**6. PRONÓSTICOS PARA EL PERIODO *EX POST***

Los pronósticos realizados para el periodo *ex post*, como estrategia de validación cruzada, se harán nuevamente con origen en la observación n = 239. Es preciso recordar que para nuestro modelo no es posible obtener intervalos de pronóstico, en tanto el modelo global del que se partió es parcialmente multiplicativo, y en este sentido, únicamente se presentarán pronósticos puntuales y medidas de error asociados a este tipo de pronósticos (RMSE, MAE y MAPE). Además, se presentará la ecuación de pronósticos para el modelo **tres,** el cual se refleja en la ***tabla XXXX1.***

|  |
| --- |
| **Tabla XXXX1.** Ecuación de pronósticos para el ***modelo tres*** con origen en |
| **Modelo 3**  Donde    En las anteriores ecuaciones se tiene que |

Por otro lado, se tiene que las medidas de precisión para los pronósticos puntuales de los cuatros modelos están presentes en la tabla **XXXX2.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tabla XXXX2.** Precisión de los Pronósticos puntuales | | | | |  |
| Medidas | Modelo 1 | Modelo 2 | Modelo 3 | Modelo 4 | Modelo local de descomposición multiplicativa y *loess* lineal |
| RMSE | 0.84 | 0.79 | 1.07 | 0.76 | 0.67 |
| MAE | 0.73 | 0.65 | 0.86 | 0.66 | 0.53 |
| MAPE (%) | 0.63 | 0.56 | 0.75 | 0.56 | 0.44 |
| Amplitud. Media I.P | NA | NA | NA | NA | NA |
| Cobertura (%) I.P | NA | NA | NA | NA | NA |

De la tabla se puede rescatar que los modelos en general presentan resultados aceptables ya que sus medidas de error son bajas, y para para cada una de ellas se presentan que el modelo **cuatro** es el que tiene cifras de error más bajas si no tenemos en cuenta las obtenidas por el modelo local, seguidas por el modelo **dos** y el modelo **tres**, en ese orden, y finalmente se tiene al modelo **uno,** que es el que peores indicadores de error presenta. Entonces, tomando al modelo cuatro, que a la luz de estos valores se asume como el mejor modelo, según el RMSE el modelo cuatro se equivocó en promedio en cada pronóstico en 0.76 puntos del índice y según el MAE, se equivoca en promedio en 0.66 puntos del índice de ventas nominales, toda vez que según el MAPE este modelo se equivocó en promedio en cada pronóstico en un 0.56% respecto a cada valor real. Si ponemos a competir al modelo dos con el modelo local en estos términos, el modelo local es el ganador.

Se presenta la ***figura XXXX3,*** en la cual se compara el pronóstico realizado por cada uno de los modelos presentados y en el que se compara con los valores realmente observados del índice de ventas nominales de los meses del periodo *ex post.* En este es claro que todos los modelos realizan unos pronósticos adecuados para cada mes, pues los valores pronosticados se aproximan de forma aceptable a los índices reales. De esta gráfica se destaca al modelo **cuatro,** la cual está en color cian y se puede evidenciar su proximidad con los índices reales, especialmente para los primeros meses del periodo *ex post.*

|  |
| --- |
| Chart, line chart  Description automatically generated |
| **Figura XXX3.** Comparación de los pronósticos. |

**7. COMPARACIÓN DE LOS MODELOS DE REGRESIÓN CON ERRORES ARMA CONTRA UN MODELO LOCAL DE DESCOMPOSICIÓN MULTIPLICATIVA Y *LOESS* LINEAL**

Chart, histogram

Description automatically generated Previo al planteamiento de los modelos de regresión con errores estructurales ARMA, se plantearon dos modelos globales y dos locales, siendo uno de ellos *Holt-Winters* y el segundo un modelo local de descomposición multiplicativa y *loess* lineal, donde se encontró que este último conseguía unos ajustes y pronósticos aceptables para la serie temporal del índice de ventas nominal. De esta manera, con el objeto de comparar a los modelos planteados en este documento, se va a traer nuevamente a colación al mejor modelo local antes mencionado y contrastarlo con los modelos ARMA. De esta forma, se presenta inicialmente la ***figura XXXX4,*** en la cual se puede el ajuste realizado por esta gráfica.

|  |
| --- |
| **Figura XXXX4.** Gráficas de los ajustes del modelo de descomposición multiplicativa y *loess* lineal. |

Asimismo, se presenta en la ***tabla XXXX5*** la compilación de los criterios de información de Akaike (AIC) y Bayesiano (BIC) para los cuatro modelos ajustados y el modelo local que está siendo considerado.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tabla XXXX5.** Valores de los criterios de información AIC y BIC. | | | | |  |
| **Medidas** | **Modelo local** | **Modelo 1** | **Modelo 2** | **Modelo 3** | **Modelo 4** |
| AIC | 4.24 | 5.18 | 4.38 | 5.29 | 5.30 |
| BIC | 7.27 | 8.75 | 7.29 | 8.19 | 7.63 |

Y como se observa, entre los cuatro modelos de regresión con errores ARMA, el modelo **dos** es el que obtiene el menor **AIC** y **BIC,** por lo que se puede considerar al modelo **dos** como el que mejor ajustes realiza; no obstante, si tiene en cuenta también al modelo de descomposición multiplicativa y *loess* lineal, rotulado en la ***tabla XXXX5*** como “*modelo local*”, se advierte que este obtiene valores de AIC y BIC ligeramente inferiores al **modelo dos,** por lo que resulta competitivo en cuanto a ajustes se refiere. No obstante, es importante tener presente que cada uno de estos modelos supone que los errores de ajuste son un ruido blanco, de manera que vale la pena ver la ***figura XXXX5,*** donde se incluyen los gráficos de los residuales del modelo local:

|  |  |
| --- | --- |
| (a) | (b) |
| **Figura XXXX5:** Gráfica Residuales del modelo local. **(a)** ; **(b)**. | |

Entonces, es posible observar que no existen evidencias gráficas en contra de media constante e igual a cero. De igual forma, tampoco parece que exista evidencia en contra de la homocedasticidad de los errores, y tampoco es evidente a simple vista la existencia de ciclos. Ahora bien, es necesario verificar que los errores estén incorrelacionados mediante sus residuales, y si es el caso, chequear su normalidad. De esta manera, se presenta en la ***figura XXXX6*** las gráficas de la ACF y la PACF muestrales.

|  |  |
| --- | --- |
| (a) | (b) |

**Figura XXXX6.**  **(a)** Gráfica ACF sobre residuales de ajuste del modelo local. **(b)** Gráfica PACF sobre residuales de ajuste del modelo local.

De la ***figura XXXX6*** se rescata que los patrones de ambas gráficas son tipo corte, lo cual es un aspecto característico de un proceso de ruido blanco, pero se ve en la ***figura XXXX6 (a)***  que no se cumple que para k = 1, 4, 6, hecho que no se puede ignorar ya que se sabe que a los primeros valores de *k* la prueba es bastante confiable. Similarmente en la ***figura XXXX6* *(b),*** se evidencia queque para k =1, 4, 7, 8; esta evidencia es suficiente para rechazar el supuesto de que el error del modelo local es un proceso ruido blanco, y dada la evidencia en contra del supuesto de independencia, no se procede con el test de normalidad para los residuales del modelo local. Es importante tener presente que este hecho imposibilidad considerar a este modelo aunque sus medidas de error para los ajustes y los pronósticos sean mejores.

**7.1. Conclusiones**

Se encuentra que para el modelo global no se cumplía el supuesto de que los errores estructurales provenían de un proceso ruido blanco debido a que los residuos estructurales no tenían media constante y además se encontró que no eran incorrelacionados gracias a las pruebas ACF y PACF, y coherente con esto, la prueba *Ljung-Box* rechazó la hipótesis nula, además, no fue posible realizar la prueba de *Durbin-Watson* ya que esta solo se aplica a modelos que son lineales en su vector de parámetros, lo cual no se corresponde con el modelo global considerado. Sin embargo, se obtuvo que los errores estructurales eran un proceso estacionario. Por todo lo anterior, se procedió a proponer modelos exponencial polinomial de grado seis con componente estacional representada por funciones trigonométricas con errores ARMA, dado que anteriormente se obtuvo una gráfica con patrón tipo cola exponencial sinusoidal en la ACF, hecho que supone no considerar ningún modelo que proponga ) y tipo corte en la gráfica PACF con últimos dos valores que corta los límites de rechazo en . Así, se consideró pertinente postular un modelo AR(19) para los errores estructurales debido a la probabilidad de cometer error tipo I a mayor *k* en dicha prueba. Ahora bien, gracias a los resultados de la gráfica EACF se postuló un modelo ARMA(7,11), luego, se empleó la herramienta de identificación automática con la funcion auto.arima() que arrojó como propuesta un modelo ARMA(1, 2)(0, 2)[12] y otros modelos AR(p), los cuales no fueron tenidos en cuenta debido a la condición que ya se había impuesto respecto a que p debe ser mayor o igual a dos. Más tarde, algo similar sucedió con la función SelectModel(), ya que se prefirió seguir las evidencias gráficas de la PACF para definir p; por tanto el modelo tres se definió tal y como se indica en la guía de trabajo. Por último, se postuló al modelo cuatro como un ARMA(12, 10), donde solo se incluyeron los parámetros usando armasubsets().

Hablando en términos de ajuste, el modelo con valores más pequeños en los criterios de información AIC y BIC es el **modelo do**s seguido por el modelo cuatro con el segundo valor más pequeño en BIC; en términos de cumplimiento de supuestos el mejor modelo es el **modelo dos,** siendolo único reprochable de dicho modelo que en el gráfico Q-Q plot se ve evidencia que hacia el final trata de desviarse un poco de la línea recta pero rápidamente vuelve a direccionarse sobre la recta de cuantiles teóricos, lo que se considera un patrón aceptable, además de algunos datos atípicos para los cuales resulta pertinente ignorar para aceptar el supuesto de que los residuos de ajuste provienen de un R. B ~ N(0, ). Adicionalmente, en términos de pronóstico, se halló con ayuda de las medidas de error que los mejores modelos son el **dos** y el **cuatro,** dado que ambos empatan en el MAPE, el dos cuenta con menor MAPE y el cuatro con menor RMSE, si bien las diferencias son de dos y una centésima para cada medida respectivamente.

Después de todos los análisis pertinentes, postula al **modelo dos** como el mejor modelo en general. Ahora bien, resulta interesante comparar este modelo con el mejor modelo local que obtuvo en el trabajo número uno, el cual fue el filtro descomposición combinado con loess lineal, usando criterio GCV para escoger parámetro suavizamiento loess, ya que teniendo las herramientas para identificar la incorrelación de los errores estructurales de dicho modelo, se encuentra que los errores estructurales no son incorrelacionados, y por ello en términos de cumplimiento de supuestos el ganador sigue siendo el modelo **dos,** dejando así en segundo plano que los valores de los criterios de información obtenidos por este modelo lo hacían el mejor en términos de ajuste aunque con una pequeña diferencia, de igual forma que como lo había hecho en términos de predicción obteniendo los menos valores en RMSE, MAE y MAPE. De este modo, como siempre se tiene como prioridad el cumplimiento de los supuestos, se llegamos a la conclusión de que el mejor modelo para la serie de tiempo de ventas nominales es un modelo exponencial polinomial de grado seis con componente estacional representada por funciones trigonométricas y errores ARMA(7, 11), el cual tiene la desventaja que no es parsimonioso, en tanto este modelo cuenta con 37 parámetros, de tal suerte que resultaría interesante indagar algún método que permita proponer algún modelo más parsimonioso y cuyos supuestos se cumplan.

|  |  |
| --- | --- |
| **Tabla 1:** Parámetros estimados Modelos de regresión global | |
| **Tabla 1a.** Parámetros estimados en Modelo 1   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Parámetros | Estimación | Error Std |  |  | |  | 6.25081 | 0.02487 | 251.34474 | 0.00000 | |  | 0.01988 | 0.00129 | 15.39242 | 0.00000 | |  | -1.290×10-4 | 0.00002 | -6.54274 | 0.00000 | |  | 3.168×10-7 | 8.526×10-8 | 3.71609 | 0.00029 | |  | -0.13167 | 0.01574 | -8.36568 | 0.00000 | |  | -0.01853 | 0.01574 | -1.17723 | 0.24105 | |  | 0.01446 | 0.01574 | 0.91883 | 0.35972 | | (escala log); AIC= 8995.101 BIC= 10345.524 | | | | | | **Tabla 1c.** Parámetros estimados en Modelo 2   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Parámetros | Estimación | Error Std |  |  | |  | 6.25467 | 0.04752 | 131.63476 | 0.00000 | |  | 0.02015 | 0.00209 | 9.62945 | 0.00000 | |  | -1.323×10-4 | 0.00003 | -4.75668 | 0.00000 | |  | 3.301×10-7 | 1.101×10-7 | 2.99833 | 0.00320 | |  | -0.14307 | 0.01832 | -7.80790 | 0.00000 | |  | -0.02956 | 0.01719 | -1.71946 | 0.08768 | |  | 0.00512 | 0.01686 | 0.30392 | 0.76163 | | , AIC= 8956.667, BIC= 10301.320 | | | | | |
| **Tabla 1b.** Parámetros estimados en Modelo 1b   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Parámetros | Estimación | Error Std |  |  | |  | 6.21688 | 0.02275 | 273.28662 | 0.00000 | |  | 0.01988 | 0.00129 | 15.39242 | 0.00000 | |  | -1.290×10-4 | 0.00002 | -6.54274 | 0.00000 | |  | 3.168×10-7 | 8.526×10-8 | 3.71609 | 0.00029 | |  | -0.07307 | 7.818e-03 | -9.34583 | 0.00000 | |  | 0.00926 | 0.00787 | 1.17723 | 0.24105 | |  | 0.02467 | 0.00555 | 4.44922 | 0.00002 | | (escala log); AIC= 8995.101 BIC= 10345.524 | | | | | | **Tabla 1d.** Parámetros estimados en Modelo 2b   |  |  |  |  |  | | --- | --- | --- | --- | --- | | Parámetros | Estimación | Error Std |  |  | |  | 6.21279 | 0.04641 | 133.85500 | 0.00000 | |  | 0.02015 | 0.00209 | 9.62945 | 0.00000 | |  | -1.323×10-4 | 0.00003 | -4.75668 | 0.00000 | |  | 3.301×10-7 | 1.101×10-7 | 2.99832 | 0.00320 | |  | -0.07410 | 0.00907 | -8.16517 | 0.00000 | |  | 0.01478 | 0.00860 | 1.71946 | 0.08768 | |  | 0.02709 | 0.00625 | 4.33595 | 0.00003 | | , AIC= 8956.667, BIC= 10301.320 | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Tabla 2.** Pronósticos puntuales y por I.P del 95% de confianza | | | | | | | | | | | | |
|  | Modelo 1 | | | Modelo 1b | | | Modelo 2 | | | Modelo 2b | | |
| Período | 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 1b | Modelo 2 | Modelo 2b |
| RMSE | 160.31 | 160.31 | 154.47 | 154.47 |
| MAE | 130.98 | 130.98 | 124.55 | 124.55 |
| MAPE (%) | 7.10 | 7.10 | 6.74 | 6.74 |
| Amplitud. Media I.P | 464.7828 | 464.7828 | NA | NA |
| Cobertura (%) I.P | 100% | 100% | NA | NA |

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

|  |  |
| --- | --- |
| (a) | (b) |
| (c) | (d) |
| **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-cubico estacional usando variables trigonométricas en la representación de la componente estacional, en frecuencias , ), 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 , ), respectivamente,

**,** . (1)

**,** . (2)

**,** . (3)

**,** . (4)

**5.4.3.2 Ecuaciones como parte de una tabla**

|  |
| --- |
| **Tabla 4.** Ecuaciones de los modelos propuestos |
| **Modelo 1**  **,** |
| **Modelo 1b**  **,** |
| **Modelo 2**  **,** |
| **Modelo 2b**  **,** |