**3.5 Pronósticos para la validación cruzada**

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. Además, para los intervalos de pronóstico se usará un nivel de confianza del 95% y se presentará la gráfica de los pronósticos contra los datos reales en los tiempos de pronóstico *ex post.* De esta manera, vale la pena comenzar estudiando la ***tabla 9.***

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Tabla 9.** 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 9*** 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 del sector manufacturero colombiano en pesos nominales será de 107.034 puntos y se situará entre los 99.232 y los 115.449 puntos con un nivel de confianza del 95 %. Por su parte, el ***modelo dos*** pronostica que el índice de ventas del sector manufacturero colombiano en pesos 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 un nivel de confianza del 95 %. Por último, el ***modelo cuatro*** pronostica que el índice de ventas del sector manufacturero colombiano en pesos nominales de febrero de 2021 será de 106.827 puntos y con un nivel de confianza del 95% el índice de ventas del sector manufacturero colombiano en pesos 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 10***.

***Tabla 10.*** Métricas para el análisis de precisión de los pronósticos puntuales y de los intervalos de predicción (I.P.) del 95 % de confianza.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 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% |

A partir de las métricas presentadas en la ***tabla 10*** se ve 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 del sector manufacturero colombiano en pesos nominales y según MAE este mismo modelo se equivoca en promedio 0.1019 puntos del índice, mientras que según MAPE el ***modelo dos*** erró en promedio en cada pronóstico en un 0.0874% respecto al valor real del índice de ventas del sector manufacturero colombiano en pesos nominales. En general los cuatro modelos presentan buenos resultados en cuanto pronóstico según las medidas mencionadas anteriormente.

Luego, en cuanto a los intervalos de predicción, se ve 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 13.*** Comparación de los pronósticos. |

Podemos concluir con la ***figura 13*** que no hay diferencias prácticas de importancia entre los pronósticos puntuales y los valores reales. Como estamos interesados en escoger el modelo que mejor pronostica sin dejar de lado la validez de supuestos de dicho modelo, y teniendo en cuenta que no hay diferencias importantes, se preferirán los modelos más parsimoniosos , ya que ninguno de los modelos paso el supuesto de que sus errores tienen una distribución normal, nos quedamos con el modelo uno por las razones comentadas anteriormente.

**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, con lo que se consigue que la varianza se estabilice, pero la serie sigue sin ser estacionaria por que aún se mantiene la componente de tendencia y la componente estacional. Estas mismas observaciones se reflejan en la respectiva gráfica ACF, y por tanto es pertinente analizar el comportamiento del logaritmo natural de la serie cuando se le aplica la primera diferencia regular, aunque con esto se logra eliminar de la tendencia, en la ACF asociada se tiene un decaimiento lento para, es decir, indica necesidad de aplicar también la diferencia estacional. Con lo anterior se concluye que este proceso no es estacionario en sentido débil. Entonces, al aplicar solo la primera diferencia estacional sobre el logaritmo de la serie se obtiene que aunque para la gráfica ACF en los múltiplos de 12 ya no se encuentran correlaciones significativas, no se puede considerar el proceso como estacionario en covarianza porque se encontraron evidencias en contra de los supuestos media constante y homocedasticidad; después, al aplicar el filtro mixto (primera diferencia regular y primera diferencia estacional), se puede comentar de la gráfica ACF que a pesar de que se encuentran tanto para la parte regular como para la parte estacional autocorrelaciones muestrales estadísticamente significativas, se pueden considerar ambas partes como ergódicas, y adicionalmente no se encuentran evidencias gráficas en contra de los supuestos de homocedasticidad y media constante, y así, se concluye 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.

A continuación, según el test HEGY, la serie transformada usando logaritmo natural tiene tanto raíz unitaria regular como estacional, luego, es apropiado diferenciar regular y estacionalmente a esta serie, y con ello se confirma lo evaluado con los análisis descriptivos que se realizaron justo antes en cuanto a la necesidad de tomar la primera diferencia regular y estacional para el logaritmo natural del índice de ventas del sector manufacturero colombiano en pesos nominales.

Ahora, ante la necesidad de ajustar modelos SARIMA, se procede con la identificación de los patrones de las gráficas ACF y PACF para la serie diferenciada por tendencia y estacionalidad. De este modo, para la parte regular se identifica un patrón tipo cola exponencial sinusoidal mientras que la PACF se identifica un patrón tipo corte donde el último 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 se tiene un ARIMA(2,1,0); hecho esto, se prosigue con el análisis de la parte estacional, y comenzando con la ACF es posible identificar un patrón tipo corte con último rezago estacional significativo en 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 un ARMA(0,1,2)[12]. Entonces, mezclando ambos análisis se propone el ***modelo uno:*** ARIMA(2,1,0)(0,1,2)[12] para .

A continuación, 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, y también se descartó aquel modelo que proponía un proceso de medias móviles para la parte regular debido a que no concuerda con las evidencias gráficas que se mencionaron anteriormente, por lo que el modelo candidato resultante de auto.arima() es: ***modelo 2****:* ARIMA(4,1,0)(1,1,2)[12] sin deriva para .

Pasando a la identificación con armasubsets con el primer renglón del tablero 1212 sobre usando el método “ols” y agregando al parámetro se obtiene el ***modelo 3*:** ARIMA(6, 1, 10)(0, 1, 1)[12], con únicos parámetros no nulos . Ahora con el primer renglón del tablero 18 × 18 de armasubsets sobre usando el método “ols” se identifica el **modelo 4. ARIMA(9, 1, 10)(0, 1, 1)[12]** que solo emplea los términos

Pasando al ajuste de estos modelos se puede comentar de la gráfica de la serie real y el modelo ajustado que se observa que sigue adecuadamente la tendencia y la estacionalidad presente en la serie; además, para los cuatro modelos no se observan diferencias significativas en cuanto al ajuste, pero, aun así, no está de más comentar que el ***modelo uno*** es el que cuenta con un menor AIC y BIC.

Asimismo, considerando la validación de los supuestos bajo los cuales fueron ajustados los modelos, para ninguno de los cuatro se encontraron evidencias gráficas en contra de que sus medias fueran iguales a cero, las varianzas sean constantes o que los errores sean un proceso ergódico. Con el test Ljung-Box se confirmó que no hay evidencia suficiente en contra de que los errores sean incorrelacionados, y como para todos los modelos se obtuvo que los errores eran un proceso de ruido, luego se pasó a probar si para todos los modelos los errores son un proceso ruido blanco mediante el chequeo del supuesto de normalidad. De esta forma, con ayuda de los gráficos cuantil-cuantil se rechazó el supuesto de normalidad para los cuatro modelos, concluyendo así que ningún modelo SARIMA tiene errores ruido blanco. Con esto en mente, se escoge como mejor modelo al ***modelo uno***, ya que en términos de pronósticos no existen diferencias prácticas de importancia entre los modelos y además este es el modelo que contiene un menor número de parámetros. Por último, es importante comentar que ya que el ***modelo uno*** no es un modelo válido en términos de cumplimiento de supuestos, sus pronósticos no son confiables.

Ahora vale la pena hacer un repaso sobre algunos de los modelos obtenidos a lo largo de los tres trabajos: mejores modelos local y global de regresión global del trabajo uno, mejor modelo de regresión global con errores ARMA en el trabajo dos, mejor modelo SARIMA del trabajo tres, cuyas ecuaciones son presentadas en la ***tabla 11.***

|  |
| --- |
| ***Tabla 11*.** Ecuaciones de los mejores modelos ajustados. |
| **Mejor modelo global.** Logpolinomial de grado seis estacional con funciones trigonométricas en cinco frecuencias Fj =j/12 , j = 1, 2, 3, 4, 5. |
| **Mejor modelo local.** Descomposición multiplicativa y *loess* lineal  En la vecindad de un tiempo donde se quiere el ajuste  con para todo t en la vecindad de , y  con parámetros de la recta local. |
| **Mejor modelo con errores ARMA .**  Logpolinomial de grado seis estacional con funciones trigonométricas en cinco frecuencias Fj =j/12 , j = 1, 2, 3, 4, 5; con error estructural ARMA: ARMA(12,10) con 𝜙7 y 𝜃10.  donde |
| **Mejor modelo SARIMA** ARIMA(2, 1, 0)(0, 1, 2)[12]  **,** |

Hecho esto, en la ***figura 14*** se presentan las gráficas que comparan la serie en su escala original comparadas con los modelos ajustados. Así pues, es importante rescatar que para los mejores modelos global, ARMA y SARIMA se consigue en general seguir de forma adecuada la dinámica de la serie, lo cual se refleja en el hecho que las respectivas líneas rojas, asociadas a los modelos ajustados, se apegan de forma aceptable a la línea negra, correspondiente a los datos reales de la serie temporal; empero, para el mejor modelo global se puede observar que, si bien logra seguir bien la dinámica de la serie en cuanto a su tendencia y estacionalidad, cuando se tienen ciclos el ajuste no es adecuado, como se puede verificar por ejemplo en el año 2009, cuando el modelo sobreestima los índices de ventas del sector manufacturero colombiano en pesos nominales para los periodos de dicho año. Este análisis se verifica claramente al observar las métricas de ajuste con los criterios de información de Akaike y bayesiano presentados en la ***tabla 12,*** pues este par de métricas son máximas para el mejor modelo global, dejando claro entonces que es el modelo que peores ajustes consigue a la luz de los índices realmente observados. Ahora bien, en cuanto a los otros tres modelos, se tiene que según el AIC el modelo que mejores ajustes realiza es el modelo local de descomposición multiplicativa y *loess* lineal, y los mejores modelos ARMA y SARIMA tienen AIC semejantes y no muy distantes entre sí, y ambos separados una unidad aproximadamente del AIC del mejor modelo local; no obstante, con el criterio BIC el análisis cambia en cuanto el modelo con el menor valor de criterio de información bayesiano es ahora el modelo SARIMA y con una mayor distancia respecto al segundo y tercer modelo en orden de acuerdo con este criterio que en el caso del AIC. Así pues, considerando que prefiere al criterio de información bayesiano, se concluye que el modelo que mejores ajustes realiza es el modelo SARIMA, que corresponde al ***modelo uno*** presentado en este trabajo.

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |

***Figura 14.*** Gráficas de los ajustes (a) en mejor modelo global (modelo uno del trabajo uno); (b) en mejor modelo local (modelo cuatro del trabajo uno); (c) en mejor modelo con errores ARMA(modelo cuatro del trabajo dos); (d) en modelo 2b.

**Tabla 12.** Valores criterios AIC y BIC

|  |  |  |
| --- | --- | --- |
| Medidas | AIC | BIC |
| Mejor modelo global | 9.11703 | 11.674705 |
| Mejor modelo local | 4.241348 | 7.265074 |
| Mejor modelo con errores ARMA | 5.303542 | 7.629496 |
| Mejor modelo SARIMA | 5.514306 | 5.844665 |

En cuanto a los pronósticos puntuales de estos modelos, se ve en la ***figura 15*** que los pronósticos del modelo global son los que más se alejan de los valores reales que toma la serie en el periodo *ex post*, pero el resto de modelos alcanzan un conjunto de predicciones aceptable, pero destaca la buena predicción del modelo SARIMA para el periodo *ex post*. Para respaldar estas evidencias gráficas se recurre a los criterios para la precisión de los pronósticos puntuales y de los intervalos de predicción al 95% de confianza que se muestran en la ***tabla 13***, donde, a excepción del modelo global, los demás modelos tienen valores similares y resulta interesante que el modelo local es el cuenta con menor MAE y MAPE, mientras que el modelo SARIMA presenta el menor RMSE. A propósito de los intervalos de predicción, es importante tener en cuenta que por la forma en que se ajustaron el mejor modelo local y el mejor modelo ARMA, no es posible obtener para estos intervalos de predicción, sino pronósticos puntuales únicamente, pero de todos modos, entre el mejor modelo local y el mejor modelo SARIMA, vale la pena rescatar que ambos cubren el 100 % de los índices de ventas del sector manufacturero colombiano en pesos nominales para los meses del periodo *ex post,* pero se diferencian porque el mejor modelo SARIMA permite obtener intervalos de predicción más angostos, lo cual es claramente más beneficioso. Así, se puede determinar que el mejor modelo en cuanto a los pronósticos es el SARIMA ya que tiene el menor RMSE, su MAE y su MAPE no difieren mucho de los del mejor modelo local y además permite obtener intervalos de predicción que permiten tener una mejor idea del rango de valores que puede tomar el índice de interés.

|  |
| --- |
|  |

**Figura 15.** Comparación de los pronósticos de los mejores modelos con respecto a los índices realmente observados en el periodo *ex post.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Tabla 13.** Precisión de los pronósticos puntuales y de los I.P del 95% | | | | |
| Medidas | **Mejor modelo global** | Mejor modelo local | Mejor modelo con errores ARMA | Mejor modelo SARIMA |
| RMSE | 2.906349 | 0.6728501 | 0.7601474 | 0.5962 |
| MAE | 2.510166 | 0.5283713 | 0.6555441 | 0.5462 |
| MAPE (%) | 2.108491 | 0.4439505 | 0.5586861 | 0.4683 |
| Amplitud. Media I.P | 26.42722 | NA | NA | 22.7103 |
| Cobertura (%) I.P | 100% | NA | NA | 100% |

Por último, se va a hacer un recuento del cumplimiento de los supuestos bajo los cuales se construyen todos los modelos en los tres trabajos, que estriban básicamente en el hecho de que sus errores sean procesos de ruido blanco, lo cual implica que sean estacionarios en sus medias sean nulas, sean homocedásticos e independientes entre sí y que se distribuyan siguiendo una normal. Para ello, se presenta la ***tabla 14.***

***Tabla 14.*** Resumen del cumplimiento del supuesto de ruido blanco para los mejores modelos ajustados.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Modelo | Supuestos  sobre | | ¿Hay evidencia fuerte en contra de? | | | | | | | | | | Modelo  Válido | |
| Media cero | | Homocedasticidad | | Independencia | | | Normalidad | | |
|  |  |  | Sí | No | Sí | No | Sí | No | *NA* | Sí | No | *NA* | Si | No |
| Mejor modelo global | ✘ |  | ✘ |  | ✘ |  |  |  | ✘ |  |  | ✘ |  | ✘ |
| Mejor modelo local | ✘ |  |  | ✘ |  | ✘ | ✘ |  |  |  |  | ✘ |  | ✘ |
| Mejor modelo con errores ARMA |  | ✘ |  | ✘ |  | ✘ |  | ✘ |  | ✘ |  |  |  | ✘ |
| Mejor modelo SARIMA | ✘ |  |  | ✘ |  | ✘ |  | ✘ |  | ✘ |  |  |  | ✘ |

De esta manera, se debe iniciar mencionando que aunque el modelo global hubiera cumplido con los supuestos aún no sería recomendable para modelar la serie temporal asociada le índice de ventas del sector manufacturero colombiano en pesos nominales, ya que los modelos globales asumen que las relaciones establecidas por el modelo son estables en el tiempo lo que no es un hecho viable para índices como, puesto que las relaciones comerciales, industriales y económicas evolucionan con el paso del tiempo. Así pues, una solución a esto serían los modelos locales, que son precisamente aquellos que admiten cambios en los parámetros; aunque un modelo local nos proporciona buenos pronósticos dado que estos buscan la última estimación que se ajustó, esto es peligroso ya que nada asegura que el último patrón local y el patrón futura van a coincidir.

En cuanto a los problemas enfrentados en la modelación, se puede empezar comentando que para el modelo global fue relativamente fácil ajustar el modelo ya que desde los enunciados se especificaba el orden del polinomio que modelaba la tendencia, pero determinar dicho orden no es una tarea fácil ya que para llegar a él se requiere una serie de pruebas que comparen diferentes órdenes para determinar cuál resulta más apropiado. Por el lado del modelo local, afortunadamente se cuentan con métodos automáticos que ayudan a proponer dichos modelos ya que encontrar un parámetro de suavizamiento es una tarea exhaustiva. Por su lado, en el modelo con errores ARMA, además de que se debe encontrar cómo representar tanto tendencia como estacionalidad (trabajo que ya se había adelantado con el modelo global), se debe definir cómo modelar el error, lo que tal vez puede representar problemas ya que no siempre es fácil identificar los órdenes de y sin la ayuda de métodos automáticos. Este último inconveniente también lo comparte el modelo SARIMA, y en este caso identificar patrones y órdenes en gráficas ACF y PACF llega a ser más retador aún por la interacción que se da entre la parte regular y la parte estacional, además que se limita a pocos rezagos (lo cual es muy complejo para modelos trimestrales para el análisis asociado a la parte regular).

Mencionado todo lo anterior, se puede concluir que los modelos ajustados, si bien son adecuados siguiendo la dinámica de la serie (con las salvedades explicadas para el mejor modelo global), todos fallan en el cumplimiento de los supuestos sobre los, pero como nunca se van a encontrar modelos perfectos ,no está demás comentar que **el modelo SARIMA arrojó unos buenos pronósticos a corto plazo y aunque la confiabilidad de estos no es asegurada sobre todo para los intervalos de predicción no, es una mala idea considerar dicho modelo.**

**BIBLIOGRAFÍA**

Álvarez, N. G. (2022). *Notas de Clase Estadística III*.