

6. Pronósticos para la validación cruzada

Los pronósticos realizados para el periodo ex post para llevar a cabo la 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, únicamente se presentará la ecuación general de los pronósticos para el Modelo 3.

Tabla 4. Ecuación de pronóstico con origen $n = 239$

Modelo 3

$$\hat{Y}_{239}(L) = \exp[3.55 + 2.182 \times 10^{-3}(239 + L) - 2.801 \times 10^{-4}(4239 + L)^2 - 4.481 \times 10^{-6}(239 + L)^3 + 2.919 \times 10^{-8}(239 + L)^4 - 8.674 \times 10^{-11}(239 + L)^5$$

$$+ 9.781 \times 10^{-14}(239 + L)^6 - 4.125 \times 10^{-2} \sin\left(\frac{\pi}{6}(239 + L)\right) + 1.473 \times 10^{-2} \cos\left(\frac{\pi}{6}(239 + L)\right) - 2.968 \times 10^{-2} \sin\left(\frac{\pi}{3}(239 + L)\right)$$

$$+ 1.357 \times 10^{-2} \cos\left(\frac{\pi}{3}(239 + L)\right) - 1.925 \times 10^{-2} \sin\left(\frac{\pi}{2}(239 + L)\right) + 2.035 \times 10^{-2} \cos\left(\frac{\pi}{2}(239 + L)\right) - 1.534 \times 10^{-2} \sin\left(\frac{2\pi}{3}(239 + L)\right)$$

$$+ 2.3456 \times 10^{-2} \cos\left(\frac{2\pi}{3}(239 + L)\right) + 4.155 \times 10^{-2} \sin\left(\frac{5\pi}{6}(239 + L)\right) + 2.378 \times 10^{-2} \cos\left(\frac{5\pi}{6}(239 + L)\right) + \hat{E}_{239}(L)]$$

Donde

$$\hat{E}_{239}(L) = 5.10 \times 10^{-1} \hat{E}_{239}(L - 1) + 8.65 \times 10^{-1} \hat{E}_{239}(L - 2) - 5.91 \times 10^{-1} \hat{E}_{239}(L - 3) - 3.423 \times 10^{-1} \hat{a}_{239}(L - 1) - 5.80 \times 10^{-1} \hat{a}_{239}(L - 2) + 6.98 \times 10^{-1} \hat{a}_{239}(L - 3)$$

$$- 3.423 \times 10^{-1} \hat{a}_{239}(L - 4) + 2.50 \times 10^{-2} \hat{a}_{239}(L - 5) + 1.19 \times 10^{-1} \hat{a}_{239}(L - 6) - 2.11 \times 10^{-1} \hat{a}_{239}(L - 7) + 1.77 \times 10^{-1} \hat{a}_{239}(L - 8) + 2.37 \times 10^{-1} \hat{a}_{239}(L - 9)$$

$$+ 9.20 \times 10^{-2} \hat{E}_{239}(L - 12) + (9.20 \times 10^{-2} \times 5.10 \times 10^{-1}) \hat{E}_{239}(L - 13) + (9.20 \times 10^{-2} \times 8.65 \times 10^{-1}) \hat{E}_{239}(L - 14) + (9.20 \times 10^{-2} \times -5.91 \times 10^{-1}) \hat{E}_{239}(L - 15)$$

En las anteriores ecuaciones se tiene que

$$\hat{E}_{239}(L - j) = \begin{cases} \text{Residuo estructural } \hat{E}_{239+L-j} & \text{si } L - j \leq 0 \\ \text{pronóstico de los ciclo } L - j \text{ periodos después de } t = 239, & \text{si } L - j > 0 \end{cases}$$

$$\hat{a}_{239}(L - i) = \begin{cases} \text{Residuo de ajuste } \hat{a}_{239+L-i} & \text{si } L - j \leq 0 \\ 0, \text{ el pronóstico del R.B } L - i \text{ periodos después de } t = 239, & \text{si } L - j > 0 \end{cases}$$

Tabla 3. Precisión de los Pronósticos puntuales

Medidas	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo local
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 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 mientras que según 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.

Comparación de los pronósticos para el periodo ex post

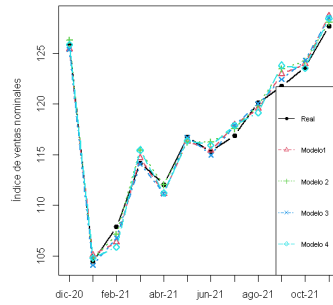
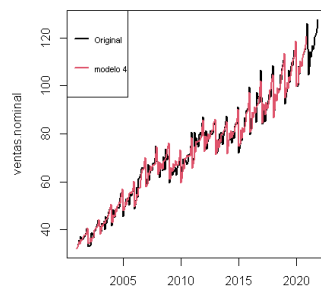
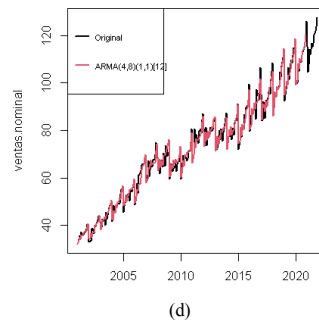
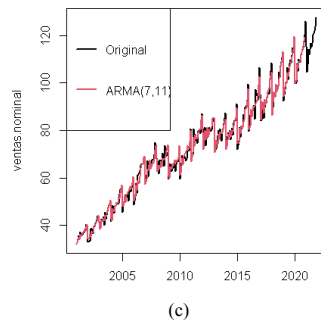
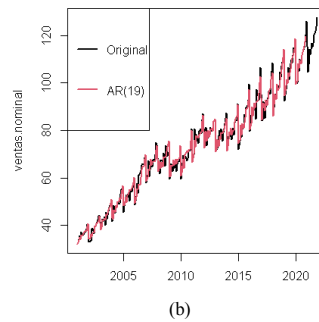
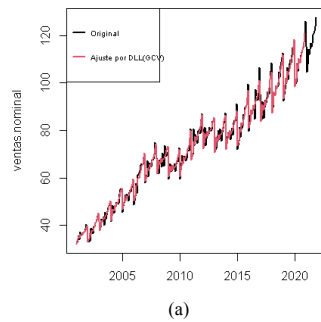


Figura 1: Comparación de los pronósticos.

7. Comparación de modelos de regresión con errores ARMA vs. mejor local.



(e)

Figura 1: Gráficas de los ajustes. (a) En mejor modelo local; (b) En modelo 1; (c) En modelo 2; (d) En modelo 3 ; (e) En modelo 4

Tabla 3. 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

Según los criterios de información el modelo que tiene mejor ajuste es el modelo **uno**, no obstante como se puede observar en la **Figura** los demás modelos presentan resultados aceptables; Vale la pena recordar en este caso los gráfico de los residuales del mejor modelo local

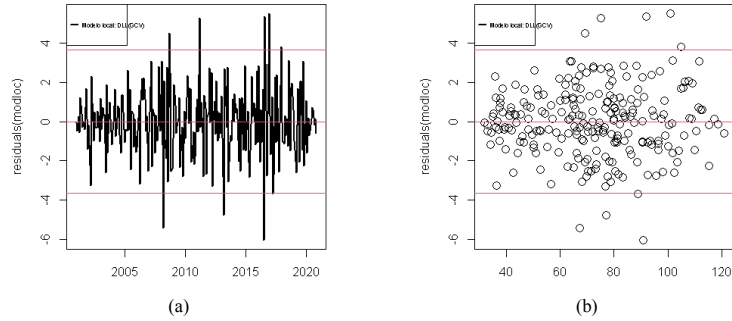


Figura 1: Gráfica Residuales del modelo local. (a) \hat{E}_t vs t ; (b) \hat{E}_t vs \hat{Y}

No existen evidencias gráficas en contra de media constante e igual a cero de igual forma que para el supuesto de varianza constante, tampoco es evidente a simple vista ciclos, podemos ahora que tenemos la herramienta probar si los errores de este modelo son un proceso R.B, para ello verifiquemos si el proceso es ergódico con las siguientes gráficas

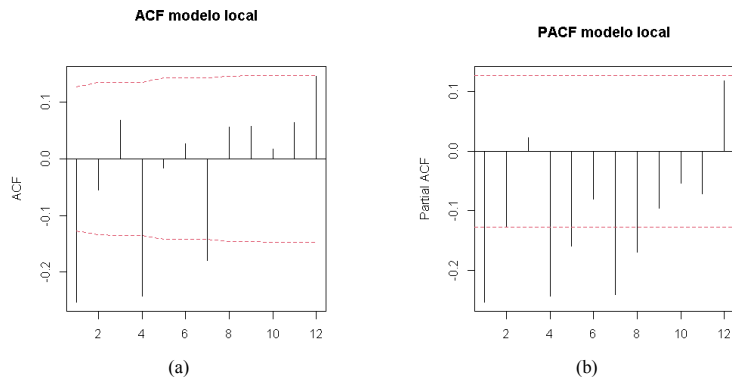


Figura 1: (a) Gráfica ACF sobre residuales de ajuste del modelo local; (b) Gráfica PACF sobre residuales de ajuste del modelo local

Los patrones de ambas gráficas son tipo corte un aspecto característico de un proceso R.B pero se ve en la **Figura (a)** que no se cumple $\rho(k) = \text{Corr}(E_t, E_{t+k}) = 0$ para $k = 1, 4, 6$ hecho que no se puede ignorar ya que sabemos que a los primeros valores de k la prueba es bastante confiable similarmente en la **Figura (b)** que $\phi_{kk} = \text{Corr}(E_t, E_{t+k} | E_{t+1}, \dots, E_{t+k-1}) \neq 0$ para $k = 1, 4, 7, 8$; esta evidencia es suficiente para rechazar el supuesto de que el error del modelo local es un proceso R.B por tanto se considera innecesaria realizar la prueba de normalidad.

4.1 Conclusiones.

Encontramos que para el modelo global no se cumplía el supuesto de que los errores estructurales provenían de un proceso R.B debido a que los residuos estructurales no tenían media constante y además encontramos que no eran incorrelacionados gracias a las pruebas ACF y PACF, coherente con esto la prueba LJung Box falló a favor de la hipótesis alternativa, no realizamos la prueba Durbin-Watson ya que esta solo se aplica a modelos que son lineales en su vector de parámetros lo que no es nuestro caso sin embargo obtuvimos que los errores estructurales eran un proceso estacionario, por todo lo anterior procedimos a proponer modelos exponencial polinomial de grado seis con componente estacional representada por funciones trigonométricas con errores ARMA; dado que anteriormente obtuvimos una gráfica con patrón tipo cola exponencial sinusoidal en la ACF (Este hecho ya nos predispone a no considerar ningún modelo que proponga $p < 2$) y tipo corte en la gráfica PACF con últimos dos valores que corta los límites de rechazo en $k = 19, 33$ consideramos 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 gracias a los resultados de la gráfica EACF postulamos un modelo ARMA(7,11); continuamos usando la herramienta de identificación automática con la función `auto.arima()` que nos arroja como propuesta un modelo ARMA(1, 2)(0, 2)[12] y otros modelos AR(p) los cuales no consideramos debido a la condición que ya habíamos inpuesto en la decíamos que nuestro p debe ser mayor o igual a dos algo similar sucedió con la función `SelectModel()` ya que preferimos seguir las evidencias gráficas de la PACF para definir p ; por tanto el modelo tres lo definimos tal y como se indica en la guía de trabajo por ultimo de igual manera definimos el modelo 4 como un ARMA(12, 10) donde solo incluimos los parámetros $\phi_1, \phi_2, \phi_3, \phi_7, \phi_{12}, \theta_4, \theta_9$ y θ_{10} usando `armasubsets()`.

Hablando en términos de ajuste el modelo con valores más pequeños en los criterios de información AIC y BIC es el **modelo dos** 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** lo único reprochable de dicho modelo es que en el gráfico Q-Q plot se ve como al final trata de desviarse un poco de la línea recta pero rápidamente vuelve a direccionarse lo que consideramos un patrón aceptable, además de algunos datos atípicos que consideramos pertinente ignorar para aceptar el supuesto de que los residuos de ajuste provienen de un $R. B \sim N(0, \sigma_a^2)$.

En términos de pronóstico el mejor modelo es el modelo cuatro seguido por el modelo dos, después de todos lo análisis pertinentes postulamos el **modelo dos** como el mejor modelo en general. Ahora resulta interesante comparar este modelo con el mejor modelo local que obtuvimos en el trabajo número uno dicho modelo fue el Filtro descomposición combinado con loess lineal, usando criterio GCV para escoger parámetro suavizamiento loess, ya que tenemos las herramientas para identificar la incorrelación de los errores estructurales de dicho modelo con estas herramientas llegamos a la conclusión de 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 por poco, de igual forma que como lo había hecho en términos de predicción obteniendo los menos valores en RMSE, MAE y MAPE; Como siempre ponemos como prioridad el cumplimiento de los supuestos 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), la crítica que podemos hacer a nuestro mejor modelo es que este cuenta con 37 parámetros, como siempre preferimos modelos parsimoniosos sería prudente postular un modelo que no contengan aquellos parámetros tanto de tendencia, estacionalidad y parámetros en el modelo del error estructural que no resultan significativos.

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 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.

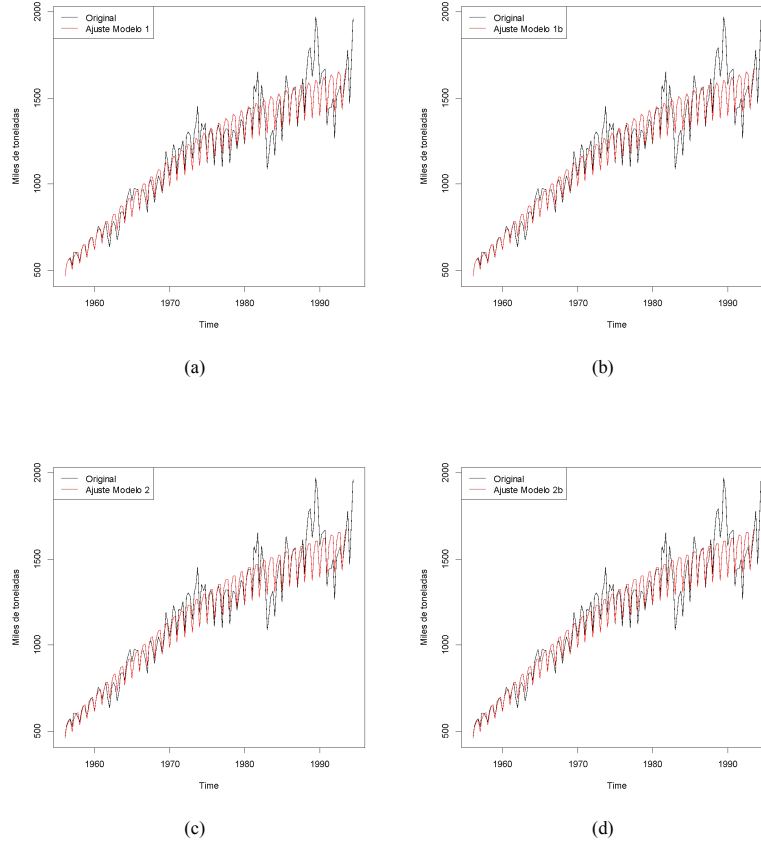


Figura 1: Gráficas de los ajustes. (a) En modelo 1; (b) En modelo 1b; (c) En modelo 2; (d) En modelo 2b

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,

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

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

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

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

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

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