

TÍTULO TRABAJOJuan Jaramillo¹, Fernando Jaramillo² y Pedro Jaramillo³

Fecha de entrega: dd – mm – 2022

3.5 Pronósticos para la validación cruzada

Para cada modelo ajustado presente la ecuación general de los pronósticos $\hat{Y}_n(L)$ usando los valores estimados para los parámetros, y proporcione los pronósticos puntuales para la validación cruzada junto con sus intervalos del 95% de confianza, debidamente tabulados y fechados. Interprete las cifras obtenidas según los datos, calcule y compare las medidas de precisión de los pronósticos puntuales (MAE, MAPE y RMSE de pronóstico) y las medidas de precisión para los intervalos de pronóstico (amplitud media y cobertura alcanzada); Presente una gráfica comparativa de los pronósticos y valores reales dejados para la validación cruzada. Dé una conclusión respecto a cuál modelo pronostica mejor para el horizonte de pronóstico seleccionado, teniendo en cuenta tanto la precisión de los pronósticos puntuales como la de sus intervalos de predicción.

A continuación, se verán los pronósticos realizados para el periodo ex post para llevar a cabo validación cruzada, lo cual se hace teniendo en cuenta que el origen ocurre en $n = 239$. Además, para los intervalos de pronóstico se usará una confianza del 95 % y se presentará la gráfica de los pronósticos contra los datos reales en los tiempos de pronóstico ex post.

Tabla 2. Pronósticos puntuales y por I.P del 95% de confianza

	Modelo 1			Modelo 2			Modelo 3			Modelo 4		
Período	Pronóstico	Lim. Inf	Lim. Sup	Pronóstico	Lim. Inf	Lim. Sup	Pronóstico	Lim. Inf	Lim. Sup	Pronóstico	Lim. Inf	Lim. Sup
Dec 2020	125.004	116.960	133.600	125.912	117.839	134.538	125.877	117.754	134.560	126.074	117.920	134.791
En 2021	104.026	97.088	111.461	104.564	97.546	112.086	104.015	97.091	111.433	103.913	97.006	111.313
Feb 2021	107.034	99.232	115.449	107.980	100.005	116.591	106.799	99.030	112.177	106.827	99.104	115.152
Mar 2021	113.871	104.628	123.931	114.277	104.834	124.570	113.601	104.401	123.612	113.495	104.331	123.464
Abr 2021	11.305	101.784	121.716	112.172	102.520	122.732	111.057	101.603	121.390	111.115	101.716	121.382
May 2021	116.444	105.854	128.093	116.870	106.057	120.785	116.142	105.631	127.700	115.996	105.574	127.447
Jun 2021	114.782	103.765	126.968	115.432	104.201	127.874	114.568	103.388	126.957	114.580	103.723	126.575
Jul 2021	117.292	105.522	130.375	116.968	105.097	130.179	117.014	105.202	130.152	117.041	105.471	129.881
Ago 2021	119.176	106.690	133.124	120.196	107.415	134.498	119.584	106.894	133.781	119.632	107.292	133.392
Sep 2021	121.367	108.148	136.201	121.901	108.461	137.007	121.101	107.690	136.181	121.133	107.884	136.009
Oct 2021	123.385	109.461	139.081	123.752	109.608	139.720	123.262	108.991	139.403	123.295	109.165	139.255
Nov 2021	127.263	112.415	144.073	127.839	112.729	144.975	126.865	111.600	144.219	126.819	111.640	144.062

Teniendo los pronósticos que se muestran en la tabla 2 para todos los modelos propuestos, vale la pena interpretar el resultado con algún periodo particular, como lo puede ser el mes de febrero de 2021, para el cual el modelo uno pronostica que el índice de ventas nominales será de 107.034 puntos y se situará entre los 99.232 y los 115.449 puntos con una confianza del 95 %. Por su parte, el modelo dos pronostica que el índice de ventas nominales de febrero de 2021 será de 107.980 y se ubicará entre los 100.005 y 116.591 puntos, mientras que el modelo tres proyecta que será de 106.799 puntos y se situará entre los 99.030 y los 112.177 puntos con una confianza del 95 %. Por último, el modelo cuatro pronostica que el índice de ventas nominales de febrero de 2021 será de 106.827 puntos y con una confianza del 95% el índice de ventas nominales estará entre 99.104 y 115.152 puntos. Ahora bien, es importante comparar los cuatro modelos a partir de la tabla de diferentes medidas de error gracias al conocimiento de los valores reales para el periodo ex post, para lo que se presenta la tabla 3.

Tabla 3. Precisión de los Pronósticos puntuales y de los I.P del 95%

Medidas	Modelo 1	Modelo 2	Modelo 3	Modelo 4
RMSE	0.5962	0.1083	0.6586	0.6737
MAE	0.5462	0.1019	0.5886	0.6174
MAPE (%)	0.4683	0.0874	0.5091	0.5338
Amplitud. Media I.P	22.7103	23.1035	22.9409	22.6579
Cobertura (%) I.P	100%	100%	100%	100%

¹ Estudiante Ingeniería Industrial, Universidad Nacional de Colombia – Sede Medellín.² Estudiante Ingeniería Administrativa, Universidad Nacional de Colombia – Sede Medellín³ Estudiante Ingeniería Administrativa, Universidad Nacional de Colombia – Sede Medellín

Vemos que el modelo dos presenta el menor valor tanto como en RMSE, MAE y MAPE, siguiendo con esta idea según el RMSE el modelo dos se equivocó en promedio en cada pronóstico del periodo ex post en 0.1083 puntos del índice de ventas nominales y según MAE este mismo modelo se equivoca en promedio 0.1019 puntos del índice mientras que según MAPE el modelo dos se equivocó en promedio en cada pronóstico en un 0.0874% respecto al valor real del índice de ventas nominales, en general los cuatro modelos presentan buenos resultados en cuanto pronóstico según las medidas mencionadas anteriormente; en tanto a los intervalos de predicción vemos que todos contienen el valor real de la serie para cada uno de los periodos ex post en este caso el modelo con un intervalo de predicción más estrecho es el modelo cuatro seguido por el modelo uno; para ver más fácilmente la calidad de la predicción se presenta en la figura 1 las predicciones puntuales de cada modelo y los valores reales del índice de ventas para el periodo ex post

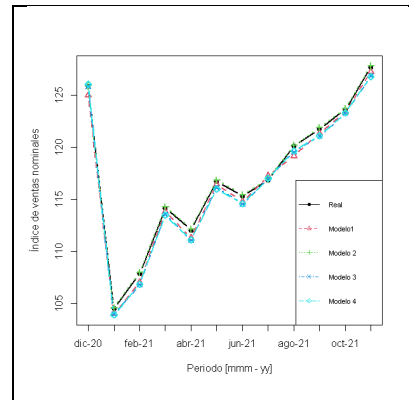


Figura 1: Comparación de los pronósticos.

Podemos concluir con la gráfica que no hay diferencias prácticas de importancia entre los pronósticos puntuales y los valores reales. Como estamos interesados en escoger el modelo que mejor pronóstica sin dejar de lado la validez de supuestos de dicho modelo y teniendo en cuenta que no hay diferencias importantes se preferirá los modelos más parsimoniosos

3.6 Conclusiones

A Partir del análisis descriptivo se evidencia que la serie tiene tanto componente de tendencia como componente de estacionalidad adicionalmente su varianza no es constante y aumenta en la misma dirección de la tendencia, por lo que se procede a modelar el logaritmo de la serie, a partir de esto obtenemos que la varianza se estabiliza pero la serie sigue sin ser estacionaria por que aún se mantiene la componente de tendencia y la componente estacional, estas mismas conclusiones se reflejan en la respectiva gráfica ACF, por tanto es pertinente analizar el comportamiento del logaritmo natural de la serie cuando se le aplica la primera diferencia regular, aunque con esto logramos deshacernos de la tendencia, en la respectiva ACF existe un decaimiento lento para $k = 12, 24, 36$, es decir, indica necesidad de aplicar también la diferencia estacional, con esto concluimos que este proceso no es estacionario en sentido débil; al aplicar solo la diferencia estacional sobre el logaritmo de la serie obtenemos que aunque para la gráfica ACF en los k múltiplos de 12 ya no se encuentran correlaciones significativas no podemos considerar el proceso como estacionario en covarianza porque se encontraron evidencias en contra de los supuestos media constante y homocedasticidad; ahora al aplicar el filtro mixto podemos comentar de la gráfica ACF que aunque se encuentran tanto como para la parte regular como para la parte estacional autocorrelaciones muestrales estadísticamente significativas podemos considerar ambas partes como ergódicas además no se encuentran evidencias graficas en contra de los supuestos varianza constante y media constante así concluimos que el logaritmo natural de la serie diferenciado por tendencia y estacionalidad es un proceso ergódico, lo cual muestra además que ya no hay evidencia de existencia de raíces unitarias regulares y estacionales, por lo que no es necesario diferenciar más por tendencia o estacionalidad; según el test HEGY, la serie $\log(Y_t)$ tiene tanto raíz unitaria regular como estacional, luego, es apropiado diferenciar regular y estacionalmente a esta serie, y con ello se confirma lo que se vio a partir del análisis gráfico sobre la serie de logaritmos y sus diferencias: regular, estacional y combinado. Para empezar con la identificación de modelos comenzamos por identificar los patrones de las gráficas ACF y PACF para la serie $\log(Y_t)$ diferenciada por tendencia y estacionalidad en la respectiva gráfica ACF para la parte regular identificamos un patrón tipo cola exponencial sinusoidal mientras que la PACF se identifica un patrón tipo corte donde el ultimo rezago estadísticamente diferente de cero es el segundo, por lo que para la parte regular se identifica un AR(2), lo que implica que para la parte regular de $\log(Y_t)$ se tiene un ARIMA(2,1,0); pasando a la parte estacional en la ACF identificamos un patrón tipo corte con ultimo rezago estacional significativo en $k = 24$ y en la gráfica PACF un patrón tipo cola por lo que los modelos propuestos deben cumplir la condición de que la parte estacional se modela con un MA(2)[12] y en el caso de $\log(Y_t)$ un ARMA(0,1,2)[12], mezclando ambos análisis se proponer el modelo ARIMA(2,1,0)(0,1,2)[12] para $\log(Y_t)$; pasando a la identificación de modelos SARIMA(p,d,q)(P,D,Q)[s] con ayuda de métodos automáticos, con el método auto.arima() solo se consideraron aquellos que contemplaban tanto la diferencia regular como la estacional, también se descartó aquel modelo que proponía un proceso de medias móviles para la parte regular esto debido a que no concuerda con las evidencias graficas que mencionamos anteriormente, finalmente el modelo candidato resultante de auto.arima() es: ARIMA(4,1,0)(1,1,2)[12] sin deriva para $\log(Y_t)$; pasando a la identificación con armasubsets con el primer renglón del tablero 12x12 s sobre $\nabla \nabla_{12} \log Y_t$ usando el método "ols" y agregando al parámetro ϕ_6

obtenemos el modelo ARIMA(2, 1, 10)(0, 1, 1)[12], con únicos parámetros no nulos $\phi_1, \phi_2, \phi_6, \theta_{10}, \theta_1$; ahora con el primer renglón del tablero 18×18 de armassubsets sobre $\nabla \nabla 12 \log Yt$ usando el método “ols” se identifica un modelo ARIMA(9, 1, 10)(0, 1, 1)[12] que solo emplea los términos $\phi_1, \phi_2, \phi_9, \theta_{10}, \theta_1$ **comentar patron parte regular y estacional**

Pode

4. Referencias y citas

Toda figura, tabla, ecuación, sección, bibliografía, dirección electrónica, etcétera, debe ser referenciada, como se ejemplifica en este documento.

4.1 Ejemplo (Citaciones y referencias bibliográficas). Las citaciones bibliográficas deben estar acompañadas por un número en el texto empleando el ambiente “insertar” seguido de “nota al pie...” Las referencias bibliográficas deben ir al final del documento^{1, 2, 3]} y su formato es como se muestra en la Sección Referencias de este documento.

5. Objetos

Toda figura, tabla y ecuación es un objeto. Los objetos deben ser colocados lo más cerca posible del párrafo donde son referenciados por primera vez y no se admite que floten en el texto.

5.1 Tablas

Se deben colocar en una sola tabla, los resultados de todos los modelos que sean de la misma naturaleza (por ejemplo, tablas de parámetros estimados, tablas de predicciones) y que su colocación sea centrada. Las tablas deben ser tituladas y numeradas en su parte superior, como muestran las Tablas 1, 2 y 3 del Ejemplo en 5.4.1.

5.2 Figuras

Las figuras deberán titularse y numerarse en su parte inferior como se muestra en el Ejemplo en 5.4.2. Se deberá dejar en el texto el espacio suficiente para ubicar la figura en el sitio que le corresponde. Las dimensiones de las figuras deberán ser de 5cm x 5cm excepto para aquellas en las que sea necesario un tamaño mayor, pero no deberán ser excesivas al punto de ocupar cada una media página o más.

5.3 Ecuaciones

Si se escriben como párrafo, deben ser centradas y con la numeración a la derecha y entre paréntesis. También puede presentarlas dentro de una tabla como se ilustra en la sección 5.4.3. Si se colocan individualmente, en ese caso forman parte de un párrafo y observan las reglas de puntuación. Solo enumere las ecuaciones a las que se refiera en el texto. Para referirse a las ecuaciones utilice la palabra ecuación seguida de su número como se ilustra en el Ejemplo en Sección 5.4.3 Las ecuaciones deben ser escritas en el editor de ecuaciones de Word, no pegadas como imagen capturada de otros documentos.

5.4 Ejemplos de tablas, figuras y ecuaciones

5.4.1 Ejemplo Tablas. Si las tablas tienen el mismo tipo de información deben colocarse juntas y además si tienen mismo número de filas pueden colocarse una al lado de la otra. En caso contrario se colocan una debajo de la otra, por ejemplo como se ilustra a continuación.

Tabla 1: Parámetros estimados Modelos de regresión global

Tabla 1a. Parámetros estimados en Modelo 1					Tabla 1c. Parámetros estimados en Modelo 2				
Parámetros	Estimación	Error Std	T_0	$P(t_{144} > T_0)$	Parámetros	Estimación	Error Std	T_0	$P(t_{144} > T_0)$
β_0	6.25081	0.02487	251.34474	0.00000	β_0	6.25467	0.04752	131.63476	0.00000
β_1	0.01988	0.00129	15.39242	0.00000	β_1	0.02015	0.00209	9.62945	0.00000
β_2	-1.290×10^{-4}	0.00002	-6.54274	0.00000	β_2	-1.323×10^{-4}	0.00003	-4.75668	0.00000
β_3	3.168×10^{-7}	8.526×10^{-8}	3.71609	0.00029	β_3	3.301×10^{-7}	1.101×10^{-7}	2.99833	0.00320
δ_1	-0.13167	0.01574	-8.36568	0.00000	δ_1	-0.14307	0.01832	-7.80790	0.00000
δ_2	-0.01853	0.01574	-1.17723	0.24105	δ_2	-0.02956	0.01719	-1.71946	0.08768
δ_3	0.01446	0.01574	0.91883	0.35972	δ_3	0.00512	0.01686	0.30392	0.76163
$\sqrt{MSE} = 0.06813$ (escala log); AIC= 8995.101 BIC= 10345.524					$\sqrt{MSE} = 92.52$, AIC= 8956.667, BIC= 10301.320				

Tabla 1b. Parámetros estimados en Modelo 1b					Tabla 1d. Parámetros estimados en Modelo 2b				
Parámetros	Estimación	Error Std	T_0	$P(t_{144} > T_0)$	Parámetros	Estimación	Error Std	T_0	$P(t_{144} > T_0)$
β_0	6.21688	0.02275	273.28662	0.00000	β_0	6.21279	0.04641	133.85500	0.00000
β_1	0.01988	0.00129	15.39242	0.00000	β_1	0.02015	0.00209	9.62945	0.00000
β_2	-1.290×10^{-4}	0.00002	-6.54274	0.00000	β_2	-1.323×10^{-4}	0.00003	-4.75668	0.00000
β_3	3.168×10^{-7}	8.526×10^{-8}	3.71609	0.00029	β_3	3.301×10^{-7}	1.101×10^{-7}	2.99832	0.00320
α_1	-0.07307	7.818e-03	-9.34583	0.00000	α_1	-0.07410	0.00907	-8.16517	0.00000
γ_1	0.00926	0.00787	1.17723	0.24105	γ_1	0.01478	0.00860	1.71946	0.08768
γ_2	0.02467	0.00555	4.44922	0.00002	γ_2	0.02709	0.00625	4.33595	0.00003
$\sqrt{MSE} = 0.06813$ (escala log); AIC= 8995.101 BIC= 10345.524					$\sqrt{MSE} = 92.52$, AIC= 8956.667, BIC= 10301.320				

Tabla 2. Pronósticos puntuales y por I.P del 95% de confianza

Período	Modelo 1			Modelo 1b			Modelo 2			Modelo 2b		
	Pronóstico	Lim. Inf	Lim. Sup	Pronóstico	Lim. Inf	Lim. Sup	Pronóstico	Lim. Inf	Lim. Sup	Pronóstico	Lim. Inf	Lim. Sup
1993 Q4	1649.702	1429.387	1903.976	1649.702	1429.387	1903.976	1668.857	---	---	1668.857	---	---
1994 Q1	1450.015	1255.496	1674.671	1450.015	1255.496	1674.671	1450.493	---	---	1450.493	---	---
1994 Q2	1628.089	1408.739	1881.594	1628.089	1408.739	1881.594	1629.504	---	---	1629.504	---	---
1994 Q3	1687.280	1458.900	1951.411	1687.280	1458.900	1951.411	1691.936	---	---	1691.935	---	---

Tabla 3. Precisión de los Pronósticos puntuales y de los I.P del 95%

Medidas	Modelo 1	Modelo 2	Modelo 3	Modelo 4
RMSE	0.5962	0.1083	0.6586	0.6737
MAE	0.5462	0.1019	0.5886	0.6174
MAPE (%)	0.4683	0.0874	0.5091	0.5338
Amplitud. Media I.P	22.7103	23.1035	22.9409	22.6579
Cobertura (%) I.P	100%	100%	100%	100%

5.4.2 Ejemplo Figuras. Figuras del mismo tipo deben colocarse juntas como se ilustra a continuación.

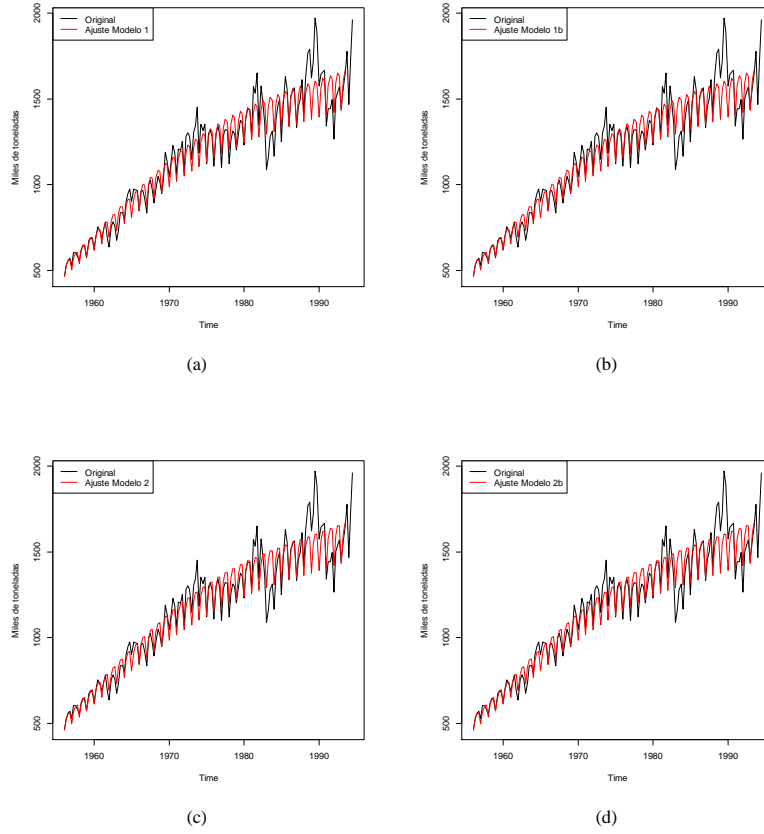


Figura 1: Gráficas de los ajustes. (a) En modelo 1; (b) En modelo 1b; (c) En modelo 2; (d) En modelo 2b

5.4.3 Ejemplo ecuaciones

5.4.3.1 Ecuaciones cada una en un párrafo independiente

Las ecuaciones (1) a (4) corresponden a los modelos 1 (log-cúbico estacional con indicadoras, nivel de referencia Q4), modelo 1b (log-cúbico estacional usando variables trigonométricas en la representación de la componente estacional, en frecuencias $F_1 = 1/4$, $F_2 = 1/2$), modelo 2 (exponencial-cúbico estacional con variables indicadoras, trimestre de referencia Q4 (se usan las indicadoras, nivel de referencia Q4) y modelo 2b (exponencial-cúbico estacional usando variables trigonométricas en la representación de la componente estacional, en frecuencias $F_1 = 1/4$, $F_2 = 1/2$), respectivamente,

$$\log(Y_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \delta_3 I_{3,t} + E_t, E_t \sim iid N(0, \sigma^2). \quad (1)$$

$$\log(Y_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \alpha_1 \sin\left(\frac{\pi t}{2}\right) + \gamma_1 \cos\left(\frac{\pi t}{2}\right) + \gamma_2 \cos(\pi t) + E_t, E_t \sim iid N(0, \sigma^2). \quad (2)$$

$$Y_t = \exp(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \delta_3 I_{3,t}) + E_t, E_t \sim iid N(0, \sigma^2). \quad (3)$$

$$Y_t = \exp\left(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \alpha_1 \sin\left(\frac{\pi t}{2}\right) + \gamma_1 \cos\left(\frac{\pi t}{2}\right) + \gamma_2 \cos(\pi t)\right) + E_t, E_t \sim iid N(0, \sigma^2). \quad (4)$$

5.4.3.2 Ecuaciones como parte de una tabla

Tabla 4. Ecuaciones de los modelos propuestos

Modelo 1
$\log(Y_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \delta_3 I_{3,t} + E_t, E_t \sim iid N(0, \sigma^2)$
Modelo 1b
$\log(Y_t) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \alpha_1 \sin\left(\frac{\pi t}{2}\right) + \gamma_1 \cos\left(\frac{\pi t}{2}\right) + \gamma_2 \cos(\pi t) + E_t, E_t \sim iid N(0, \sigma^2)$
Modelo 2
$Y_t = \exp(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \delta_3 I_{3,t}) + E_t, E_t \sim iid N(0, \sigma^2)$
Modelo 2b
$Y_t = \exp\left(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \alpha_1 \sin\left(\frac{\pi t}{2}\right) + \gamma_1 \cos\left(\frac{\pi t}{2}\right) + \gamma_2 \cos(\pi t)\right) + E_t, E_t \sim iid N(0, \sigma^2)$

Referencias

- [1] Chuang, I. L. and Nielsen, M. A. (2000), “Quantum computation and quantum information”, Cambridge: Cambridge University Press.
- [2] Moreno, L. F. (2004), “Factorización cuántica de números enteros. Una introspectiva al algoritmo de Shor”, Universidad EAFIT.
- [3] Shor, P. W. (1994), “Algorithms for quantum computation. Discrete logarithms and factoring,” In 35th Annual Symposium on Foundations of Computer Science, pp. 124–134, IEEE.