha cobrado especial importancia como herra-

mienta para la planeación energética, ayudando a los hacedores de política en el desarrollo de un adecuado sistema tributario y de fijación de precios, aportando información fundamental para la toma de decisiones de inversión en infraestructura que garanticen la seguridad energética, coadyuvando en el diseño de políticas ambientales mediante estimaciones adelantadas sobre emisiones a la atmósfera, así como la planificación de necesidades energéticas futuras y la identificación de necesidades de investigación y desarrollo.

# Capítulo 4

## Motivación del proyecto

Una predicción confiable de la demanda es un elemento clave para la toma de decisiones a lo largo de toda la cadena productiva de los petrolíferos, pero con principal énfasis en las actividades de gestión de inventarios, planeación de la producción y control de procesos.

En el contexto mexicano, la predicción de la demanda cobra una especial relevancia debido a la situación actual de la infraestructura logística y la reciente transformación estructural del mercado de petrolíferos. Si bien se tienen identificadas las terminales con menor inventario para la provisión de combustibles, las regiones no presentan un patrón de demanda homogéneo y no se espera que en el futuro lo hagan, por lo que un pronóstico diferenciado podrá influir en las decisiones de inversión para garantizar la seguridad energética y la implementación de políticas ambientales diseñadas a la medida.

El objetivo principal de este trabajo de investigación es el desarrollo de un sistema para la predicción de la demanda mensual de gasolinas de las TAR del país. La metodología estadística aplicada se centra en un modelo de regresión dinámico multivariable, el cual considera variables externas en un proceso estacional autorregresivo integrado independiente para cada clúster de terminales que presenten una fuerte correlación.

Los clústeres que agrupan de mejor manera las terminales resultaron ser las entidades fedrativas, derivado de que cada estado es abastecido por aquella infraestructura de almacenamiento que se localiza dentro de su perímetro geográfico. Las excepciones a lo anterior son los estados de Tlaxcala y Quintana Roo, los cuales no cuentan con terminales de almacenamiento y son provistas por TAR ubicadas en Puebla y Yucatán respectivamente.

# Capítulo 5

## Datos

La muestra usada para la estimación y evaluación de los modelos incluye 190 observaciones mensuales de consumo promedio diario de gasolina de enero de 2002 a octubre de 2017 para 78 terminales distribuidas en 30 entidades de la República Mexicana. La información fue provista por el área de planeación de Pemex Logística. La variable dependiente en cada ecuación es el promedio de consumo diario de gasolina para el próximo mes medido en miles de barriles diarios.

### **Las variables exógenas**

#### **Población**

Considera el número de habitantes de cada unidad de análisis. Los datos fueron obtenidos a partir de los factores de expansión de la Encuesta Nacional de Ocupación y Empleo, la cual es generada trimestralmente por el INEGI y es representativa a nivel estatal, por lo que para el cálculo de los datos mensuales se realizó una interpolación lineal.

#### **Parque vehicular**

Constituye el número de vehículos registrados en cada unidad de análisis. Los datos fueron extraídos de la estadística de Vehículos de Motor Registrados en Circulación (VMRC) del INEGI, misma que proporciona información mensual y anual sobre la composición y evolución del parque vehicular registrado que circula en el territorio nacional, identificando la clase de vehículo y el tipo de servicio.

A nivel nacional, la serie de los VMRC tiene una frecuencia mensual, sin embargo, a nivel entidad la estadística es anual, por lo que para completar estas series, se computó la proporción anual de cada estado en relación al parque total, usando las cifras estatales al cierre de cada año y el número de vehículos totales en ese periodo. Mediante esta fracción con variación anual y el número de autos registrados por mes en el país para cada año, se generaron los

datos para el parque vehicular estatal.

### **Actividad económica**

Como la literatura relacionada lo sugiere, el PIB suele concentrar cierto poder explicativo del proceso de demanda de combustibles, empero este indicador para las entidades es presentado por el INEGI con una frecuencia anual para los estados. En razón de lo anterior, este análisis utiliza como variable proxy el Indicador Trimestral de la Actividad Económica Estatal (ITAE), el cual brinda información de corto plazo que permite tener un panorama de la evolución económica de las entidades federativas, y cuya elaboración sigue los conceptos, criterios metodológicos, clasificadores y datos fuente, que se emplean en los cálculos del Producto Interno Bruto Trimestral y el PIB por Entidad Federativa (PIBE). A partir de las series del ITAE, únicamente fue necesario hacer uso de una interpolación lineal para obtener los datos mensuales.

### **Precio promedio**

El Banco de Información Económica que pone a disposición el INEGI reporta el precio nominal de las gasolinas magna y premium para la zona de la frontera norte y el resto del país desde enero de 1997 a diciembre de 2016. Los precios promedio mensuales por estado para 2017 fueron obtenidos de la Comisión Reguladora de Energía (CRE). Para la preparación de una sola serie de precios de gasolina se ponderó el precio por litro para cada octanaje usando el consumo puntual, de esta forma, los precios varían entre las entidades incluso antes de la liberación de los mismos. Los precios nominales son pasados a pesos de enero de 2017 mediante deflatores calculados a partir de del Índice Nacional de Precios al Consumidor.

# Capítulo 6

## Modelo para la demanda nacional

Derivado del análisis de la literatura relacionada y el objetivo fundamental de la presente investigación, misma que persigue obtener pronósticos de demanda precisos y no así la estimación de elasticidades de variables exógenas, se justifica la aplicación de un modelo basado en series de tiempo, mismos que han sido usados extensamente por su alto poder predictivo. Específicamente los modelos autorregresivos de media móvil (ARMA) enuncian que el valor actual de una serie depende linealmente de sus valores previos más una combinación de los valores, actual y anteriores, y del término del error, el cual sigue un proceso de ruido blanco. La función de autocorrelación (acf) del proceso ARMA despliega una mezcla entre los comportamientos de las partes AR y MA, pero para rezagos mayores al orden  $p$  del modelo la acf será igual a la del modelo AR individual, siendo ésta la parte que dominará en el largo plazo. Por otro lado, la función de autocorrelación parcial (pacf) es particularmente útil en este contexto debido a que puede distinguir entre un proceso puramente autorregresivo o un proceso combinado.

El correlograma característico de un proceso con ambos componentes mostrará que tanto la acf como la pacf decrecen geoméricamente.

Los primeros en diseñar una aproximación sistemática para la estimación de modelos ARMA fueron Box y Jenkins (1976), la cual está organizada en tres etapas: identificación, estimación y evaluación. La identificación al estilo de Box y Jenkins reside exclusivamente en el análisis del correlograma para hacer una primera propuesta con respecto al orden del modelo. Una técnica que elimina la subjetividad inherente a la interpretación gráfica es la conocida como criterios de información. Las fórmulas más usadas son las de Akaike, Schwarz y Hannan-Quinn, y están basadas en dos términos, el primero es común a las tres y corresponde al logaritmo de la varianza residual del modelo, el segundo término se añade con el objeto de penalizar la pérdida de grados de libertad asociada a la incorporación de un rezago adicional al modelo ARMA.

El presente análisis usa el criterio de Schwarz ya que éste penaliza más drásticamente la

inclusión de términos extra, con lo cual se promueve la estimación de modelos parsimoniosos que únicamente incluirán términos que consigan minimizar el error residual más allá del detrimento en los grados de libertad.

Dada la relevancia de la precisión en la predicción, esta investigación examina otras formas de discriminación entre modelos, como el error cuadrado medio y su descomposición en proporciones de sesgo, varianza y covarianza, además de un diagnóstico de los residuales estimados.

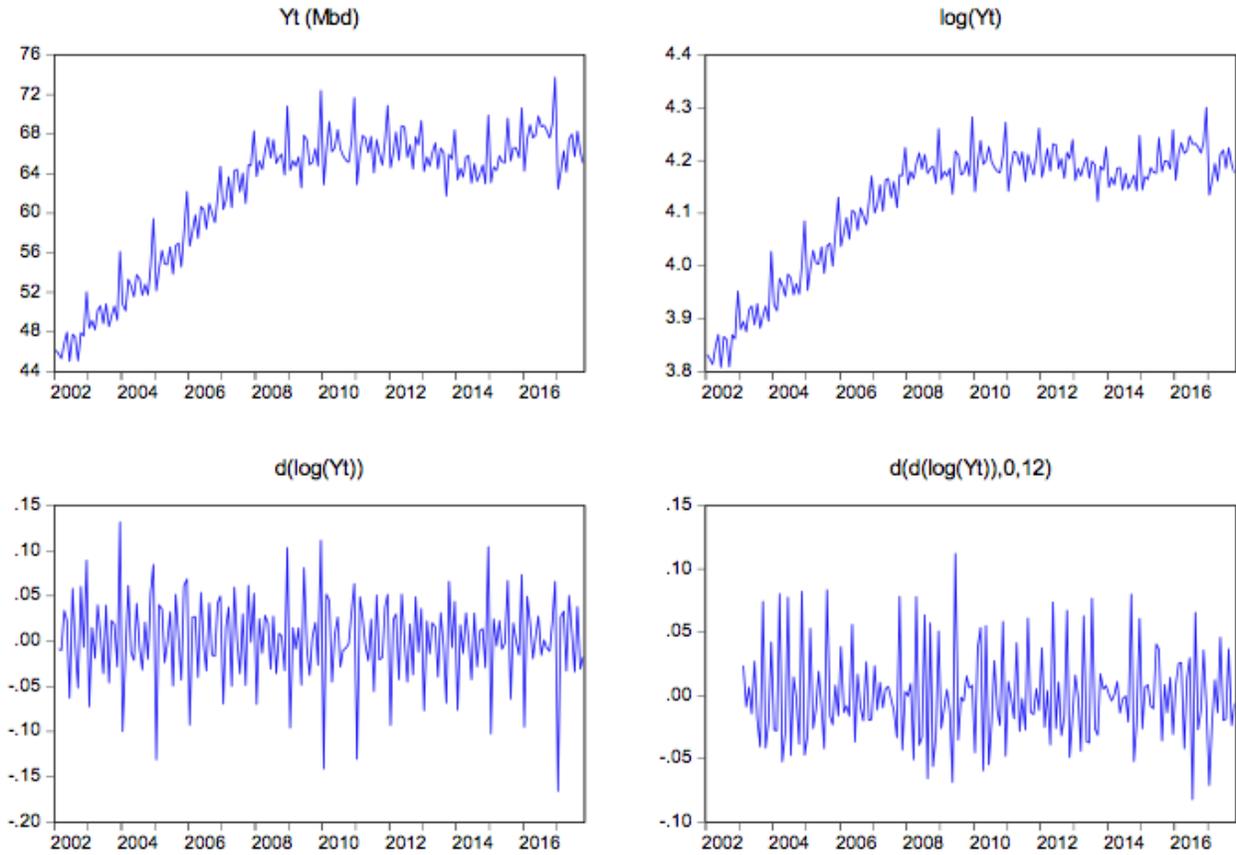
La demanda de gasolina no es directamente observable, por lo que este trabajo supone que el consumo de gasolina, es decir, la gasolina suministrada por cada una de las TAR es proporcional a su demanda y depende de las mismas variables, lo cual no resulta ilógico si tenemos en cuenta la limitada capacidad de almacenamiento a nivel nacional. Por los motivos anteriores, las demandas agregadas corresponderán a las entregas mensuales de las TAR hacia las entidades y estarán expresadas en miles de barriles diarios (Mbd).

Las 190 observaciones que integran la muestra fueron divididas en dos subconjuntos. Los 168 primeros datos, de enero de 2002 a diciembre de 2015, se usaron para la estimación del modelo y los 22 restantes, enero de 2016 a octubre de 2017, se destinaron a la evaluación de la predicción.

En consonancia con la metodología de modelos predictivos basados en series de tiempo, se realizaron tres transformaciones a la variable dependiente. El primero corresponde a una transformación logarítmica debido a que se observa que la varianza aumenta marginalmente en el tiempo, con esto conseguimos que la curtosis de la distribución pase de 2.79 a 3.08. Posteriormente se calculó la primera diferencia dado que el proceso muestra una clara tendencia creciente, sin embargo, esto no fue suficiente para rechazar la hipótesis nula de la existencia de una raíz unitaria en el polinomio característico, por lo que se realizó una primera diferencia estacional, a partir de lo cual no sólo estamos en posición de rechazar la hipótesis nula de la prueba de Dickey-Fuller aumentada con un nivel de significancia de 99.95 %, sino que también fue posible mitigar el comportamiento estacional.

Las transformaciones y el proceso a modelar se presentan en el siguiente gráfico:

Figura 6.1: Transformaciones y proceso a modelar



Teniendo listo el proceso a modelar, se calcularon las funciones de correlación, obteniendo así el siguiente correlograma.

Figura 6.2: Correlograma

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
█	█	1 -0.542	-0.542	52.877	0.000
█	█	2 -0.097	-0.553	54.587	0.000
█	█	3 0.328	-0.110	74.171	0.000
█	█	4 -0.254	-0.165	85.959	0.000
█	█	5 0.036	-0.138	86.203	0.000
█	█	6 0.193	0.093	93.115	0.000
█	█	7 -0.209	0.071	101.27	0.000
█	█	8 -0.060	-0.211	101.96	0.000
█	█	9 0.364	0.191	126.88	0.000
█	█	10 -0.393	-0.059	156.24	0.000
█	█	11 0.262	0.276	169.31	0.000
█	█	12 -0.100	-0.083	171.22	0.000
█	█	13 -0.196	-0.260	178.67	0.000
█	█	14 0.344	-0.069	201.64	0.000
█	█	15 -0.110	0.095	204.03	0.000
█	█	16 -0.213	-0.122	212.96	0.000
█	█	17 0.244	-0.040	224.75	0.000
█	█	18 0.014	0.106	224.79	0.000
█	█	19 -0.256	0.044	237.96	0.000
█	█	20 0.269	-0.162	252.60	0.000
█	█	21 -0.115	0.030	255.28	0.000
█	█	22 -0.122	-0.024	258.34	0.000
█	█	23 0.276	0.069	273.95	0.000
█	█	24 -0.158	0.012	279.13	0.000

El proceso evidencia una fuerte autocorrelación parcial en los primeros dos rezagos, a partir de ahí se presenta una caída brusca, y vuelve a sobrepasar las bandas a partir del octavo rezago. Con respecto a la autocorrelación total vemos que ésta decae después del primer rezago e igualmente recobra significancia en los rezagos nueve y diez.

El correlograma es considerablemente complejo de interpretar, pero a partir de su análisis podemos inferir que la parte autorregresiva debe dominar y que algún rezago entre el octavo y el décimo contiene información sustancial sobre la dinámica de la serie.

Teniendo lo anterior en mente, se evaluaron varias especificaciones usando los criterios descritos arriba. Algunas de ellas incluyeron variables exógenas como las segundas diferencias del log precio de la gasolina, del Indicador Global de la Actividad Económica (IGAE) del sector terciario<sup>3</sup>, del parque vehicular, del número de vehículos por habitante y de la población. Se encontró que la ecuación que mejor pronostica la demanda nacional de gasolinas es un modelo SARIMA(9, 1, 4) × (0, 1, 1)<sub>12</sub>, con coeficientes no significativos para los rezagos tres a ocho en el proceso autorregresivo y uno a dos en el proceso media móvil:

$$\phi_9(B)\nabla_{12}^1\nabla^1\log(y_t) = \Theta_1(B^{12})\theta_4(B)u_t$$

Donde:

- $\phi_9(B) = (1 - \phi_1B - \phi_2B^2 - \phi_9B^9)$  es el proceso AR estándar de orden 9
- $\nabla_{12}^1 = (1 - B^{12})$  representa las primeras diferencias estacionales
- $\nabla^1 = (1 - B)$  representa las primeras diferencias estándar
- $\Theta_1(B^{12}) = (1 - \Theta_1B^{12})$  es el proceso estacional MA de orden 1
- $\theta_4(B) = (1 - \theta_3B^3 - \theta_4B^4)$  es el proceso MA estándar de orden 4
- $u_t$  es el proceso de ruido blanco

Para evaluar la dependencia lineal en el proceso de los residuales se calculó el estadístico de Ljung-Box (1978), con lo cual fue posible descartar que los errores conservaran alguna estructura discernible y por lo tanto puede concluirse que siguen un proceso de ruido blanco con una media estadísticamente igual a cero.

---

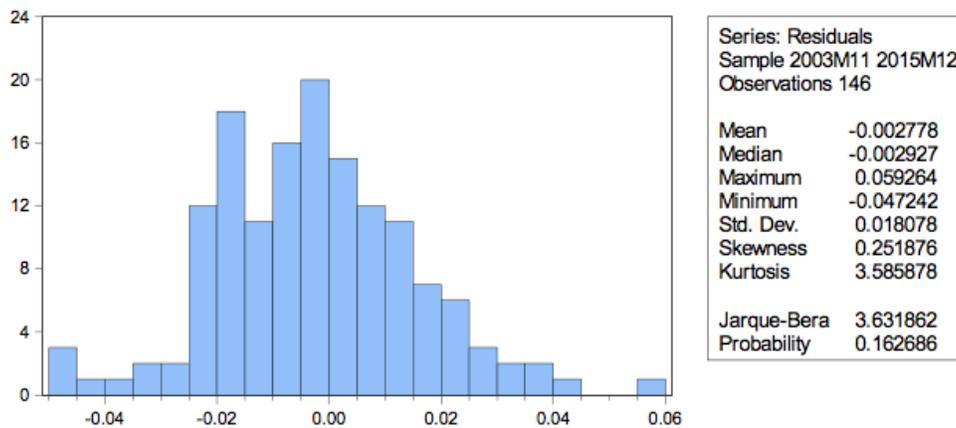
<sup>3</sup>El sector terciario engloba las actividades relacionadas con los servicios no productores o transformadores de bienes materiales y resultó ser el que mayor correlación presentó con la serie de tiempo del consumo de gasolina, incluso por encima del IGAE correspondiente al sector transporte.

Figura 6.3: Cálculo del estadístico de Ljung-Box

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.025	0.025	0.1016	
		2	-0.098	-0.098	1.6153	
		3	0.001	0.006	1.6156	
		4	-0.033	-0.043	1.7919	
		5	0.093	0.097	3.1961	
		6	-0.003	-0.017	3.1976	
		7	-0.027	-0.006	3.3150	0.069
		8	-0.004	-0.009	3.3182	0.190
		9	0.060	0.066	3.9135	0.271
		10	0.062	0.048	4.5611	0.335
		11	-0.061	-0.053	5.1845	0.394
		12	-0.062	-0.048	5.8480	0.440
		13	0.090	0.090	7.2483	0.403
		14	0.131	0.114	10.219	0.250
		15	-0.131	-0.142	13.199	0.154
		16	-0.004	0.034	13.202	0.213
		17	-0.025	-0.034	13.312	0.273
		18	0.027	0.030	13.439	0.338
		19	0.034	-0.021	13.644	0.399
		20	0.021	0.062	13.721	0.471
		21	0.056	0.068	14.297	0.503
		22	-0.001	0.001	14.297	0.577
		23	0.004	-0.024	14.301	0.646
		24	-0.038	-0.032	14.563	0.692

Como parte del diagnóstico de residuales, también se encontró que éstos presentan una distribución en forma de campana cuya curtosis y asimetría son consistentes con el supuesto de normalidad de la prueba de Jarque-Bera (1987).

Figura 6.4: Prueba de Jarque-Bera



La estimación de los coeficientes de la regresión de LS fue la siguiente:

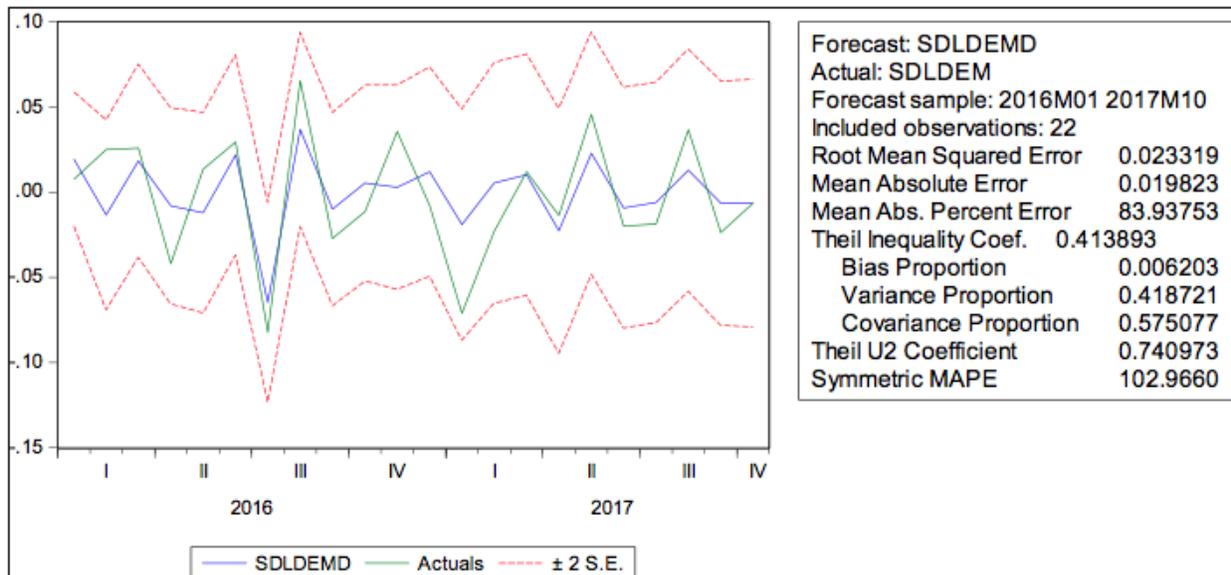
Tabla 6.1: Coeficientes de la regresión de LS

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-1.005427	0.062537	-16.077400	0.0000
AR(2)	-0.703846	0.082041	-8.579225	0.0000
AR(9)	0.201518	0.062450	3.226881	0.0015
MA(3)	-0.285333	0.092891	-3.071710	0.0025
MA(4)	0.263699	0.090542	2.912440	0.0041
SMA(12)	-0.710876	0.080734	-8.805139	0.0000
SIGMASQ	0.000357	0.000038	9.419218	0.0000
R-squared	0.741357	Mean dependent var		-0.000239
Adjusted R-squared	0.730872	S.D. dependent var		0.037277
S.E. of regression	0.019338	Akaike info criterion		-4.936238
Sum squared resid	0.055349	Schwarz criterion		-4.798793
Log likelihood	389.5584	Hannan-Quinn criter.		-4.880411
Durbin-Watson stat	1.895686			

Es importante destacar el hecho de que ninguno de los coeficientes correspondientes a las variables exógenas resultó significativo, por lo fueron descartadas del modelo final.

En el siguiente gráfico podemos apreciar el ajuste en la predicción dinámica con respecto a los valores reales de la submuestra separada para evaluación del pronóstico, además de algunas medidas de su precisión.

Figura 6.5: Ajuste en la predicción dinámica con respecto a los valores reales de la submuestra



De las estadísticas precedentes destacamos que la proporción del error cuadrático medio asociada con el sesgo es prácticamente cero y que el valor de éste es considerablemente bajo.

## Capítulo 7

# Modelos para la predicción de la demanda a las TAR

Las 77 terminales propiedad de PEMEX que actualmente suministran combustibles hacia las casi 12,000 estaciones de servicio del país están distribuidas en 29 entidades federativas, únicamente Tlaxcala y Quintana Roo no cuentan con infraestructura de almacenamiento dentro de sus límites territoriales. Debido a la extensión, geografía y nivel de consumo, estados como Aguascalientes, Baja California Sur, Campeche, Nayarit, Querétaro, Tabasco y Zacatecas albergan solamente una TAR, mientras que algunos otros como Tamaulipas, Sonora y Veracruz concentran cinco, seis y ocho, respectivamente.

Sin perder de vista que el objetivo fundamental de este estudio es generar pronósticos, lo más precisos posibles, para la demanda de gasolina a cada TAR y dada la alta correlación entre las ventas de las terminales ubicadas dentro de un mismo estado, se observó la necesidad de generar un modelo por entidad. En consecuencia, el número de variables endógenas (dependientes) a pronosticar por cada modelo es igual al número de TAR que abastecen a cierto estado en particular.

En razón de lo anterior, para los estados con una terminal se implementó un procedimiento muy similar al aplicado en la estimación de la demanda nacional, un modelo ARIMA que considera las variables exógenas descritas anteriormente.

Como ejemplo ilustrativo se presenta el modelo obtenido para el estado de Baja California Sur, el cual es aprovisionado por la terminal La Paz, misma que cuenta con 10 tanques que

se traducen en una capacidad operativa de 161,793 barriles, que representan menos de seis días de abastecimiento y es alimentada mediante buque-tanque desde la refinería de Salina Cruz.

El modelo determinado para Baja California Sur es un SARIMAX  $(2, 1, 12) \times (1, 0, 0)_{12}$ :

$$\phi_2(B)\nabla_{12}^1\nabla^1\log(y_t) = \theta_{12}(B)u_t + \psi\nabla^1\log(AE_t)$$

Donde:

- $\phi_2(B) = (1 - \phi_1B - \phi_2B^2)$  es el proceso AR estándar de orden 2
- $\nabla_{12}^1 = (1 - B^{12})$  representa las primeras diferencias estacionales del proceso AR
- $\nabla^1 = (1 - B)$  representa las primeras diferencias estándar
- $\theta_{12}(B) = (1 - \theta_{12}B^{12})$  es el proceso estacional MA estándar de orden 12
- $u_t$  es el proceso de ruido blanco
- $\psi\nabla^1\log(AE_t) = (1 - \psi B)\log(AE_t)$  representa las primeras diferencias del logaritmo del índice de actividad económica

En este caso, la variable dependiente  $y_t$  constituye las ventas de gasolina, ambos octanajes, de la TAR La Paz en el tiempo  $t$ . En la siguiente tabla se muestran los valores estimados para los coeficientes de los rezagos y del efecto del cambio logarítmico en la variable proxy de actividad económica, misma que resultó significativa con un nivel de confianza de 97.7%.

Tabla 7.1: Valores estimados para los coeficientes de los rezagos y del efecto del cambio logarítmico en la variable proxy de actividad económica

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLAE	0.196658	0.085531	2.299267	0.0228
AR(1)	-0.865401	0.063409	-13.647980	0.0000
AR(2)	-0.486717	0.056815	-8.566630	0.0000
SAR(12)	0.987051	0.013762	71.72429	0.0000
MA(12)	-0.783389	0.098559	-7.948464	0.0000
SIGMASQ	0.001019	0.000103	9.893232	0.0000

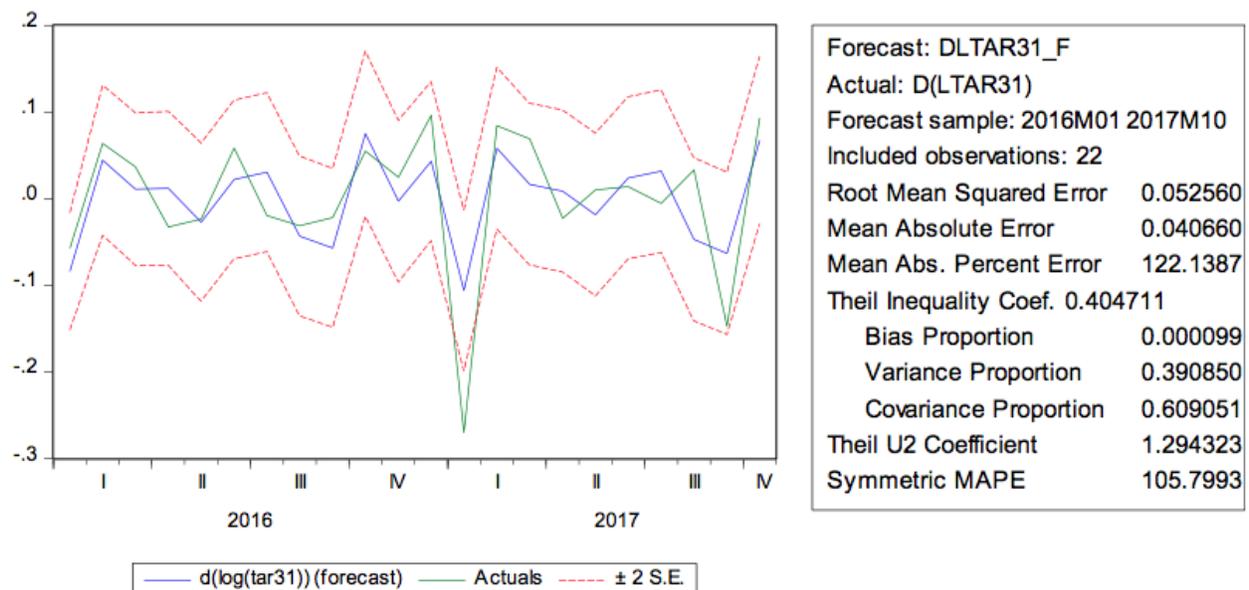
El modelo apunta que el incremento de un punto porcentual en el índice de la actividad económica de Baja California Sur, provoca un incremento cercano al 0.2% en la demanda

de gasolina.

Los residuales generados por el modelo pasan la prueba de autocorrelación de Ljung-Box, por lo que puede concluirse que siguen un proceso de ruido blanco con media constante y estadísticamente igual a cero, además de que la curtosis y asimetría de su distribución son consistentes con las de una distribución normal.

El nivel de precisión en las predicciones obtenidas a partir del modelo estimado es considerablemente bueno. Esta aseveración se respalda mediante el cálculo de las medidas típicas para evaluar desviaciones entre valores reales y pronosticados. El modelo arroja un error cuadrado medio de 0.0027 y una proporción de sesgo muy cercana a cero, usando la variable dependiente transformada. El siguiente gráfico contrasta las ventas reales contra las predicciones dinámicas del modelo.

Figura 7.1: Contraste entre las ventas reales y las predicciones dinámicas del modelo



Al igual que en el caso nacional, observamos que el modelo subestima la caída en la demanda acontecida en enero de 2017. Esta desviación es comprensible debido a que corresponde a un suceso inédito como lo fue la liberalización de los precios de la gasolina, del cual no existe ningún antecedente en la submuestra usada para la estimación de los modelos.

## Modelos multivariable

Los modelos de vectores autorregresivos (VAR) representan la generalización natural de los modelos autorregresivos univariados, éstos se usan cuando un sistema tiene más de una variable endógena (Sims 1980).

Para los 23 estados cuya demanda es abastecida por más de una terminal de almacenamiento se desarrolló un modelo VAR. Estos modelos incorporan variables exógenas en los casos en los que éstas mejoraron el nivel predictivo y resultaron estadísticamente significativas.

Particularmente, el estado de Nuevo León cuenta con dos terminales de almacenamiento, Cadereyta y Santa Catarina, las cuales están separadas por 64 kilómetros y tienen una capacidad operativa conjunta de 661 miles barriles de combustible y ambas son alimentadas por ducto proveniente de la refinería de Cadereyta.

El modelo con mejor ajuste para el estado de Nuevo León es un VAR(12) que considera únicamente los rezagos 1, 2, 10 y 12 de las variables independientes. La expresión en notación matricial es:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} a_{1,1}^1 & a_{1,2}^1 \\ a_{2,1}^1 & a_{2,2}^1 \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1,1}^2 & a_{1,2}^2 \\ a_{2,1}^2 & a_{2,2}^2 \end{bmatrix} \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \end{bmatrix} + \begin{bmatrix} a_{1,1}^{10} & a_{1,2}^{10} \\ a_{2,1}^{10} & a_{2,2}^{10} \end{bmatrix} \begin{bmatrix} y_{1,t-10} \\ y_{2,t-10} \end{bmatrix} \\ + \begin{bmatrix} a_{1,1}^{12} & a_{1,2}^{12} \\ a_{2,1}^{12} & a_{2,2}^{12} \end{bmatrix} \begin{bmatrix} y_{1,t-12} \\ y_{2,t-12} \end{bmatrix} + \begin{bmatrix} B_{1,1} & B_{1,2} & B_{1,3} \\ B_{2,1} & B_{2,2} & B_{2,3} \end{bmatrix} \begin{bmatrix} x_{1,t} \\ x_{2,t} \\ x_{3,t} \end{bmatrix} + \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix}$$

Donde:

- $y_{1,t}$  es la primera diferencia del logaritmo de las ventas de la TAR Santa Catarina (TAR43)
- $y_{2,t}$  es la primera diferencia del logaritmo de las ventas de la TAR Cadereyta (TAR6)
- $x_{1,t}$  es la primera diferencia del logaritmo de la población en millones (POBM)
- $x_{2,t}$  es la primera diferencia del precio real ponderado por litro (PREC\_R)
- $x_{3,t}$  es la primera diferencia del logaritmo del índice de actividad económica (AE)

Expresando el modelo en notación compacta tenemos:

$$Y_t = C + A_1Y_{t-1} + A_2Y_{t-2} + A_{10}Y_{t-10} + A_{12}Y_{t-12} + BX_t + U_t$$

Al igual que en el caso univariado, la identificación del orden del modelo se basó primeramente en la minimización del criterio de información de Schwarz. Posteriormente, se calcularon las funciones de autocorrelación y autocorrelación parcial para los procesos de los residuales de ambas ecuaciones, una por variable endógena, y se aplicó la prueba de Ljung-Box para evaluar conjuntamente si las correlaciones, hasta el rezago 36, son diferentes a cero. Con esto se busca garantizar que los residuales no conserven ninguna estructura, es decir, hayamos extraído la dinámica del proceso que queremos explicar mediante la inserción de los rezagos adecuados.

En razón de lo anterior, y teniendo en mente la estacionalidad y frecuencia del proceso, todos los procesos multivariados incluyen el décimo segundo rezago. El siguiente cuadro contiene los valores estimados para los coeficientes del modelo.

Tabla 7.2: Valores estimados para los coeficientes del modelo

	Santa Catarina d(log(TAR43))	Cadereyta d(log(TAR6))
d(log(TAR43(-1)))	-0.630106 (0.07816) [-8.06159]	0.016132 (0.14327) [0.11259]
d(log(TAR43(-2)))	-0.308289 (0.06982) [-4.41573]	0.099013 (0.12797) [0.77370]
d(log(TAR43(-10)))	0.052056 (0.06359) [0.81856]	0.31232 (0.11657) [2.67923]
d(log(TAR43(-12)))	0.244645 (0.06968) [3.51107]	0.525749 (0.12772) [4.11636]
d(log(TAR6(-1)))	-0.096969 (0.04498) [-2.15602]	-0.291156 (0.08244) [-3.53167]
Continúa		

Continuación		
	Santa Catarina d(log(TAR43))	Cadereyta d(log(TAR6))
d(log(TAR6(-2)))	0.034935 (0.04626) [0.75522]	-0.178961 (0.08479) [-2.11059]
d(log(TAR6(-10)))	-0.041135 (0.04453) [-0.92387]	-0.245989 (0.08162) [-3.01398]
d(log(TAR6(-12)))	0.10554 (0.04584) [ 2.30245]	-0.028325 (0.08402) [-0.33711]
C	-0.010412 (0.01212) [-0.85940]	-0.065621 (0.02221) [-2.95471]
d(log(POBM))	5.315951 (8.23629) [0.64543]	59.0123 (15.09730) [3.90879]
d(log(PREC_R))	1.549947 (0.65913) [2.35151]	-2.568583 (1.20820) [-2.12596]
d(log(AE))	0.632337 (0.29129) [2.17082]	0.242127 (0.53394) [0.45347]
Fin de la tabla		
Nota: Errores estándar en paréntesis y estadísticos t en corchetes		

Si centramos nuestra atención en las estimaciones obtenidas para las variables exógenas, podemos advertir que el crecimiento poblacional tiene poder explicativo únicamente para la terminal de Cadereyta, mientras que el crecimiento en la actividad económica es significativo exclusivamente para las ventas de la terminal de Santa Catarina.

Con respecto al precio ponderado real se observa un fenómeno interesante, y es que éste resulta significativo para ambas series de demanda, no obstante, los coeficientes tienen signo opuesto, lo que refleja un efecto sustitución. Esto podría interpretarse como que ante un incremento en el precio del combustible, la demanda de la terminal de Cadereyta disminuye en favor de un mayor consumo de la terminal de Santa Catarina. Dado que la TAR Santa Catarina y el centro de la ciudad de Monterrey están separados por una distancia de aproxi-

madamente la mitad de la que divide al centro y la terminal de Cadereyta, un incremento en el precio puede generar incentivos para que los minoristas de gasolinas adquieran el producto en la terminal más cercana a los puntos de venta.

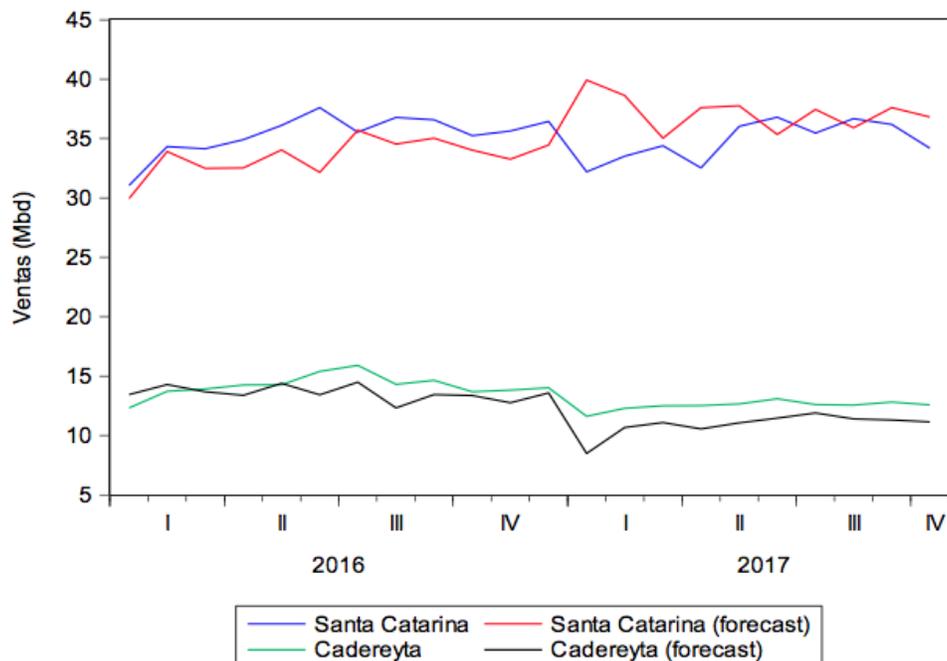
La evaluación de la predicción del modelo se basa en el nivel de ajuste de los valores pronosticados con respecto a los reales. La medida usada para este propósito fue la raíz del error cuadrático medio (RMSE) asociado a las predicciones generadas por cada ecuación, los cuales se presentan en el siguiente cuadro junto con otras medidas comúnmente reportadas en la literatura.

Tabla 7.3: RMSE asociado a las predicciones generadas por cada ecuación

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
$d(\log(\text{TAR43}))$	22	0.073391	0.04827	170.6507	0.602324
$d(\log(\text{TAR6}))$	22	0.07554	0.052474	146.7666	0.417237

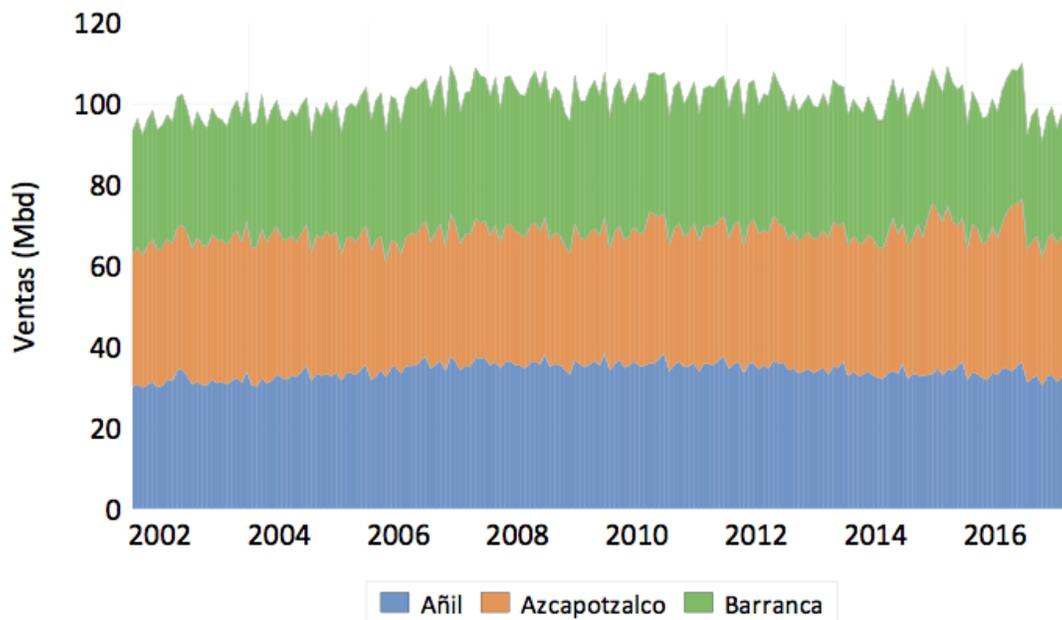
El siguiente gráfico muestra las ventas reales y los pronósticos calculados a partir del modelo para ambas TAR:

Figura 7.2: Ventas reales y pronósticos calculados a partir del modelo



El último modelo que presentaremos es el de la Ciudad de México. La capital del país es abastecida por tres terminales: Añil, Azcapotzalco y Barranca del Muerto, y tienen una capacidad operativa en miles de barriles de 158, 1,108 y 92, respectivamente, siendo Azcapotzalco la terminal con mayor capacidad de México. A pesar de la diferencia de tamaños, la demanda de gasolina de la ciudad es cubierta en proporciones prácticamente iguales.

Figura 7.3: Tendencia del consumo de gasolinas de la Ciudad de México



Separando la tendencia del consumo de gasolinas de la Ciudad de México mediante el filtro Hodrick-Prescott, vemos que ésta presenta un máximo global de 103.9 Mbd en octubre de 2008, razón por la cual su participación con respecto al consumo nacional pasa de 16.2% en 2003 a 12.5% en 2016.

Si bien las series correspondientes a las ventas de las terminales, no parecen seguir una tendencia clara, tanto el proceso asociado a la TAR Añil, como Barranca del Muerto exhiben un comportamiento que asemeja al de una parábola, por lo que fue necesario calcular segundas diferencias para poder asegurar su estacionariedad. Teniendo en cuenta lo anterior, y aunque la serie de la terminal Azcapotzalco es estacionaria con un nivel de confianza de 99.95%, se usaron segundas diferencias de las transformaciones logarítmicas para cada variable depen-

diente.

Usando los mismos criterios que en el caso de Nuevo León, el modelo en notación compacta identificado para la Ciudad de México es un VAR(12) que incorpora todos los rezagos intermedios y no contempla variables exógenas al sistema<sup>4</sup>:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + A_3 Y_{t-3} + A_4 Y_{t-4} + A_5 Y_{t-5} + A_6 Y_{t-6} + A_7 Y_{t-7} + A_8 Y_{t-8} \\ + A_9 Y_{t-9} + A_{10} Y_{t-10} + A_{11} Y_{t-11} + A_{12} Y_{t-12} + U_t$$

Donde:

- $y_{1,t}$  es la segunda diferencia del logaritmo de las ventas de la TAR Añil (TAR3)
- $y_{2,t}$  es la segunda diferencia del logaritmo de las ventas de la TAR Azcapotzalco (TAR4)
- $y_{3,t}$  es la segunda diferencia del logaritmo de las ventas de la TAR Barranca del Muerto (TAR5)

El modelo anterior no parece seguir el principio de parsimonia, aún más si consideramos que el modelo que minimiza el criterio de Schwarz requiere sólo de cinco rezagos. Sin embargo, los procesos de las terminales de Añil y Barranca del Muerto presentan una prolongada memoria y, por lo tanto, la eliminación de la correlación en sus residuales precisó la adición de rezagos más lejanos. Para verificar lo anterior se realizó una prueba de exclusión de rezagos basada en el estadístico Wald, los resultados se muestran a continuación.

Tabla 7.4: Prueba de exclusión de rezagos

	Añil	Azcapotzalco	Barranca del Muerto	Joint
	$y_1$	$y_2$	$y_3$	
Lag 1	358.1490 [0.000000]	259.2824 [0.000000]	391.2329 [0.000000]	787.1732 [0.000000]
Lag 2	129.9495 [0.000000]	101.9168 [0.000000]	142.7920 [0.000000]	320.2304 [0.000000]
Lag 3	63.92563 [8.52e-14]	47.38166 [2.88e-10]	67.44414 [1.51e-14]	157.7768 [0.000000]
Continúa				

<sup>4</sup>Ninguna de las variables exógenas resultó estadísticamente significativa.

Continuación				
	Añil	Azcapotzalco	Barranca del Muerto	Joint
	$y_1$	$y_2$	$y_3$	
Lag 4	49.46465 [1.04e-10]	44.12937 [1.42e-09]	47.22222 [3.12e-10]	115.6336 [0.000000]
Lag 5	33.87988 [2.10e-07]	33.14769 [3.00e-07]	37.33997 [3.90e-08]	76.47886 [8.06e-13]
Lag 6	23.64730 [2.96e-05]	23.74416 [2.82e-05]	23.12367 [3.81e-05]	58.17979 [3.00e-09]
Lag 7	23.61091 [3.01e-05]	19.56973 [0.000208]	21.53151 [8.16e-05]	49.61477 [1.27e-07]
Lag 8	23.33880 [3.43e-05]	18.67544 [0.000319]	23.47131 [3.22e-05]	42.69233 [2.45e-06]
Lag 9	20.47415 [0.000135]	12.60917 [0.005563]	23.65951 [2.94e-05]	33.08526 [0.000129]
Lag 10	21.30364 [9.10e-05]	10.40865 [0.015394]	18.44060 [0.000357]	36.47001 [3.27e-05]
Lag 11	21.56714 [8.03e-05]	7.346064 [0.061648]	13.76927 [0.003237]	38.30582 [1.54e-05]
Lag 12	4.920432 [0.177717]	3.094573 [0.377272]	8.551399 [0.035890]	19.51800 [0.021132]
df	3	3	3	9

Fin de la tabla

Nota: Valores  $p$  entre corchetes

Debido a la longitud del modelo, el cuadro con los coeficientes estimados y algunos estadísticos más se muestran en el anexo A.

Las medidas de precisión de la predicción obtenidas para las tres ecuaciones están contenidas en la siguiente tabla:

Tabla 7.5: Medidas de precisión de la predicción

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
d(d(log(TAR3)))	22	0.03207	0.025499	75.13705	0.167078
d(d(log(TAR4)))	22	0.064326	0.048822	161.2024	0.267473
d(d(log(TAR5)))	22	0.042776	0.029874	129.1279	0.200667

La metodología VAR ofrece una herramienta analítica muy poderosa, la función impulso-respuesta (IRF). La utilidad de la IRF radica en que facilita el monitoreo de las respuestas de las variables del sistema ante shocks aplicados directamente al término del error idio-

sincrático.

La IRF en el contexto de demanda de gasolina puede ser aplicada con un enfoque operativo o estratégico. El análisis impulso-respuesta de un sistema que considera exclusivamente las ventas de las terminales nos permitirá evaluar el impacto de un shock en la demanda en una determinada TAR, tanto sobre ella misma como hacia las demás. Esto representa un valioso insumo para la planeación de una logística con niveles de inventario mínimos que garantice el abasto de combustible ante cualquier contingencia.

En otro orden de ideas, pensemos que un VAR con variables exógenas puede ser representado como un VAR convencional restringido a que los coeficientes de las variables exógenas sean iguales a cero. En este sentido, los modelos desarrollados para las entidades pueden ser modificados para visualizar el efecto de un shock, tanto en magnitud como en persistencia, en el precio, actividad económica, parque vehicular y población, en la dinámica de la demanda hacia las terminales de almacenamiento y en conjunto hacia la demanda de gasolina de un estado en particular. Esto robustecería la toma de decisiones de los hacedores de política pública.

El anexo B incluye las funciones de impulso-respuesta para el modelo del estado de Nuevo León, el cual fue seleccionado por la facilidad en la interpretación<sup>5</sup>.

---

<sup>5</sup>Las variables dependientes del modelo desarrollado para NL son primeras diferencias.

# Capítulo 8

## Conclusiones

El presente trabajo de investigación estima 31 modelos dinámicos de demanda de gasolina a partir de un panel compuesto por datos mensuales para las 78 terminales de almacenamiento y reparto de PEMEX.

Los modelos fueron estimados usando métodos basados en series de tiempo, los cuales han sido usados extensamente debido a la precisión de los pronósticos generados a partir de ellos. No obstante, en todos los casos se exploró la inclusión de algunas de las variables exógenas más recurrentes en la literatura como lo son: el crecimiento poblacional, el aumento en el número de vehículos, el precio del combustible y la actividad económica, además de algunas combinaciones de ellas para el modelo nacional.

Aquellos modelos que admiten una parte estructural representan una minoría. Solamente 14 modelos introducen al menos un regresor exógeno. Tanto el precio real, como la variable proxy de actividad económica, coadyuvan en explicar la dinámica de la variable dependiente en ocho de los modelos desarrollados; mientras que el cambio en el parque vehicular y el crecimiento poblacional sólo fueron incluidos en tres y dos casos, respectivamente.

Es importante no perder de vista que el mercado mexicano de los energéticos y particularmente el de los combustibles fósiles está sufriendo una transformación sin precedentes y, en consecuencia, las predicciones de los modelos especificados pueden subestimar las reacciones de un mercado en donde el precio era artificialmente bajo.

Por otro lado, la apertura del sector trae consigo una mayor disponibilidad de información

que facilitará el desarrollo de modelos estructurales que puedan ser calibrados con datos que reflejen la nueva dinámica del proceso de demanda. Una mayor precisión en los pronósticos sentará bases sólidas para el diseño de políticas públicas encaminadas a la reducción de emisiones y el desarrollo eficiente de infraestructura logística que impulse el crecimiento económico.

# Anexos



# Anexo A

## Coeficientes estimados para cada estado y medidas de la precisión en la estimación

Tabla A.1: Coeficientes estimados para la Ciudad de México

	D(D(LOG(TAR3)))	D(D(LOG(TAR4)))	D(D(LOG(TAR5)))
D(D(LOG(TAR3(-1))))	-1.808309 (0.14586) [-12.3973]	-0.525502 (0.19984) [-2.62966]	-0.567869 (0.15671) [-3.62368]
D(D(LOG(TAR3(-2))))	-2.037800 (0.23826) [-8.55269]	-0.627566 (0.32643) [-1.92252]	-0.777627 (0.25598) [-3.03779]
D(D(LOG(TAR3(-3))))	-1.841467 (0.30479) [-6.04170]	-0.479828 (0.41758) [-1.14908]	-0.555957 (0.32746) [-1.69778]
D(D(LOG(TAR3(-4))))	-1.833082 (0.34887) [-5.25439]	-0.751102 (0.47796) [-1.57148]	-0.699567 (0.37481) [-1.86645]
D(D(LOG(TAR3(-5))))	-1.506816 (0.38015) [-3.96377]	-0.738899 (0.52081) [-1.41874]	-0.686624 (0.40842) [-1.68118]
D(D(LOG(TAR3(-6))))	-1.206249 (0.38126) [-3.16385]	-0.415412 (0.52234) [-0.79529]	-0.452620 (0.40961) [-1.10499]
D(D(LOG(TAR3(-7))))	-1.237024 (0.37351) [-3.31192]	-0.578156 (0.51172) [-1.12984]	-0.549104 (0.40128) [-1.36836]

Continúa

Continuación			
	D(D(LOG(TAR3)))	D(D(LOG(TAR4)))	D(D(LOG(TAR5)))
D(D(LOG(TAR3(-8))))	-1.297692 (0.35991) [-3.60558]	-0.817355 (0.49309) [-1.65762]	-0.919959 (0.38668) [-2.37913]
D(D(LOG(TAR3(-9))))	-1.217930 (0.32364) [-3.76320]	-0.838678 (0.44340) [-1.89147]	-1.047054 (0.34771) [-3.01127]
D(D(LOG(TAR3(-10))))	-1.155325 (0.28041) [-4.12015]	-0.648796 (0.38417) [-1.68883]	-0.792097 (0.30126) [-2.62925]
D(D(LOG(TAR3(-11))))	-0.938488 (0.23355) [-4.01834]	-0.537512 (0.31997) [-1.67987]	-0.571086 (0.25092) [-2.27596]
D(D(LOG(TAR3(-12))))	-0.045759 (0.15974) [-0.28646]	-0.039887 (0.21884) [-0.18226]	-0.077788 (0.17162) [-0.45327]
D(D(LOG(TAR4(-1))))	-0.038264 (0.09476) [-0.40378]	-1.132918 (0.12983) [-8.72610]	-0.076451 (0.10181) [-0.75090]
D(D(LOG(TAR4(-2))))	0.007246 (0.14048) [ 0.05158]	-1.151799 (0.19247) [-5.98438]	-0.091627 (0.15093) [-0.60708]
D(D(LOG(TAR4(-3))))	0.048758 (0.18372) [ 0.26540]	-0.858070 (0.25170) [-3.40912]	-0.028719 (0.19738) [-0.14550]
D(D(LOG(TAR4(-4))))	0.002679 (0.20269) [ 0.01322]	-0.999943 (0.27770) [-3.60086]	-0.030970 (0.21777) [-0.14222]
D(D(LOG(TAR4(-5))))	-0.021164 (0.22244) [-0.09515]	-0.911195 (0.30474) [-2.99003]	-0.011459 (0.23898) [-0.04795]
D(D(LOG(TAR4(-6))))	0.085337 (0.23673) [ 0.36048]	-1.001399 (0.32433) [-3.08763]	0.008609 (0.25433) [ 0.03385]
D(D(LOG(TAR4(-7))))	0.124513 (0.24419) [ 0.50991]	-0.854003 (0.33454) [-2.55276]	-0.001495 (0.26235) [-0.00570]
D(D(LOG(TAR4(-8))))	0.009807 (0.24650) [ 0.03978]	-0.754713 (0.33772) [-2.23476]	0.019229 (0.26483) [ 0.07261]
D(D(LOG(TAR4(-9))))	-0.008099 (0.23385) [-0.03463]	-0.394741 (0.32038) [-1.23209]	-0.033383 (0.25124) [-0.13287]

Continúa

Continuación			
	D(D(LOG(TAR3)))	D(D(LOG(TAR4)))	D(D(LOG(TAR5)))
D(D(LOG(TAR4(-10))))	-0.100730 (0.21346) [-0.47189]	-0.412054 (0.29245) [-1.40898]	-0.143490 (0.22934) [-0.62567]
D(D(LOG(TAR4(-11))))	-0.191618 (0.17128) [-1.11871]	-0.179074 (0.23466) [-0.76311]	-0.119387 (0.18402) [-0.64876]
D(D(LOG(TAR4(-12))))	-0.267117 (0.13785) [-1.93773]	-0.129857 (0.18886) [-0.68759]	-0.185422 (0.14810) [-1.25199]
D(D(LOG(TAR5(-1))))	0.188164 (0.13689) [ 1.37456]	-0.075190 (0.18754) [-0.40092]	-1.256577 (0.14707) [-8.54403]
D(D(LOG(TAR5(-2))))	0.256200 (0.22947) [ 1.11647]	-0.259391 (0.31439) [-0.82507]	-1.364856 (0.24654) [-5.53604]
D(D(LOG(TAR5(-3))))	0.185069 (0.30029) [ 0.61630]	-0.731153 (0.41141) [-1.77720]	-1.481391 (0.32262) [-4.59169]
D(D(LOG(TAR5(-4))))	0.321345 (0.35894) [ 0.89527]	-0.435035 (0.49176) [-0.88466]	-1.244371 (0.38563) [-3.22684]
D(D(LOG(TAR5(-5))))	0.118327 (0.39716) [ 0.29794]	-0.444568 (0.54412) [-0.81704]	-1.177777 (0.42669) [-2.76024]
D(D(LOG(TAR5(-6))))	-0.127267 (0.41335) [-0.30789]	-0.364248 (0.56630) [-0.64320]	-1.096803 (0.44409) [-2.46977]
D(D(LOG(TAR5(-7))))	-0.090938 (0.41399) [-0.21966]	-0.194281 (0.56718) [-0.34254]	-0.933548 (0.44478) [-2.09892]
D(D(LOG(TAR5(-8))))	0.215830 (0.39820) [ 0.54202]	0.120951 (0.54554) [ 0.22171]	-0.502339 (0.42781) [-1.17420]
D(D(LOG(TAR5(-9))))	0.391726 (0.36117) [ 1.08462]	0.112374 (0.49481) [ 0.22711]	-0.157763 (0.38803) [-0.40658]
D(D(LOG(TAR5(-10))))	0.597999 (0.29772) [ 2.00857]	0.177696 (0.40789) [ 0.43565]	-0.060595 (0.31987) [-0.18944]
D(D(LOG(TAR5(-11))))	0.630528 (0.21412) [ 2.94468]	0.105911 (0.29336) [ 0.36103]	0.016708 (0.23005) [ 0.07263]

Continúa

Continuación			
	D(D(LOG(TAR3)))	D(D(LOG(TAR4)))	D(D(LOG(TAR5)))
D(D(LOG(TAR5(-12))))	0.211983 (0.13695) [ 1.54784]	-0.034992 (0.18763) [-0.18649]	-0.007570 (0.14714) [-0.05145]
R-squared	0.901500	0.859286	0.905887
Adj. R-squared	0.872284	0.817549	0.877972
Sum sq. resids	0.097804	0.183577	0.112893
S.E. equation	0.028790	0.039443	0.030931
F-statistic	30.85620	20.58807	32.45183
Log likelihood	348.3376	299.8533	337.2904
Akaike AIC	-4.056333	-3.426666	-3.912862
Schwarz SC	-3.346396	-2.716729	-3.202925
Mean dependent	6.65E-05	-0.000214	-0.000531
S.D. dependent	0.080559	0.092341	0.088545
Determinant resid covariance (dof adj.)		2.47E-10	
Determinant resid covariance		1.11E-10	
Log likelihood		1109.475	
Akaike information criterion		-13.00617	
Schwarz criterion		-10.87636	

Fin de la tabla

Nota: Errores estándar en paréntesis y estadísticos t en corchetes

Tabla A.2: RMSE asociado a las predicciones generadas

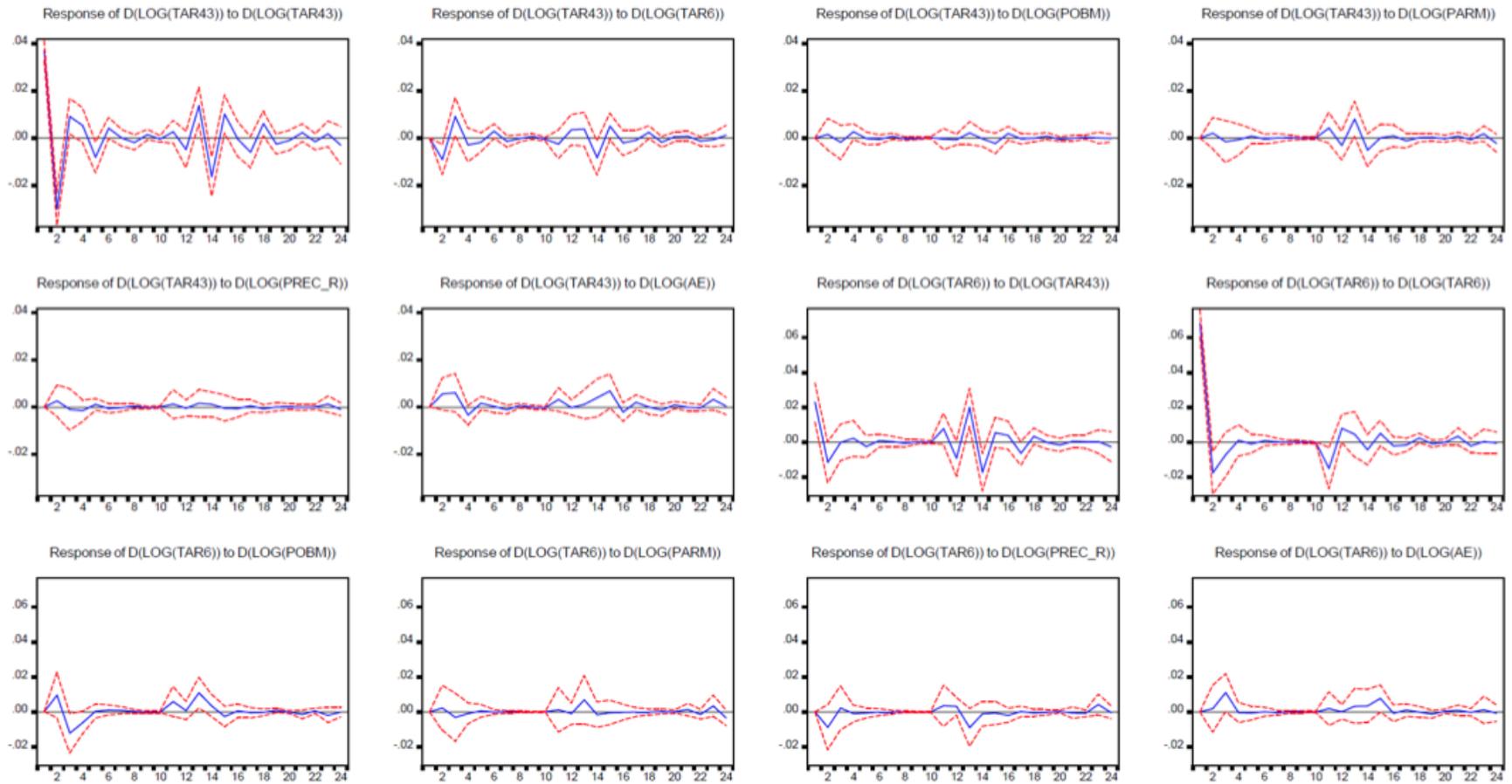
Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
D(D(LOG(TAR3)))	22	0.032070	0.025499	75.13705	0.167078
D(D(LOG(TAR4)))	22	0.06433	0.048822	161.2024	0.267473
D(D(LOG(TAR5)))	22	0.04278	0.029874	129.1279	0.200667

Los resultados obtenidos para los siguientes estados están a disposición del lector en la siguiente liga: <https://bit.ly/2kpngs3>

## Anexo B

### Función impulso respuesta (IRF) para el estado de Nuevo León

Figura B.1: Response to Cholesky One S.D. (d.f. adjusted) Innovations - 2 S.E.



# Bibliografía

- Arévalo, J. L., Herreros, Ó. P., et al. (2015). The uneven impact of the economic crisis of 2008-2009 in the labor markets of the regions of Mexico: the northern border versus the southern region. *Contaduría y Administración*, 60(6):195–218.
- Barassi, M. R. and Zhao, Y. (2017). Combination forecasting of energy demand in the UK.
- Bhattacharyya, S. C. and Timilsina, G. R. (2009). Energy demand models for policy formulation: a comparative study of energy demand models.
- Blinder, A. S. (2013). *Economic policy and the great stagflation*. Elsevier.
- Brons, M., Nijkamp, P., Pels, E., and Rietveld, P. (2006). A meta-analysis of the price elasticity of gasoline demand. a system of equations approach. Technical report, Tinbergen Institute Discussion Paper.
- Brons, M., Nijkamp, P., Pels, E., and Rietveld, P. (2008). A meta-analysis of the price elasticity of gasoline demand. A SUR approach. *Energy Economics*, 30(5):2105–2122.
- Chai, J., Wang, S., Wang, S., and Guo, J. (2012). Demand forecast of petroleum product consumption in the Chinese transportation industry. *Energies*, 5(3):577–598.
- Chris, B. (2008). Introductory econometrics for finance. *Cambridge, Cambridge University*.
- Crotte, A., Noland, R. B., and Graham, D. J. (2010). An analysis of gasoline demand elasticities at the national and local levels in Mexico. *Energy Policy*, 38(8):4445–4456.
- Espey, M. (1998). Gasoline demand revisited: an international meta-analysis of elasticities. *Energy Economics*, 20(3):273–295.
- Fullerton Jr, T. M., Salazar, J. A. I., Elizalde, M., et al. (2015). Microeconomic gasoline consumption anomalies in Mexico: 1997-2007. *Asian Economic and Financial Review*, 5(4):709–722.
- Graham, D. J. and Glaister, S. (2002). The demand for automobile fuel: a survey of elasticities. *Journal of Transport Economics and policy*, pages 1–25.
- Hanly, M., Dargay, J., and Goodwin, P. (2002). Review of income and price elasticities in the demand for road traffic. *Department for Transport, London*.

- Havranek, T., Irsova, Z., and Janda, K. (2012). Demand for gasoline is more price-inelastic than commonly thought. *Energy Economics*, 34(1):201–207.
- Hong, W.-C. (2013). *Intelligent energy demand forecasting*, volume 10. Springer Science & Business Media.
- Karathodorou, N., Graham, D. J., and Noland, R. B. (2010). Estimating the effect of urban density on fuel demand. *Energy Economics*, 32(1):86–92.
- Kongcharoen, C. and Kruangpradit, T. (2013). Autoregressive integrated moving average with explanatory variable (arimax) model for Thailand export. In *33rd International Symposium on Forecasting, South Korea*, pages 1–8.
- Labandeira, X., Labeaga, J. M., and López-Otero, X. (2017). A meta-analysis on the price elasticity of energy demand. *Energy Policy*, 102:549–568.
- Li, Z., Rose, J. M., and Hensher, D. A. (2010). Forecasting automobile petrol demand in Australia: an evaluation of empirical models. *Transportation Research Part A: Policy and Practice*, 44(1):16–38.
- Lutkepohl, H. (2004). Forecasting with VARMA models/. *European University Institute, Economics Working Papers*, (25):43.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- Mayr, J. and Ulbricht, D. (2007). Log versus level in VAR forecasting: 16 Million empirical answers-expect the unexpected. Technical report, Ifo Working Paper.
- Ocampo, S. and Rodríguez, N. (2012). An introductory review of a structural VAR-X estimation and applications. *Revista Colombiana de Estadística*, 35(3):479–508.
- Proietti, T. (2002). Forecasting with structural time series models. *A companion to economic forecasting*, 105132.
- Reyes, O., Escalante, R., and Matas, A. (2010). La demanda de gasolinas en México: Efectos y alternativas ante el cambio climático. *Economía: teoría y práctica*, (32):83–111.
- Sánchez, A., Islas, S., and Sheinbaum, C. (2015). Demanda de gasolina y la heterogeneidad en los ingresos de los hogares en México. *Investigación económica*, 74(291):117–143.
- Secretaría de Economía (2017a). Diagnóstico de la industria de los petrolíferos 2017.
- Secretaría de Economía (2017b). Prospectiva de petróleo crudo y petrolíferos 2017-2031.
- Shojaie, A. A., Zand, A. D., and Vafaie, S. (2017). Calculating production by using short term demand forecasting models: a case study of fuel supply system. *Evolving Systems*, 8(4):271–285.

- Suganthi, L. and Samuel, A. A. (2012). Energy models for demand forecasting—a review. *Renewable and sustainable energy reviews*, 16(2):1223–1240.
- Tarsitano, A. and Amerise, I. L. (2017). Short-term load forecasting using a two-stage sarimax model. *Energy*, 133:108–114.
- Urban, F., Benders, R., and Moll, H. (2007). Modelling energy systems for developing countries. *Energy Policy*, 35(6):3473–3482.
- Wirl, F. and Szirucsek, E. (1990). Energy modelling—a survey of related topics. *OPEC Energy Review*, 14(3):361–378.



# Índice de figuras

2.1. Demanda nacional de gasolinas HP (Mbd) . . . . .	7
2.2. Curvas de demanda (2002-2017) . . . . .	7
2.3. Tendencia en la demanda regional de gasolinas (Mbd) . . . . .	8
2.4. Crecimiento medio anual del consumo de gasolinas 2002-2016 (%) . . . . .	9
2.5. Consumo per cápita de gasolinas en 2016 . . . . .	10
2.6. Consumo per cápita de gasolinas (litros) . . . . .	11
2.7. Evolución 2003-2016 . . . . .	12
6.1. Transformaciones y proceso a modelar . . . . .	23
6.2. Correlograma . . . . .	23
6.3. Cálculo del estadístico de Ljung-Box . . . . .	25
6.4. Prueba de Jarque-Bera . . . . .	25
6.5. Ajuste en la predicción dinámica con respecto a los valores reales de la sub- muestra . . . . .	26
7.1. Contraste entre las ventas reales y las predicciones dinámicas del modelo . . . . .	30
7.2. Ventas reales y pronósticos calculados a partir del modelo . . . . .	34
7.3. Tendencia del consumo de gasolinas de la Ciudad de México . . . . .	35
B.1. Response to Cholesky One S.D. (d.f. adjusted) Innovations - 2 S.E. . . . .	48

# Índice de tablas

1.1. Participación y costo de los distintos medios de transporte a TAR en 2016 . . . . .	4
6.1. Coeficientes de la regresión de LS . . . . .	26
7.1. Valores estimados para los coeficientes de los rezagos y del efecto del cambio logarítmico en la variable proxy de actividad económica . . . . .	29
7.2. Valores estimados para los coeficientes del modelo . . . . .	32
7.3. RMSE asociado a las predicciones generadas por cada ecuación . . . . .	34
7.4. Prueba de exclusión de rezagos . . . . .	36

7.5. Medidas de precisión de la predicción . . . . .	37
A.1. Coeficientes estimados para la Ciudad de México . . . . .	43
A.2. RMSE asociado a las predicciones generadas . . . . .	46