

Pronóstico del PIB mediante series de Fourier con base en el IGAE. Una comparación con el análisis de series de tiempo

PIB forecast using Fourier series based on the IGAE. A comparison with time series analysis

Godfrey Orozco Lira^{1*}
Gabriel S. Cruz-Méndez^{1**}
¹ESE-IPN

Resumen

El componente cíclico es el término común para las series macroeconómicas, las fluctuaciones se pueden aproximar mediante combinaciones lineales de términos ondulatorios, esto es mediante series de Fourier que han resultado ser un buen ajuste para la serie del indicador global de la actividad económica (IGAE) en comparación con ciertos ajustes convencionales mediante términos autorregresivos y de medias móviles que son discutidos en este trabajo.

El análisis del modelo propuesto se simplifica toda vez que sus parámetros se obtienen por mínimos cuadrados. Luego, el PIB es estimado por el IGAE según la relación entre ellos a largo plazo. Sin embargo, la predicción para el PIB no ha resultado ser la mejor, por lo cual se sugiere considerar ambas técnicas de ajuste y combinar los resultados mediante un promedio, el resultado es en suma favorable ya que el error absoluto promedio es significativamente menor.

Palabras Clave: Error Absoluto, Estimación, IGAE, PIB, Series de Fourier, Proyección

Abstract

The cyclical component is the common term of macroeconomic series, fluctuations can be approached by linear combinations of sinusoidal terms, it means by Fourier series that have been result a right fit for the global indicator of the economic activity (IGAE, for its initials in spanish) compared with certain conventional adjustments by autoregressive and moving average terms argued in this work.

The proposed model analysis is easier because its parameters are obtained by least squares. Then, gross domestic product (PIB, for its initial in spanish) is estimated by the IGAE as they are related in the long run. Nevertheless, PIB predictions have not been the best, for this reason it is suggested to consider both approaches and join the outcomes by an average value, the result is favorable as the absolute error is significantly less.

Keywords: Absolute Error, Estimation, Fourier Series, IGAE, PIB, Projection

*Doctor en Ciencias Económicas. Departamento de Métodos Cuantitativos, ESE-IPN.

**Maestro en Ciencias Económicas. Departamento de Métodos Cuantitativos, ESE-IPN .

I. INTRODUCCIÓN

En el contexto económico siempre es de importancia fundamental el producto interno bruto (PIB) ya que mide la producción de bienes y servicios finales generados por los agentes económicos, además de que tradicionalmente funciona como un indicador de crecimiento y desarrollo económico, como indica Amielski (2020), “por más de 50 años (casi 70 a la fecha) el PIB sea utilizado como base para medir la prosperidad económica de las naciones”. Sin cuestionar o no que realmente sea un indicador de bienestar, en principio representa un estado que nos permite tomar decisiones en materia de producción, consumo e inversión

De acuerdo con Elizondo (2019), el indicador global de la actividad económica (IGAE) es una variable que se utiliza comúnmente como una aproximación del PIB. El ejercicio de obtener una mejor aproximación es importante porque permitirá tomar decisiones con fundamentos más cercanos a la realidad económica actual, no con base en la condición económica más reciente, indicada por el PIB, que se publica cada trimestre y a 50 días naturales después del trimestre de referencia¹.

Lo más frecuente es realizar estimaciones del PIB mensual, ya que la mayoría de los indicadores que le contribuyen se reportan mensualmente. Por ejemplo, Dynan (2020), propone que el consumo es la parte que explica más significativamente las tendencias del PIB ya que a nivel gráfico la dinámica es semejante, en México se podría considerar el indicador mensual de consumo privado en el mercado interior (IMCPMI).

Sin embargo, existen diversos enfoques y metodologías, aunque tienen el mismo propósito.

Tal como lo indican Cuche y Hess (1999), en donde se hace uso del filtro de Kalman para estimar el PIB mensual [otro caso es Elizondo (2019)] indicando que sin restricciones en los términos de error se obtienen mejores resultados para interpolar los datos trimestrales comparado con otros modelos propuestos bajo el mismo esquema y modelos de regresión lineal clásica. En ambos trabajos se considera muy restrictiva la metodología toda vez que el valor trimestral se asume como una variable dependiente de sus precedentes mensuales dentro del trimestre de referencia², de hecho, en el caso de

Cuche y Hess (1999) se indica que la suma de los valores mensuales en el trimestre corresponde al valor trimestral. Por otro lado, se propone que el índice de producción industrial es el que impulsa al PIB, en México se dispone del indicador mensual de la actividad industrial (IMAI).

Otro caso es el de generar un indicador mensual del PIB propuesto por Salazar, Weale, y Wright (1997), la estimación se realiza con base en los porcentajes de contribución de cada sector y estimando mensualmente las cantidades en cada uno (excepto el sector industrial que sí está disponible mensualmente). Se supone que el PIB mensual depende del indicador realizado y en consecuencia se obtienen valores mensuales para el PIB, se señalan ajustes que permiten hacer coincidir los valores estimados en el trimestre con los realizados.

Más reciente es el enfoque que se propone en este trabajo, un ejemplo similar es el desarrollo mostrado en Hassani, Rua, Silva, y Thomakos (2019), donde se ejerce el análisis de espectro singular multivariado para emplear series de distintas frecuencias, pero relacionadas entre sí, en este caso el PIB que se dispone trimestralmente y que está relacionado con el índice de producción industrial que se reporta mensualmente. La idea central es identificar las frecuencias múltiples en cada serie. A diferencia de los trabajos mencionados en Hassani, Rua, Silva, y Thomakos (2019), se asume que el valor trimestral corresponde al de la serie original. Se indica que el método es altamente confiable debido a su mejoramiento en términos del cuadro del error de predicción promedio y el error de predicción absoluto promedio.

En otro enfoque se propone estimar la tasa de crecimiento económico, como se discute en Hausmann, y otros (2013), por medio del PIB que se deriva con base en un indicador denominado “de complejidad económica” definido como la cantidad de recursos humanos y naturales para la producción de mercancías en un país, le llaman “producción con base en el conocimiento”. La estimación del crecimiento del PIB anual durante el periodo 2008-2020 señala que México crecerá a un 4.56% anual promedio.

El modelo que se propone como ajuste para la serie del IGAE está desarrollado con base en el

¹ Según la metodología de cálculo indicada por el INEGI en https://www.inegi.org.mx/contenidos/programas/pib/2013/doc/met_pibt.pdf

² Por ejemplo, el primer valor trimestral depende de Enero, Febrero y Marzo mismo.

análisis de Fourier o series de Fourier, como será indicado en todo lo que resta del documento, ya que al observar las primeras diferencias de la serie se puede apreciar un comportamiento cíclico o de componentes ondulatorios que es precisamente donde se involucran los términos que dependen de las funciones trigonométricas *seno* y *coseno*, en otras palabras, las fluctuaciones de la serie se pueden comparar con expresiones ondulatorias.

II. METODOLOGÍA

La idea principal del trabajo es que al comprobar que el PIB y el IGAE están relacionados en el largo plazo, después de realizar estimaciones del IGAE, se logrará estimar el PIB por medio de los datos correspondientes a este último por trimestre correspondiente, es decir que si se realiza una estimación anual del IGAE sólo los datos correspondientes a los cuatro trimestres del año serán utilizados para la predicción del PIB.

Las variables económicas se han transformado mediante el logaritmo natural, obteniendo así el logaritmo del PIB (LPIB) y el logaritmo del IGAE (LIGAE).

En las figuras 1 y 2 se muestra la tendencia de cada una, para el caso de LIGAE se ha tomado la información trimestral que corresponde a la serie del LPIB en que coinciden, lo cual sucede desde el primer trimestre del año 1993 y finalizan en el último trimestre del año 2018.

Se puede apreciar que dichas series comparten una tendencia en común, más aún, observando la evolución en cada una es posible señalar que los movimientos tanto a la baja como al alza son semejantes. Se comienza afirmando con base en la evidencia gráfica que las series están cointegradas³, es decir, se considera que

$$LPIB_t = \beta_1 + \beta_2 LIGAE_t + e_t \quad (1)$$

Para comprobar la afirmación anterior se ha realizado una prueba de raíz unitaria sobre la primera diferencia de los residuos (también: términos de error) e_t en el modelo (1), es decir que se ha realizado una regresión de mínimos cuadrados ordinarios sobre el siguiente modelo

$$\Delta e_t = \delta e_t + \varepsilon_t \quad (2)$$

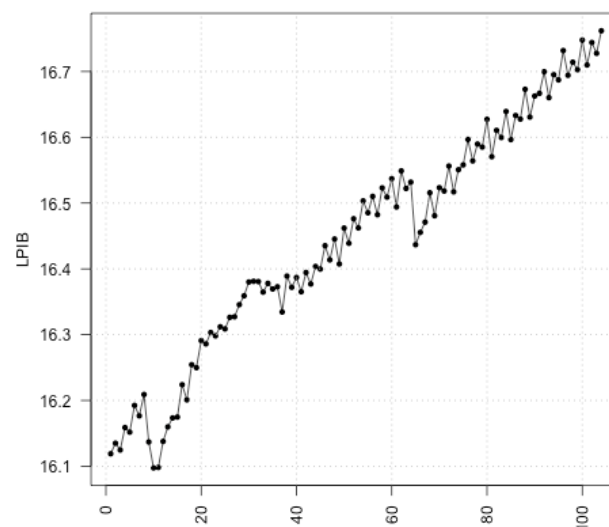
Los resultados son: $\hat{\delta} = -0.520565$ (0.0869), $d = 2.3$, $R^2 = 0.26$, y $\tau =$

-5.98 cuyos valores críticos (τ_c) son -2.6 al 1% y -1.95 al 5%, en ambos casos $|\tau| > \tau_c$ y por lo tanto es posible rechazar la hipótesis nula ($\delta = 0$) lo cual significa que la serie de los residuos es estacionaria y por lo tanto se concluye que las series LPIB y LIGAE están cointegradas.

De acuerdo con el resultado del párrafo anterior existe una relación de largo plazo entre las variables de interés, de manera que el PIB se puede explicar con base en el IGAE. Lo que sigue es estimar esta última serie, y generar valores predictivos para el PIB. El supuesto fundamental, como ya se mencionó, es que los datos trimestrales del IGAE explican el PIB, así es que al estimar la serie del IGAE simplemente se tomarán los valores trimestrales para predecir el PIB.

Figura 1

Serie LPIB



³ Todas las pruebas econométricas que serán indicadas en este trabajo se han realizado siguiendo la notación y formulaciones discutidas en Gujarati y Porter (2009).

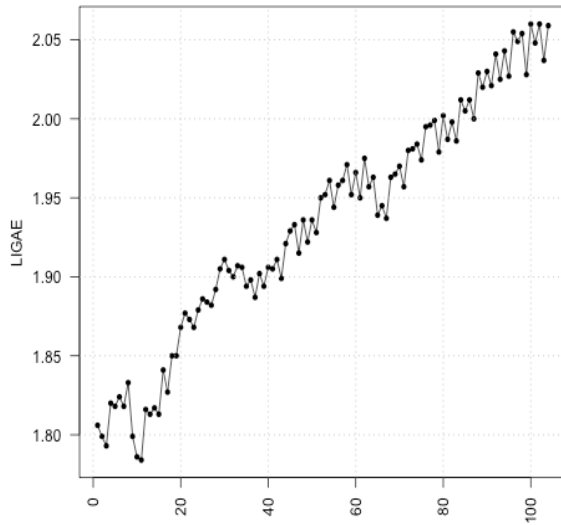


Fig. 2. Serie LIGAE datos trimestrales.

A. Estimación de LIGAE

El modelo se ha desarrollado sobre la serie completa del IGAE, cuya periodicidad es mensual, comienza en enero de 1990 y termina en diciembre de 2018.

A nivel gráfico se puede apreciar que la serie LIGAE no es estacionaria, de hecho, al realizar la prueba de raíz unitaria sin intercepto, es decir $\Delta LIGAE_t = \delta LIGAE_t + \varepsilon_t$ se obtiene que $\hat{\delta} = 4.36 \times 10^{-4}$ (prácticamente cero) y $\tau = -1.1324$, consecuentemente se puede aceptar la hipótesis nula ($\delta = 0$) ya que τ_c es -1.95 al 5%, es decir que la serie tiene raíz unitaria. Con intercepto $\Delta LIGAE_t = \alpha + \delta LIGAE_t + \varepsilon_t$ se obtiene $\delta < 0$ y $\tau = -1.6931$, la conclusión es la misma. Por lo tanto, después de todo lo anterior, se considera que LIGAE es integrada de orden 1.

Para fortalecer aún más el argumento, se realizó una prueba de raíz unitaria sobre la serie $\Delta LIGAE_t$, los resultados sin intercepto son $\delta < 0$ y $\tau = -24.1147$ que en valor absoluto es mayor que τ_c al 1% (-2.58), por lo cual se puede rechazar la hipótesis nula ($\delta = 0$). Con intercepto se obtienen $\delta < 0$ y $\tau = -24.2381$ por lo que nuevamente se puede rechazar la existencia de raíz unitaria. Es decir que, en efecto, LIGAE es integrada de orden 1.

En la figura 3 se muestra la serie de primeras diferencias de LIGAE (DLIGAE). Claramente se observa un comportamiento estacionario. Pero, más aún, una serie de tal comportamiento se puede considerar a simple vista cíclica, con periodos

aproximadamente uniformes. Esto es la razón principal por la que se propone ajustar los datos mediante series de Fourier. Parece haber un comportamiento regular de frecuencias, el objetivo es determinarlas.

De acuerdo al comportamiento observado se considera que la serie tiene dos frecuencias distintas⁴, lo cual se ve reflejado en la siguiente ecuación

$$\Delta LIGAE_t = \beta_0 + \beta_1 \cos(2\pi f_1 t) + \beta_2 \sin(2\pi f_1 t) + \beta_3 \cos(2\pi f_2 t) + \beta_4 \sin(2\pi f_2 t) \quad (3)$$

El modelo propuesto fue diseñado con base en la teoría de series de Fourier discutida en Bloomfield (2000) donde se indica, entre otras técnicas, que una forma de averiguar la frecuencia de una serie es mediante la maximización de la suma de los

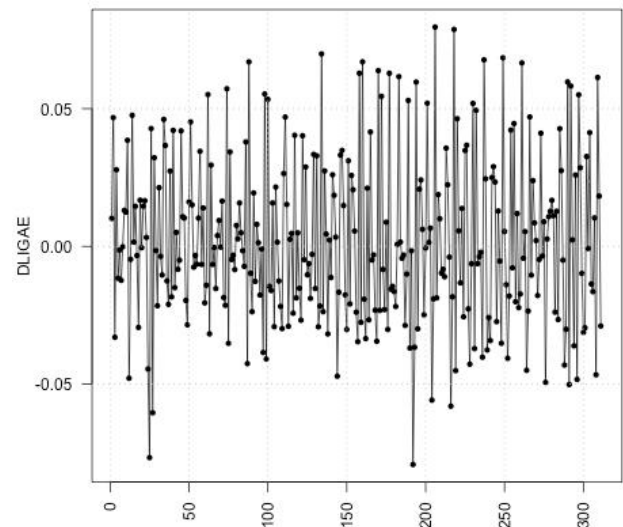


Fig. 3. Serie de primeras diferencias de LIGAE.

cuadrados que involucran a f_1 y f_2 (el problema dual corresponde a la minimización de los cuadrados de los residuos). El algoritmo utilizado para optimizar la suma de cuadrados indicada fue propuesto en Bloomfield (2000), tomando como frecuencias iniciales $f_1 = 0.415$ y $f_2 = 0.165$, que fueron seleccionadas a partir de los máximos de la gráfica ilustrada en la figura 4, el comportamiento se repite por cada intervalo separado en cantidades enteras debido a que las funciones trigonométricas son periódicas.

Sin embargo no es la combinación de frecuencias correspondientes a los picos de la gráfica en la figura 4 la que aporta el máximo. Mediante un ciclo de

⁴ Las frecuencias se pueden interpretar como ciclos, en este caso estamos asumiendo que hay un ciclo anual (f_1) y que durante el año hay diversos ciclos (f_2).

programación se utilizó el logaritmo de optimización indicado en el párrafo anterior con poco menos de 1,000 combinaciones de frecuencias en el intervalo (0,1). Los valores óptimos obtenidos son

$$f_1^* = 41.6556 \times 10^{-2} \text{ y } f_2^* = 16.6668 \times 10^{-2}$$

para los cuales el valor mínimo de la suma de los cuadrados de los residuos es 13.7289×10^{-2} .

Por todo lo anterior, ya determinados f_1 y f_2 , se puede afirmar que el modelo propuesto en (3) es lineal y por lo tanto los parámetros se estimaron con el método de mínimos cuadrados ordinarios que se resumen en la tabla 1. En la cual es posible apreciar que la mayoría de los parámetros son estadísticamente significativos al 95%, excepto por β_0 que con base en la evidencia mostrada se puede aceptar la hipótesis nula de que $\beta_0 = 0$. Sin embargo, con base en el estadístico F que es significativo al 99%, se puede rechazar la hipótesis nula de que todos los coeficientes sean cero.

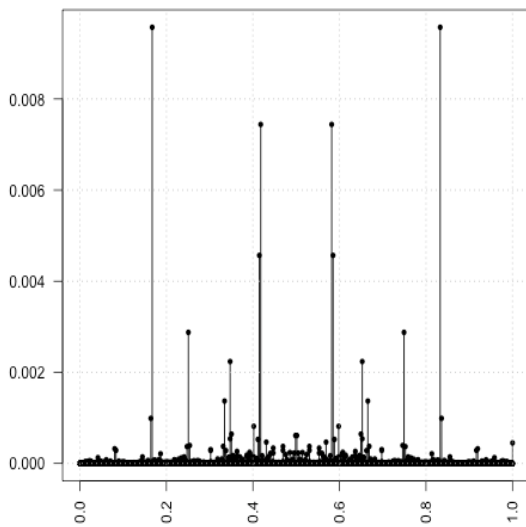


Fig. 4. Suma de cuadrados asociados con f_i para $i = 1, 2$.

Tabla 1
MÍNIMOS CUADRADOS PARA LOS PARÁMETROS DEL MODELO (3)

	Estimación	Error. Est.	t	Prob.
β_0	0.001999	0.001201	1.66	0.0971
β_1	0.012288	0.001699	7.23	0.0000
β_2	0.019750	0.001698	11.63	0.0000
β_3	-0.006083	0.001697	-3.58	0.0004
β_4	-0.018896	0.001700	-11.12	0.0000
$R^2 = 0.514429$			Desv. Est. = 0.02118	
$F = 81.04651$			Durbin-Watson = 2.653	

Aunado a todo lo anterior se realizó la regresión sin el parámetro β_0 y los resultados son similares, todos los coeficientes son significativos al mismo nivel y la desviación estándar de la regresión obtenida es 0.02124 (mayor que antes), ya que la variación es poco significativa se decidió que el modelo (3) es suficientemente representativo⁵.

B. Prueba de normalidad

Para el modelo propuesto en (3) se ha realizado la prueba de normalidad de Jarque-Bera (JB), como se muestra en la figura 5. El valor crítico obtenido es $JB = 1.0946$, la probabilidad de obtener un valor Ji cuadrada mayor a 1.0946 es 0.5785 que se considera razonablemente alta y por lo tanto se concluye que los residuos siguen la distribución normal estándar.

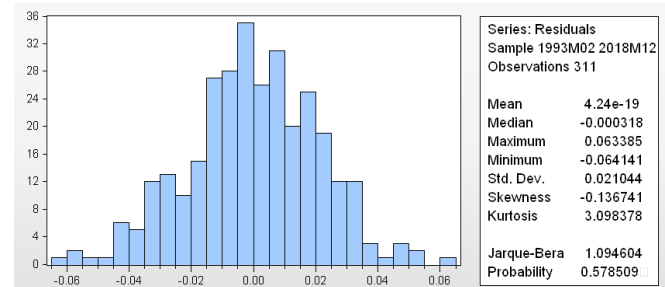


Fig. 5. Prueba de normalidad de Jarque-Bera para los residuos del modelo (3)

C. Corrección del modelo

Hay evidencia de autocorrelación mediante el modelo propuesto en (3). Esto se ha comprobado en primer lugar mediante la prueba de Durbin-Watson, considerando que los límites inferior y superior son,

⁵ Entre otras cantidades que fortalecen el argumento, por ejemplo el criterio de Akaike y Schwartz que indican una mejoría tan solo por remover variables independientes, sin embargo el R^2 es, hasta dos cifras decimales, idéntico en ambos modelos (ligeramente mayor con β_0).

respectivamente $D_L = 1.728$ y $D_U = 1.810$, se observa que para $d = 2.653$ (tabla 1) $4 - D_L < d < 4$ por lo cual se puede rechazar la hipótesis nula, en otras palabras se acepta que hay autocorrelación negativa. Más aún, la misma conclusión se obtiene mediante la prueba modificada en que $4 - d < D_U$.

Por otro lado, el argumento anterior se fortalece mediante la prueba de las rachas. En este caso el número total de observaciones es $N = 311$ compuestas por $N_1 = 155$ cantidades positivas, $N_2 = 156$ negativas y $R = 194$ rachas. Ante esta información se han calculado $E[R] = 156$ y $\delta^2 = 77$, de manera que el intervalo de confianza para el número de rachas obtenido al 95% es $[139, 174]$ y en consecuencia se puede rechazar la hipótesis nula (H_0 : los residuos no muestran autocorrelación), es decir que, hay evidencia de autocorrelación. Más aún, como el número de rachas está a poco más de 4 desviaciones estándar de la media (por lo que se considera grande), la autocorrelación es negativa.

El nivel de autocorrelación establecido es 2. En efecto, por medio de la prueba de Breusch-Godfrey (BF) el modelo más significativo para los residuos es (tomando como discriminante el criterio de Akaike y Schwartz). Los parámetros del modelo se encuentran en la tabla 2.

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t \quad (4)$$

Siguiendo el procedimiento indicado en [9] se obtuvo $R^2 = 0.2076$, $p = 2$ (longitud del rezago) y $n = 311$. Entonces, considerando que $Z = (n - p)R^2$ sigue la distribución *Ji cuadrada* con p grados de libertad, en este caso $(n - p)R^2 \approx 64.15$, se observa que $P[Z > 64.15] < 0.005$ (ya que $P[Z > 10.59] = 0.005$). En otras palabras, es prácticamente improbable obtener un valor *Ji cuadrada* mayor a 64.15 y en consecuencia al menos uno de los coeficientes en (4) es significativamente diferente de cero. Si se agregan más rezagos en (4) los demás coeficientes dejan de ser estadísticamente significativos.

Así es que, con base en todo lo anterior el modelo se corrigió de la siguiente manera:

Considere que

$$Y_t = \Delta LIGAE_t$$

$$X_{1t} = \cos(2\pi f_1 t)$$

$$X_{2t} = \sin(2\pi f_1 t)$$

$$X_{3t} = \cos(2\pi f_2 t)$$

$$X_{4t} = \sin(2\pi f_2 t)$$

Sustituyendo en (3) se obtiene

TABLA 2.
MÍNIMOS CUADRADOS PARA LOS PARÁMETROS DEL MODELO (4)

	Estimación	Error. Est.	t	Prob.
ρ_1	-0.450334	0.054036	-8.33	0.0000
ρ_2	-0.317987	0.053912	-5.89	0.0000

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} \quad (5)$$

De donde, para $i = 1, 2$

$$\rho_i Y_{t-i} = \rho_i \beta_0 + \rho_i \beta_1 X_{1t-i} + \rho_i \beta_2 X_{2t-i} + \rho_i \beta_3 X_{3t-i} + \rho_i \beta_4 X_{4t-i}$$

Luego, definiendo

$$Y_t^* = Y_t - \rho_1 Y_{t-1} - \rho_2 Y_{t-2}$$

$$X_{rt}^* = X_{rt} - \rho_1 X_{rt-1} - \rho_2 X_{rt-2} \text{ para } r = 1, 2, 3, 4$$

Es posible reescribir el modelo (5) de la siguiente forma

$$Y_t^* = a_0 + a_1 X_{1t}^* + a_2 X_{2t}^* + a_3 X_{3t}^* + a_4 X_{4t}^* + \varepsilon_t \quad (6)$$

Donde $a_0 = \beta_0(1 - \rho_1 - \rho_2)$, ε_t es el residuo de (4), mientras que ρ_1 y ρ_2 son los parámetros obtenidos en la prueba de autocorrelación de BF.

Así es que, ajustando por mínimos cuadrados los parámetros del modelo (6) se obtuvieron los resultados de la tabla 3, los cuales son favorables. En primer lugar, porque, estadísticamente dichos parámetros son estadísticamente significativos tanto a nivel individual como en conjunto. Además, con base en la prueba de Durbin-Watson, y considerando que $D_L = 1.728$ y $D_U = 1.810$, se observa que para $d = 1.983$ (tabla 3) $D_U < d < 4 - D_U$ por lo cual se puede aceptar la hipótesis nula, en otras palabras, es posible afirmar que no hay evidencia de autocorrelación. Sin embargo, el modelo no pasa la prueba de normalidad, lo cual no es grave ya que las estimaciones que se están indicando son puntales y, más aún, el histograma de frecuencias para los

residuos correspondientes sugiere que a pesar del sesgo desde una panorámica se puede apreciar que existe normalidad para la mayoría de los datos, en la figura 6 se muestra el resultado. Luego, comparando los modelos (3) y (6) por medio del coeficiente de determinación R^2 se concluye que (6) explica en mayor medida $\Delta LIGAE_t$.

TABLA 3.
MÍNIMOS CUADRADOS PARA LOS PARÁMETROS DEL MODELO (6)

	Estimación	Error. Est.	t	Prob.
a_0	0.003379	0.001071	3.16	0.0018
a_1	0.012133	0.001962	6.19	0.0000
a_2	0.019656	0.001969	9.98	0.0000
a_3	-0.005903	0.001205	-4.9	0.0000
a_4	-0.019083	0.001205	-15.84	0.0000
$R^2 = 0.575969$		Desv. Est. = 0.018818		
$F = 103.2321$		Durbin-Watson =1.983		

Es así que, con base en todos los argumentos anteriores, se ha decidido que el modelo (6) es el más adecuado para explicar las primeras diferencias de $LIGAE$ con lo cual es posible, después de realizar algunas operaciones algebraicas elementales, obtener la estimación del $LIGAE$, como se muestra enseguida.

Sustituyendo Y_t^* en (6) y dado que $Y_t = \Delta LIGAE_t$ se obtiene el modelo

$$LIGAE_t = (1 + \rho_1)LIGAE_{t-1} + (\rho_2 - \rho_1)LIGAE_{t-2} - \rho_2 LIGAE_{t-3} + a_0 + a_1 X_{1t}^* + a_2 X_{2t}^* + a_3 X_{3t}^* + a_4 X_{4t}^* + \varepsilon_t \quad (7)$$

Con el cual es posible obtener una estimación del $LIGAE$ utilizando la información de tres periodos (meses) anteriores. El objetivo es obtener los valores trimestrales del año para que mediante el modelo (1) se obtengan las estimaciones correspondientes al $LPIB$.

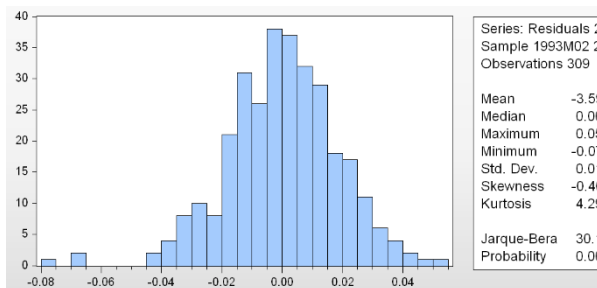


Fig. 6. Prueba de normalidad de Jarque-Bera para los residuos del modelo (6)

D. Estimación de $LPIB$

Los parámetros del modelo (1) se han ajustado por el método de los mínimos cuadrados, el resumen se muestra en la tabla 4. Los resultados, en cuanto a la significancia estadística de los parámetros tanto a nivel como individual como en conjunto, son sumamente favorables y, más aún, dado el coeficiente de determinación es alto se puede considerar este como un modelo apropiado para explicar los datos. Aunado a lo anterior, recuerde que también ya se mostró que las series $LPIB$ y $LIGAE$ están cointegradas, de manera que el modelo (1) es una relación significativa.

Así es que para obtener una estimación del PIB , se ejercerá el siguiente modelo

$$LPIB_t = 11.85 + 1.03LIGAE_t + e_t \quad (8)$$

Donde $LIGAE$ se obtiene mediante el modelo (7).

TABLA 4.
MÍNIMOS CUADRADOS PARA LOS PARÁMETROS DEL MODELO (1)

	Estimación	Error. Est.	t	Prob.
β_1	11.85027	0.621725	284.65	0.0000
β_2	1.032160	0.016407	110.46	0.0000
$R^2 = 0.991709$		Desv. Est. = 0.080081		
$F = 12200.51$		Durbin-Watson =1.68		

III. RESULTADOS

Al utilizar el modelo (7) para obtener ajustes de los valores correspondientes al $LIGAE$ se obtuvieron los siguientes resultados:

1. La suma de los cuadrados de los residuos es 0.10766
2. El error absoluto promedio es 0.01435
3. El error absoluto porcentual promedio es 0.32%

Por otro lado, mediante el modelo (1) se calcularon las estimaciones del $LPIB$ cuyos resultados son:

1. Suma de los cuadrados de los residuos = 0.02746
2. Error absoluto promedio = 0.01299
3. Error absoluto porcentual promedio = 0.08%

Recuerde que la muestra de los datos corresponde al periodo: enero de 1993 a diciembre de 2018. Los datos correspondientes al año 2019 se han utilizado para probar el modelo, y no se excluyeron antes ya que al iniciar la investigación aún no se tenían disponibles estos datos, en el apéndice se muestran los tabulados calculados.

Los resultados de las proyecciones del *LIGAE* son

- 1 Error absoluto promedio = 0.04869
- 2 Error absoluto porcentual promedio = 0.95%

Luego, se tomaron los valores correspondientes a los cuatro trimestres del año y mediante el modelo (1) se proyectaron los valores correspondientes al *LPIB* que también fueron comparados con los datos reales,

1. Error absoluto promedio = 0.04589
2. Error absoluto porcentual promedio = 0.27%

Finalmente, se calcularon los valores respectivos para el *PIB* y se compararon con los reales. Visto de esta manera el error absoluto promedio es de 868,712.77 millones de pesos. Y en porcentaje el error promedio es de 4.72%.

IV. DISCUSIÓN

En lo general, habrá que considerar el error indicado para los pronósticos del *LIGAE* y *LPIB* con base en los modelos (7) y (1), respectivamente, de manera que los resultados deberán indicarse mediante un intervalo y seleccionar el valor más representativo para la realidad según sean el entorno económico que en particular considere apropiadas el analista. A pesar de que en ambos casos se puede apreciar que el rango de error es razonablemente bajo, no debe ser ignorado.

Comparando los resultados estimados para el *LPIB*, por medio del modelo (1), con los resultados mostrados en Elizondo (2019), (donde el error absoluto promedio de las estimaciones es 0.05%), se puede concluir que el ajuste es apropiado en ambos casos, la diferencia es tan solo de 0.03%, sin embargo los modelos propuestos aquí son más simples en el contexto teórico y por lo tanto más accesibles para cualquiera que utilice el modelo (1) o que haga su propia versión del mismo.

Para refinar los puntos de referencia y comparación del modelo propuesto en este trabajo se han ajustado un par de modelos alternativos para estimar el *LIGAE* incorporando términos

autorregresivos y de medias móviles, uno de ellos mediante un esquema autorregresivo de *LIGAE* y el otro sobre las primeras diferencias (definidas como Y_t) los cuales están dados por:

$$\begin{aligned} LIGAE_t = & \alpha_0 LIGAE_{t-1} + \alpha_1 LIGAE_{t-2} \\ & + \alpha_2 LIGAE_{t-5} \\ & + \alpha_3 LIGAE_{t-10} + \alpha_4 LIGAE_{t-12} \\ & + \hat{\rho} \varepsilon_{t-13} + \varepsilon_{t-12} + \varepsilon_t \quad (9) \end{aligned}$$

$$\begin{aligned} Y_t = & \alpha_0 Y_{t-1} + \alpha_1 Y_{t-12} + \alpha_2 Y_{t-13} \\ & + \hat{\rho} \varepsilon_{t-13} + \varepsilon_{t-12} + \varepsilon_t \quad (10) \end{aligned}$$

Los parámetros del modelo (9) se muestran en la tabla 5 y los del modelo (10) en la tabla 6. En ambos casos se puede apreciar que la significancia estadística de los parámetros es relevante.

En el modelo (9) es posible que los residuos estén correlacionados, aunque se trata de una situación esperada debido a que *LIGAE* no es estacionaria, sin embargo, el coeficiente de determinación es alto. El error absoluto promedio del modelo es 0.03456, el cual es mayor que el correspondiente al modelo propuesto en (7). Se observa que hay dependencia significativa respecto a ciertos valores pasados, teniendo como longitud de rezago máxima 12, es decir que el *LIGAE* actual dependerá de ciertos valores dentro del año pasado.

Los resultados de las proyecciones son (ver apéndice)

Para el *LIGAE*

1. Error absoluto promedio = 0.02609
2. Error absoluto porcentual promedio = 0.55%

Para el *LPIB*

1. Error absoluto promedio = 0.01755
2. Error absoluto porcentual promedio = 0.10%

Por otro lado, también el modelo (10) es significativo en términos de sus parámetros y el coeficiente de determinación es mayor que en el modelo (7) y no hay evidencia de autocorrelación en los residuos. Sin embargo el error absoluto promedio del modelo es 0.02053, que también es mayor al que se obtuvo en el modelo (7). Observe que, como en el modelo (9), hay dependencia de ciertas cantidades correspondientes al año pasado.

Los resultados de las proyecciones son (ver apéndice)

Para el LIGAE

- 1 Error absoluto promedio = 0.01443
- 2 Error absoluto porcentual promedio = 0.31%

Para el LPIB

3. Error absoluto promedio = 0.01559
4. Error absoluto porcentual promedio = 0.09%

Entonces, derivado de los resultados mostrados para los modelos (9) y (10), en comparación con los resultados del modelo (7) se observa que los pronósticos son más adecuados mediante los modelos (9) y (10), más en (7) se observa un ajuste más adecuado a los datos, compare la suma de los cuadrados de los residuos y los errores absolutos en promedio. Así es que surge naturalmente la siguiente pregunta: ¿Cuál es el modelo que debería considerarse para obtener un buen pronóstico del PIB?

Quizá la mejor estrategia debería ser aquella en la que se haga un pronóstico con base en todos los modelos posibles, en este trabajo se han incorporado las series de Fourier pensando en los componentes cíclicos del crecimiento económico que cuantitativamente se observan mediante el PIB y el resultado ha sido favorable. Sin embargo, hay deficiencias en los pronósticos.

Si al resultado del pronóstico para el PIB se le resta/suma el error absoluto promedio se obtienen los límites inferior/superior del intervalo en que es probable observar el valor real del PIB. Resulta muy interesante que si se consideran únicamente los límites inferiores de aquellos intervalos, empleando cualquiera de los modelos discutidos (7), (9) o (10) los errores absolutos disminuyen en una cantidad significativa. Aunado a ello, si se consideran el promedio de los límites inferiores como una estimación entonces el error absoluto promedio del pronóstico disminuye sobremanera, resultando en un error absoluto porcentual promedio de 0.26%. Este asunto podría ser estudiado a detalle en un trabajo futuro, los resultados se muestran en el apéndice.

TABLA 5.
MÍNIMOS CUADRADOS PARA LOS PARÁMETROS DEL MODELO (9)

	Estimación	Error. Est.	t	Prob.
α_0	0.459360	0.046079	9.97	0.0000
α_1	0.196647	0.043270	4.55	0.0000
α_2	0.352536	0.052276	6.74	0.0000
α_3	-0.275328	0.055814	-4.93	0.0000
α_4	0.268149	0.043999	6.1	0.0000
$\hat{\rho}$	0.351214	0.060944	5.76	0.0000
$R^2 = 0.984844$		Desv. Est. = 0.020941		
$\sum \text{residuos}^2 = 0.128493$		Durbin-Watson = 1.761		

TABLA 6.
MÍNIMOS CUADRADOS PARA LOS PARÁMETROS DEL MODELO (10)

	Estimación	Error. Est.	t	Prob.
α_0	-0.422576	0.051537	-8.2	0.0000
α_1	1.001241	0.013252	75.55	0.0000
α_2	0.424133	0.050015	8.48	0.0000
$\hat{\rho}$	-0.789196	0.052065	-15.16	0.0000
$R^2 = 0.654846$		Desv. Est. = 0.017975		
$\sum \text{residuos}^2 = 0.086500$		Durbin-Watson = 2.222		

V. CONCLUSIONES

De acuerdo con los resultados se puede afirmar que el LPIB es una variable dependiente del LIGAE y que están relacionadas en el largo plazo, el periodo de estudio va de 1993 a 2018. Tómese en cuenta que ambas series son observadas en distintas periodicidades, la primera es trimestral mientras que la segunda es mensual, de manera que dicha relación se observó al incorporar en el modelo únicamente las observaciones trimestrales del LIGAE.

El primer paso para la estimación del PIB consiste en que se ajusten mediante un modelo la serie del LIGAE, cuyas primeras diferencias son estacionarias y dado que hay cierta periodicidad resultó conveniente ajustarla por medio de series de Fourier ya que, comparando con algunos modelos de ajuste usuales en la teoría de series de tiempo, el error absoluto promedio es mejor. Sin embargo, el modelo propuesto se queda por debajo respecto al coeficiente de determinación y la suma de los cuadrados del error (al menos para uno de los modelos con que se ha comparado). Son estos últimos indicadores los que al final podrían explicar por qué las predicciones son mejores mediante los modelos usuales.

APÉNDICE

RESULTADOS DEL MODELO (7)

PROYECCIÓN DEL LIGAE

Fecha	LIGAE pronosticado	LIGAE observado
ene/2019	4.751011868	4.72081339
feb/2019	4.799151006	4.68979293
mar/2019	4.773599487	4.72148978
abr/2019	4.756951582	4.69925848
may/2019	4.759222229	4.74670828
jun/2019	4.745903919	4.72584976
jul/2019	4.790106746	4.73003381
ago/2019	4.785911732	4.72817177
sep/2019	4.793717656	4.68810318
oct/2019	4.775573533	4.74617281
nov/2019	4.752858995	4.75416497
dic/2019	4.780444777	4.73225838

PROYECCIÓN DEL LPIB

Fecha	LIGAE pronosticado	LIGAE observado
mar/2019	16.7773884	16.7184749
jun/2019	16.7488022	16.7311455
sep/2019	16.7981536	16.7241116
dic/2019	16.7844539	16.7515066

INTERVALOS PRONOSTICADOS PARA EL PIB (MILLONES DE PESOS)

Fecha	Limite inferior	Limite superior
mar/2019	18,465,532.48	20,202,958.01
jun/2019	17,920,663.75	19,658,089.28
sep/2019	18,871,208.75	20,608,634.29
dic/2019	18,602,621.05	20,340,046.58

El ajuste mediante series de Fourier es más simple en el sentido técnico que realizar un análisis convencional de series de tiempo, ya que se evita el análisis de correlogramas para obtener los términos autorregresivos y de medias móviles, además de las pruebas de mínimos cuadrados que se deben realizar en cada paso y la pérdida de información por cada vez que la longitud del rezago aumenta.

Las series de Fourier ajustadas al modelo permiten analizar por medio de frecuencias el comportamiento de la serie y se reducen al análisis de un modelo de regresión lineal después de investigar aquellas frecuencias, en este caso para la serie de las primeras diferencias del LIGAE. Aún se puede realizar más para tratar de obtener mejores aproximaciones de las frecuencias, otra posibilidad es dividir la serie por periodos según sea el contexto y condiciones económicas que reflejan la evolución del PIB o incluso incorporar otras variables como el índice de la actividad industrial que no fue considerado en este trabajo. Se espera que en un proyecto futuro pueda ser abordado el mismo problema considerando estas últimas variables.

Al considerar mediante un criterio de selección empírico, no definido, se obtuvo un pronóstico más acertado en el que se ve involucrado el resultado de los tres modelos. En consecuencia, si se trabajan más modelos de diferente naturaleza, pero con el mismo objetivo, y se reúnen los resultados, parece ser que el pronóstico final es óptimo. Esto último se espera que pueda ser estudiado a detalle en un trabajo futuro.

RESULTADOS DEL MODELO (9)

PROYECCIÓN DEL LIGAE

Fecha	LIGAE pronosticado	LIGAE observado
ene/2019	4.74528158	4.72081339
feb/2019	4.71803632	4.68979293
mar/2019	4.72803112	4.72148978
abr/2019	4.74603458	4.69925848
may/2019	4.75895654	4.74670828
jun/2019	4.75845807	4.72584976
jul/2019	4.75892875	4.73003381
ago/2019	4.7524713	4.72817177
sep/2019	4.72998605	4.68810318
oct/2019	4.76113955	4.74617281
nov/2019	4.77123964	4.75416497
dic/2019	4.76736724	4.73225838

PROYECCIÓN DEL LPIB

Fecha	LIGAE pronosticado	LIGAE observado
mar/2019	16.7303546	16.7184749
jun/2019	16.7617601	16.7311455
sep/2019	16.7323724	16.7241116
dic/2019	16.7709558	16.7515066

INTERVALOS PRONOSTICADOS PARA EL PIB (MILLONES DE PESOS)

Fecha	Limite inferior	Limite superior
mar/2019	18,117,483.49	18,774,387.22
jun/2019	18,705,979.54	19,362,883.27
sep/2019	18,154,741.38	18,811,645.11
dic/2019	18,881,821.61	19,538,725.33

RESULTADOS DEL MODELO (10)

PROYECCIÓN DEL LIGAE

Fecha	LIGAE pronosticado	LIGAE observado
ene/2019	4.70952119	4.72081339
feb/2019	4.68300078	4.68979293
mar/2019	4.72678729	4.72148978
abr/2019	4.71483136	4.69925848
may/2019	4.75196524	4.74670828
jun/2019	4.74837175	4.72584976
jul/2019	4.73516068	4.73003381
ago/2019	4.73899175	4.72817177
sep/2019	4.70832042	4.68810318
oct/2019	4.7612337	4.74617281
nov/2019	4.77622062	4.75416497
dic/2019	4.76544927	4.73225838

PROYECCIÓN DEL LPIB

Fecha	LIGAE pronosticado	LIGAE observado
mar/2019	16.7290708	16.7184749
jun/2019	16.7513494	16.7311455
sep/2019	16.71001	16.7241116
dic/2019	16.7689761	16.7515066

INTERVALOS PRONOSTICADOS PARA EL PIB (MILLONES DE PESOS)

Fecha	Limite inferior	Limite superior
mar/2019	18,132,358.03	18,712,180.22
jun/2019	18,547,386.68	19,127,208.87
sep/2019	17,784,540.87	18,364,363.06
dic/2019	18,882,370.38	19,462,192.56

**PIB PROMEDIO PRONOSTICADO MEDIANTE LOS
LIMITES INFERIORES DE LOS INTERVALOS DERIVADOS
MEDIANTE LOS MODELOS (7), (9) Y (10)**

Fecha	Pronóstico	Observado	Error absoluto (%)
mar/2019	18,238,458.00	18,228,099.37	0.06
jun/2019	18,391,343.32	18,460,530.85	0.37
sep/2019	18,270,163.67	18,331,137.00	0.33
dic/2019	18,788,937.68	18,840,258.82	0.27

REFERENCIAS

- Amielski, M. (8 de Abril de 2020). *Pembina Institute*. Obtenido de www.jstor.org/stable/restrep00237
- Bloomfield, P. (2000). *Fourier Analysis of Time Series*. E.U.A.: John Wiley y Sons.
- Cuche, N. A., y Hess, M. K. (1999). CucheEstimating Monthly GDP In A General Kalman Filter Framework: Evidence From Switzerland. *Swiss National Bank, Study Center Gerzensee, Gerzensee*.
- Dynan, K. (9 de Abril de 2020). *GDP as a Measure of Economic Well-being*. Obtenido de Hutchins Center Working Paper #43: [4] Dynan, K. (2018). “GDP as a Measure of Economic Well-being”, Hutchins Cent<https://www.brookings.edu/research/gdp-as-a-measure-of-economic-well-being/>
- Elizondo, R. (2019). Estimaciones del PIB mensual en México basadas en el IGAE. *Estudios Económicos*, 197-241.
- Gujarati, D. N., y Porter, D. C. (2009). *Econometría*. Ciudad de México: McGraw-Hill.
- Hassani, H., R. A., Silva, E. S., y Thomakos, D. (2019). Monthly forecasting of GDP with mixed-frequency multivariate singular spectrum analysis. *International Journal of Forecasting*, 1263-1272.
- Hausmann, R., H. C., Bustos, S., Coscia, M., Simoes, A., y Yildirim, M. (2013). *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. Cambridge, Massachusetts: London, England: The MIT Press.
- Salazar, E. S., Weale, M., y Wright, S. (22 de Abril de 1997). *A Monthly Indicator of GDP*. Obtenido de National Institute Economic Review: www.jstor.org/stable/23872430