

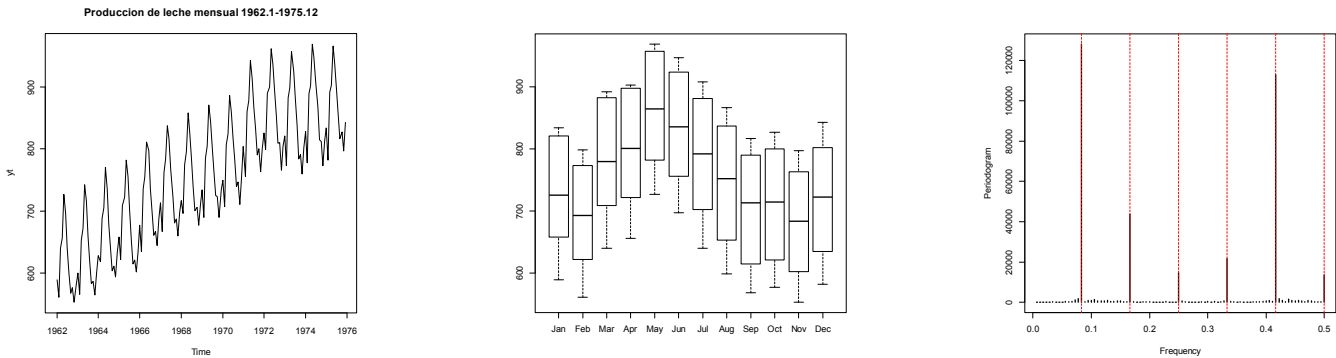
TALLER # 9: AJUSTE DE UNA SERIE POR REGRESIÓN CON ERRORES ARMA

Los siguientes datos representan la producción de leche mensual en libras por vaca, desde Enero de 1962, hasta Diciembre de 1975 (N=168 obs.)

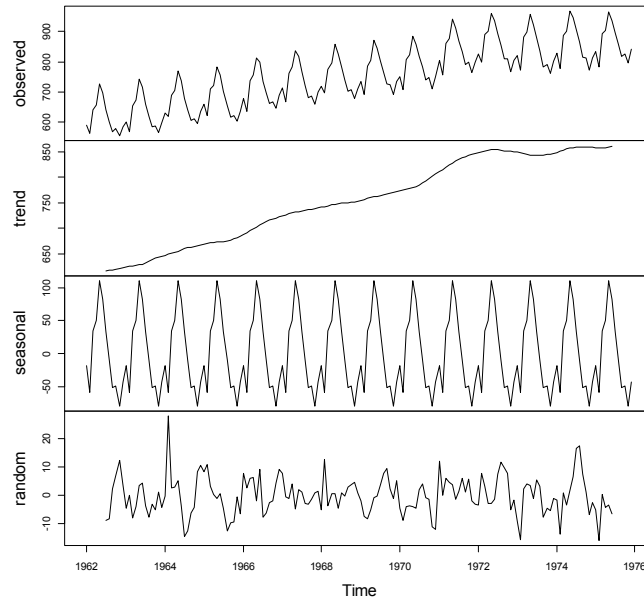
| año | ene | feb | mar | abr | may | jun | jul | ago | sep | oct | nov | dic |
|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1962 | 589 | 561 | 640 | 656 | 727 | 697 | 640 | 599 | 568 | 577 | 553 | 582 |
| 1963 | 600 | 566 | 653 | 673 | 742 | 716 | 660 | 617 | 583 | 587 | 565 | 598 |
| 1964 | 628 | 618 | 688 | 705 | 770 | 736 | 678 | 639 | 604 | 611 | 594 | 634 |
| 1965 | 658 | 622 | 709 | 722 | 782 | 756 | 702 | 653 | 615 | 621 | 602 | 635 |
| 1966 | 677 | 635 | 736 | 755 | 811 | 798 | 735 | 697 | 661 | 667 | 645 | 688 |
| 1967 | 713 | 667 | 762 | 784 | 837 | 817 | 767 | 722 | 681 | 687 | 660 | 698 |
| 1968 | 717 | 696 | 775 | 796 | 858 | 826 | 783 | 740 | 701 | 706 | 677 | 711 |
| 1969 | 734 | 690 | 785 | 805 | 871 | 845 | 801 | 764 | 725 | 723 | 690 | 734 |
| 1970 | 750 | 707 | 807 | 824 | 886 | 859 | 819 | 783 | 740 | 747 | 711 | 751 |
| 1971 | 804 | 756 | 860 | 878 | 942 | 913 | 869 | 834 | 790 | 800 | 763 | 800 |
| 1972 | 826 | 799 | 890 | 900 | 961 | 935 | 894 | 855 | 809 | 810 | 766 | 805 |
| 1973 | 821 | 773 | 883 | 898 | 957 | 924 | 881 | 837 | 784 | 791 | 760 | 802 |
| 1974 | 828 | 778 | 889 | 902 | 969 | 947 | 908 | 867 | 815 | 812 | 773 | 813 |
| 1975 | 834 | 782 | 892 | 903 | 966 | 937 | 896 | 858 | 817 | 827 | 797 | 843 |

1. Basado en las siguientes graficas

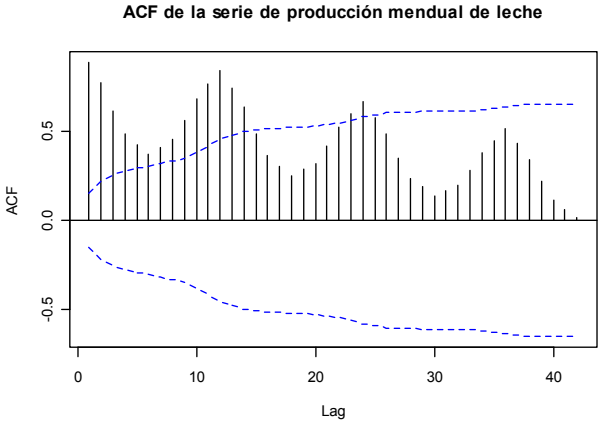
- ¿Qué tipo de tendencia se aplica a esta serie (lineal, cuadrática, cubica)?
- ¿Existe un patrón estacional? ¿Cuál es su período? ¿Qué tipo de componentes tiene la serie (aditiva, multiplicativa)?



Decomposition of additive time series



2. Analice la ACF obtenida para los datos de la serie en su escala original



- a) Concluya acerca de la estacionariedad de la serie de producción de leche.
 - b) Analice si existe algún patrón en la ACF que refleje la presencia de estacionalidad.
 - c) Considera que las conclusiones obtenidas con base en la ACF sobre la estacionariedad eran de esperarse de acuerdo a la estructura de la serie? ¿Por qué?
3. Se ajustó un modelo de regresión (aquí denominado como modelo 3) de tendencia cúbica y estacionalidad con indicadores con nivel de referencia diciembre, usando los primeros n=156 datos. Escriba la ecuación teórica y la ecuación ajustada con base en las estimaciones que se muestran a continuación. Analice la gráfica del ajuste
4. Interprete los coeficientes estacionales estimados.

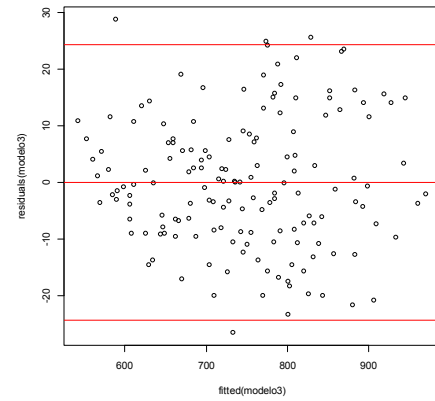
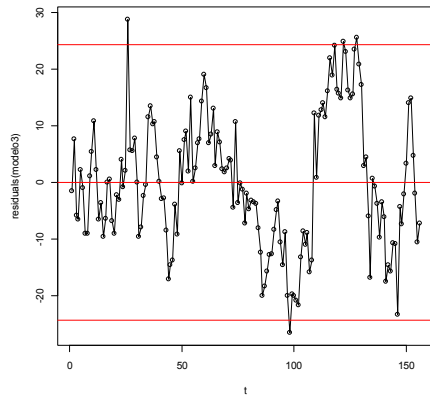
Salida R ajuste MODELO 3.

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|------------|------------|---------|--------------|
| (Intercept) | 5.641e+02 | 5.296e+00 | 106.505 | < 2e-16 *** |
| t | 1.149e+00 | 2.212e-01 | 5.193 | 7.09e-07 *** |
| t2 | 1.425e-02 | 3.269e-03 | 4.359 | 2.50e-05 *** |
| t3 | -7.107e-05 | 1.369e-05 | -5.191 | 7.16e-07 *** |
| mesJan | 2.527e+01 | 4.798e+00 | 5.267 | 5.07e-07 *** |
| mesFeb | -1.308e+01 | 4.795e+00 | -2.729 | 0.00717 ** |
| mesMar | 7.826e+01 | 4.791e+00 | 16.333 | < 2e-16 *** |
| mesApr | 9.360e+01 | 4.789e+00 | 19.547 | < 2e-16 *** |
| mesMay | 1.546e+02 | 4.786e+00 | 32.310 | < 2e-16 *** |
| mesJun | 1.265e+02 | 4.784e+00 | 26.449 | < 2e-16 *** |
| mesJul | 7.628e+01 | 4.782e+00 | 15.950 | < 2e-16 *** |
| mesAug | 3.387e+01 | 4.781e+00 | 7.085 | 6.08e-11 *** |
| mesSep | -8.605e+00 | 4.780e+00 | -1.800 | 0.07393 . |
| mesOct | -5.384e+00 | 4.779e+00 | -1.127 | 0.26176 |
| mesNov | -3.623e+01 | 4.778e+00 | -7.583 | 4.13e-12 *** |
| --- | | | | |

Residual standard error: 12.18 on 141 degrees of freedom
Multiple R-squared: 0.9866, Adjusted R-squared: 0.9852
F-statistic: 740.2 on 14 and 141 DF, p-value: < 2.2e-16

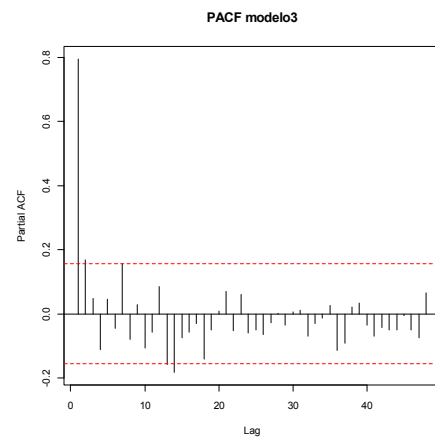
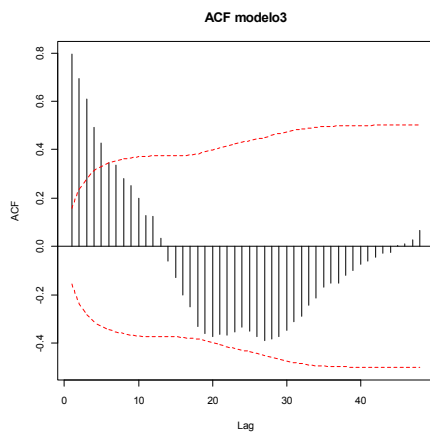
Gráfica del ajuste MODELO 3

5. Usando los residuales del ajuste del modelo 3
- a) Analice las gráficas de residuales vs. tiempo y vs. valores ajustados ¿Qué podemos concluir en cada caso?
 - b) Realice los tests de la ACF y PACF muestral, las pruebas de Box-Pierce, Lung-Box y Durbin Watson de orden 1, para los errores estructurales del modelo 3. Escriba las respectivas hipótesis, estadístico de prueba y criterio de rechazo, y concluya acerca del supuesto de independencia de los errores estructurales del modelo 3.
 - c) ¿Qué se concluye acerca de la estacionariedad de la serie de los errores del modelo 3? Indique con base en cuál resultado se evalúa tal condición. ¿Son los errores Ruido Blanco? ¿Por qué?
 - d) En caso de discordancia en la conclusión de incorrelación de los errores, ¿con cuál prueba se tiene menor probabilidad de equivocarse? ¿Se puede llegar a la misma conclusión con DW que con la ACF y/o Lung-Box?



| Modelo | TEST DURBIN-WATSON PARA AUTOCORRELACIÓN DE ORDEN 1 | | | | |
|--------|--|--------------|-----------------|----------|----------|
| | lag | rho estimado | Estadístico D-W | VP rho>0 | VP rho<0 |
| 3 | 1 | 0.7952471 | 0.4068884 | 0 | 1 |

| Box-Pierce y Ljung-Box modelo3 | | | | | | | |
|--------------------------------|----------|----|----|-----------|----------|----|----|
| Box-Pierce | | | | Ljung-Box | | | |
| m | QBP | df | VP | m | QLB | df | VP |
| 6 | 316.5142 | 6 | 0 | 6 | 326.0681 | 6 | 0 |
| 12 | 367.6739 | 12 | 0 | 12 | 380.8485 | 12 | 0 |
| 18 | 404.5112 | 18 | 0 | 18 | 422.7512 | 18 | 0 |
| 24 | 525.7595 | 24 | 0 | 24 | 565.1155 | 24 | 0 |
| 30 | 654.1543 | 30 | 0 | 30 | 722.9998 | 30 | 0 |
| 36 | 706.8380 | 36 | 0 | 36 | 790.5185 | 36 | 0 |
| 42 | 716.0286 | 42 | 0 | 42 | 802.8575 | 42 | 0 |



6. A continuación se presentan resultados para la identificación y ajuste de diferentes modelos con ciclos ARMA para los errores estructurales del modelo 3. De acuerdo a estos resultados y patrones observados en ACF determine cuáles modelos para los errores estructurales pueden ser más factibles, dé las ecuaciones del modelo de regresión con errores ARMA correspondientes

```
R Console (64-bit)
Archivo Editar Misc Paquetes Ventanas Ayuda

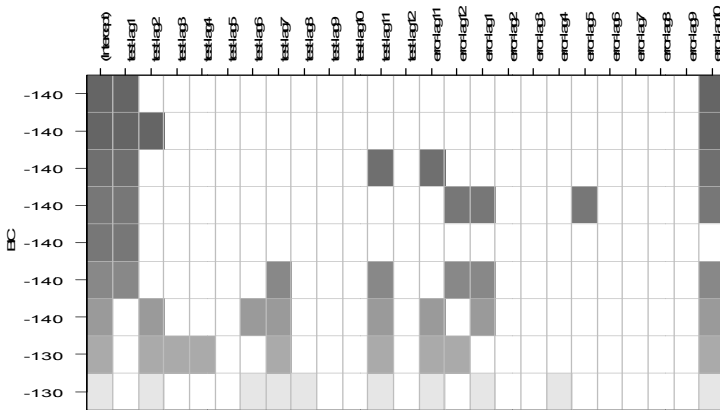
> #Identificación con criterio AIC
> auto.arima(residuals(modelo3),ic="aic")
Series: residuals(modelo3)
ARIMA(3,0,2) with zero mean

Coefficients:
      ar1      ar2      ar3      ma1      ma2
      1.0886 -1.0592  0.6885 -0.4055  0.9366
s.e.      0.0817  0.1139  0.0781  0.0535  0.0647

sigma^2 estimated as 46.73: log likelihood=-519.68
AIC=1051.36 AICc=1051.92 BIC=1069.66
> #lo siguiente puede arrojarse un modelo SARMA
> et3=ts(residuals(modelo3),frequency=12, start=c(1962,1)) #serie de tiempo residuos modelo 3
> auto.arima(et3,ic="aic")
Series: et3
ARIMA(4,0,0)(2,0,0)[12] with zero mean

Coefficients:
      ar1      ar2      ar3      ar4      sar1      sar2
      0.6633  0.1752  0.1270 -0.1354  0.2220  0.0676
s.e.      0.0790  0.0966  0.0555  0.0799  0.0836  0.0912

sigma^2 estimated as 45.94: log likelihood=-517.86
AIC=1049.73 AICc=1050.49 BIC=1071.08
> |
```



```
R Console (64-bit)
Archivo Editar Misc Paquetes Ventanas Ayuda

> SelectModel(et3, lag.max = 36, Criterion="AIC",ARModel = "AR")
p AIC-Exact AIC-Approx
1 2 609.7097 -153.8402
2 14 610.8201 -155.1390
3 3 611.4309 -154.2548
> SelectModel(et3, lag.max = 36, Criterion="BIC",ARModel = "AR")
p BIC-Exact BIC-Approx
1 1 618.4765 -146.4076
2 2 618.8593 -141.6408
3 3 623.6304 -139.0055
> |
```

```
R Console (64-bit)
Archivo Editar Misc Paquetes Ventanas Ayuda

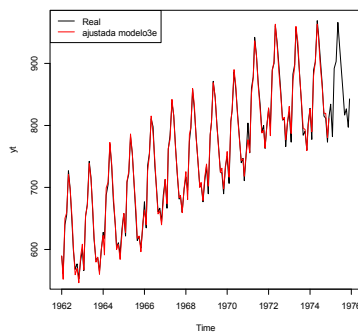
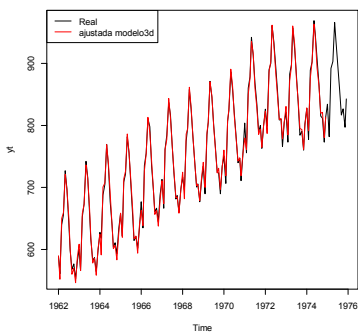
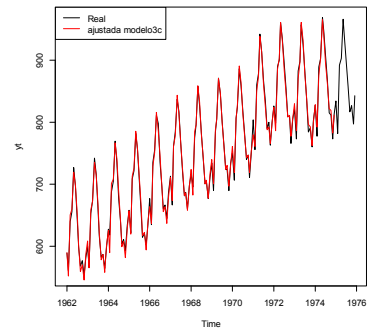
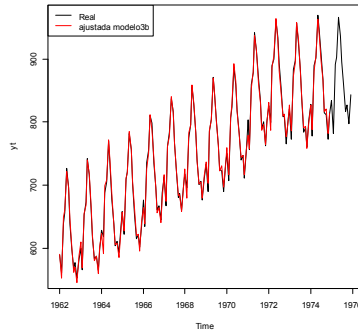
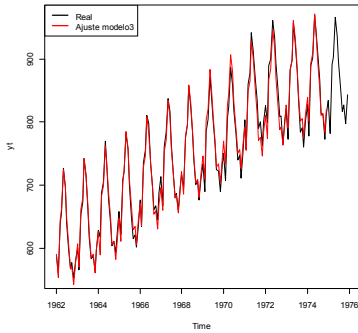
> eacf(residuals(modelo3),ar.max=24,ma.max=24)
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
0 x x x x x x x x x x x x x x x x x x x x x x x x x x
1 x o o o o x o o o o o o o o o o o o o o o o o o o o
2 x o o o o o o o o o o o o o o o o o o o o o o o o o
3 x x o o o o o o o o o o x o o o o o o o o o o o o o
4 x x x o o o o o o o o o x o o o o o o o o o o o o o
5 x o x o o o o o o o o o x o o o o o o o o o o o o o
6 x o x x o o o o o o o o o o o o o o o o o o o o o o
7 x o x x o o o o o o o o o o o o o o o o o o o o o o
8 x o x x o o o o o o o o o o o o o o o o o o o o o o
9 x x x x o o o o o o o o o o o o o o o o o o o o o o
10 x x x o o o o o o o o o o o o o o o o o o o o o o o
11 x o x o o o o o o o o o o o o o o o o o o o o o o o
12 x o x x x x x x x x x x x x x x x x x x x x x x x
13 x o o o x x x x x x x x x x x x x x x x x x x x x
14 x o o o x x x x x x x x x x x x x x x x x x x x x
15 x o o o x x x x x x x x x x x x x x x x x x x x x
16 x o o o x x x x x x x x x x x x x x x x x x x x x
17 o x x x o o o o o o o o o o o o o o o o o o o o o o
18 x x x x x x x x x x x x x x x x x x x x x x x x x
19 x o x o x x x x x x x x x x x x x x x x x x x x x
20 o x x o o x x x x x x x x x x x x x x x x x x x x
21 x x o o o x x x x x x x x x x x x x x x x x x x x
22 x x o x x x x x x x x x x x x x x x x x x x x x x
23 x o x x o o o o o o o o o o o o o o o o o o o o o
24 x o x x x x x x x x x x x x x x x x x x x x x x x
> |
```

7. Ajuste modelos de regresión con errores ARMA identificados y realice lo siguiente.
- Compare ajustes: Gráficas y medidas de ajuste (calcule AIC y BIC como $\exp(C_n^*(p))$)
 - Valide supuestos sobre los errores de ajuste
 - Escriba en cada caso la ecuación de los pronósticos. Compare predicciones (MAE, MAPE, RMSE, amplitud media de los I.P y cobertura alcanzada).
 - Finalmente seleccione el mejor modelo.

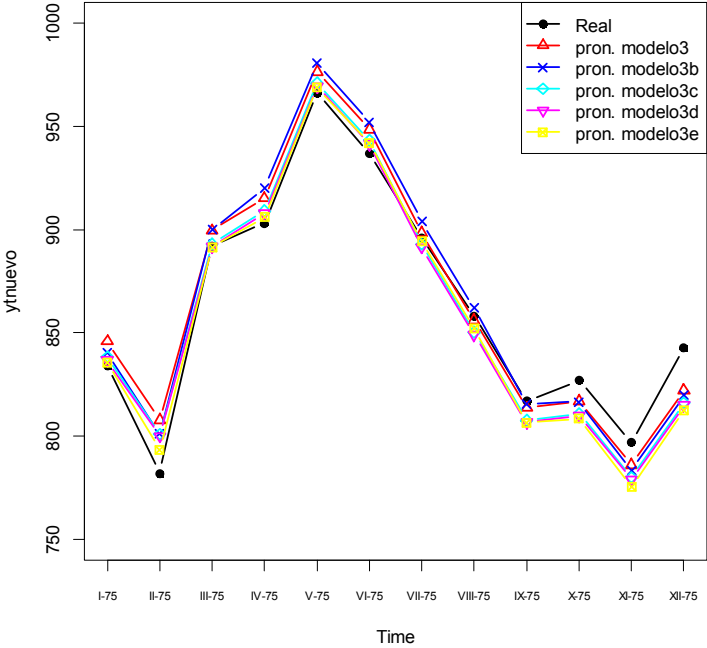
Tablas de parámetros estimados modelos con errores ARMA

| Resultados R Estimaciones Modelo3b (con errores AR(14)) | | | | | Resultados R Estimaciones Modelo3c (con errores ARMA(2,10) pero sólo $\phi_1, \phi_2, \theta_{10} \neq 0$) | | | | |
|---|-------------|------------|----------|---------------|---|-------------|------------|----------|---------------|
| | Estimate | Std. Error | t value | Pr(> t) | | Estimate | Std. Error | t value | Pr(> t) |
| ar1 | 6.6476e-01 | 7.7579e-02 | 8.5688 | 3.016e-14 *** | ar1 | 6.5453e-01 | 7.8910e-02 | 8.2947 | 8.762e-14 *** |
| ar2 | 1.4642e-01 | 9.5552e-02 | 1.5323 | 0.127934 | ar2 | 1.7019e-01 | 7.9743e-02 | 2.1343 | 0.034588 * |
| ar3 | 1.3669e-01 | 9.2722e-02 | 1.4742 | 0.142897 | ma10 | 2.9685e-02 | 9.4008e-02 | 0.3158 | 0.752652 |
| ar4 | -1.7434e-01 | 9.4561e-02 | -1.8437 | 0.067563 . | intercept | 5.6504e+02 | 1.0124e+01 | 55.8117 | < 2.2e-16 *** |
| ar5 | 2.0761e-02 | 9.6746e-02 | 0.2146 | 0.830427 | t | 1.0665e+00 | 5.0968e-01 | 2.0925 | 0.038228 * |
| ar6 | -1.2574e-01 | 9.4255e-02 | -1.3340 | 0.184579 | t2 | 1.5885e-02 | 6.6855e-03 | 2.3761 | 0.018870 * |
| ar7 | 2.3769e-01 | 9.7533e-02 | 2.4370 | 0.016194 * | t3 | -7.9217e-05 | 3.4152e-05 | -2.3195 | 0.021835 * |
| ar8 | -5.8585e-02 | 9.8314e-02 | -0.5959 | 0.552305 | Jan | 2.4714e+01 | 2.0974e+00 | 11.7834 | < 2.2e-16 *** |
| ar9 | 6.7380e-02 | 9.8339e-02 | 0.6852 | 0.494477 | Feb | -1.3566e+01 | 2.4514e+00 | -5.5342 | 1.522e-07 *** |
| ar10 | -1.0840e-01 | 9.7444e-02 | -1.1124 | 0.268065 | Mar | 7.7843e+01 | 2.7720e+00 | 28.0816 | < 2.2e-16 *** |
| ar11 | -5.1806e-02 | 1.0170e-01 | -0.5094 | 0.611344 | Apr | 9.3279e+01 | 2.9574e+00 | 31.5409 | < 2.2e-16 *** |
| ar12 | 2.5126e-01 | 9.9906e-02 | 2.5150 | 0.013151 * | May | 1.5437e+02 | 3.0573e+00 | 50.4921 | < 2.2e-16 *** |
| ar13 | -2.5256e-02 | 1.0107e-01 | -0.2499 | 0.803066 | Jun | 1.2631e+02 | 3.0844e+00 | 40.9496 | < 2.2e-16 *** |
| ar14 | -2.1922e-01 | 7.5196e-02 | -2.9153 | 0.004201 ** | Jul | 7.6090e+01 | 3.0430e+00 | 25.0049 | < 2.2e-16 *** |
| intercept | 5.6467e+02 | 9.2036e+00 | 61.3535 | < 2.2e-16 *** | Aug | 3.3702e+01 | 2.9282e+00 | 11.5097 | < 2.2e-16 *** |
| t | 1.1043e+00 | 4.8792e-01 | 2.2632 | 0.025319 * | Sep | -8.7759e+00 | 2.7372e+00 | -3.2062 | 0.001672 ** |
| t2 | 1.4662e-02 | 7.2780e-03 | 2.0146 | 0.046056 * | Oct | -5.5089e+00 | 2.4030e+00 | -2.2925 | 0.023390 * |
| t3 | -7.2044e-05 | 2.8151e-05 | -2.5592 | 0.011664 * | Nov | -3.6307e+01 | 2.0259e+00 | -17.9212 | < 2.2e-16 *** |
| Jan | 2.5502e+01 | 2.5906e+00 | 9.8443 | < 2.2e-16 *** | | | | | |
| Feb | -1.2848e+01 | 3.0021e+00 | -4.2797 | 3.653e-05 *** | | | | | |
| Mar | 7.8710e+01 | 3.2124e+00 | 24.5017 | < 2.2e-16 *** | | | | | |
| Apr | 9.4762e+01 | 3.5387e+00 | 26.7786 | < 2.2e-16 *** | | | | | |
| May | 1.5584e+02 | 3.7715e+00 | 41.3189 | < 2.2e-16 *** | | | | | |
| Jun | 1.2721e+02 | 4.1037e+00 | 30.9983 | < 2.2e-16 *** | | | | | |
| Jul | 7.6831e+01 | 3.8557e+00 | 19.9266 | < 2.2e-16 *** | | | | | |
| Aug | 3.4443e+01 | 3.6954e+00 | 9.3203 | 4.648e-16 *** | | | | | |
| Sep | -7.9644e+00 | 3.3544e+00 | -2.3743 | 0.019080 * | | | | | |
| Oct | -4.7672e+00 | 3.1077e+00 | -1.5340 | 0.127523 | | | | | |
| Nov | -3.5807e+01 | 2.5254e+00 | -14.1787 | < 2.2e-16 *** | | | | | |

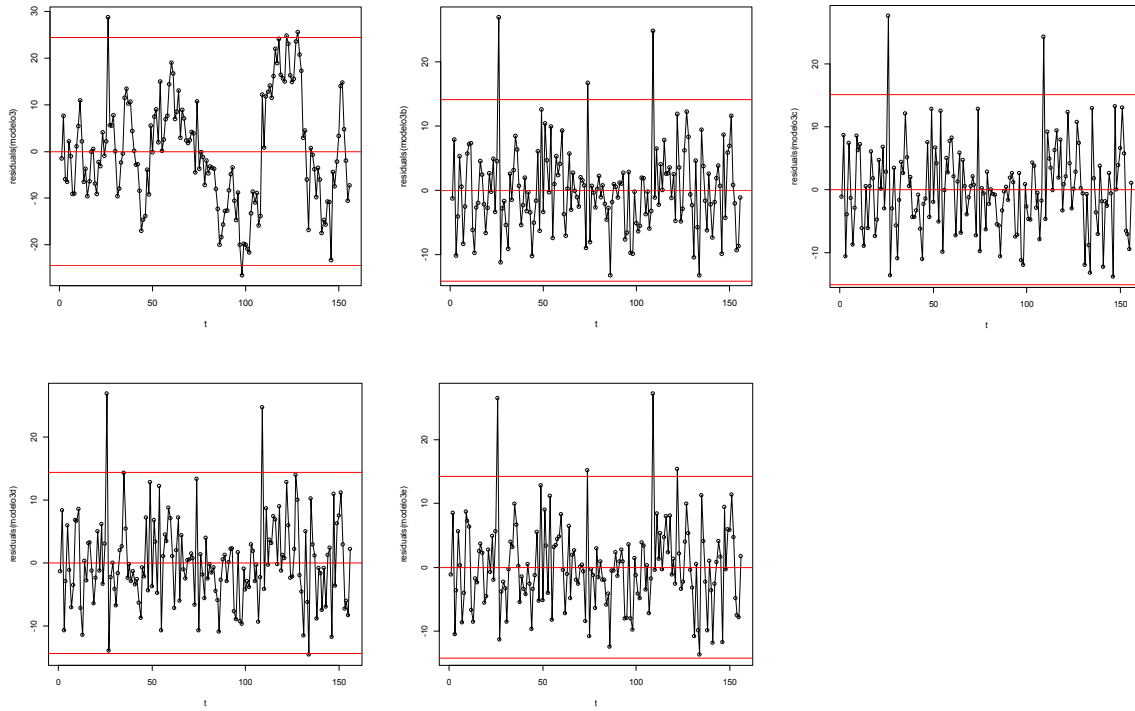
| Resultados R Estimaciones Modelo3d (con errores ARMA(3,2)) | | | | | Resultados R Estimaciones Modelo3e (con errores ARMA(4,0)(2,0)[12]) | | | | |
|--|-------------|------------|----------|---------------|---|-------------|------------|----------|---------------|
| | Estimate | Std. Error | t value | Pr(> t) | | Estimate | Std. Error | t value | Pr(> t) |
| ar1 | -3.9323e-01 | 3.2251e-01 | -1.2193 | 0.224843 | ar1 | 6.6132e-01 | 7.8985e-02 | 8.3728 | 6.347e-14 *** |
| ar2 | 6.6514e-01 | 2.0018e-01 | 3.3227 | 0.001145 ** | ar2 | 1.7958e-01 | 9.6582e-02 | 1.8594 | 0.065152 . |
| ar3 | 3.5516e-01 | 3.2076e-01 | 1.1072 | 0.270143 | ar3 | 1.2861e-01 | 9.5522e-02 | 1.3463 | 0.180449 |
| ma1 | 1.0728e+00 | 3.3440e-01 | 3.2080 | 0.001667 ** | ar4 | -1.3909e-01 | 8.0026e-02 | -1.7380 | 0.084483 . |
| ma2 | 1.6339e-01 | 4.1781e-01 | 0.3911 | 0.696361 | sar1 | 2.2181e-01 | 8.3532e-02 | 2.6555 | 0.008873 ** |
| intercept | 5.6545e+02 | 1.0046e+01 | 56.2892 | < 2.2e-16 *** | sar2 | 7.3561e-02 | 9.0567e-02 | 0.8122 | 0.418088 |
| t | 1.0374e+00 | 5.0058e-01 | 2.0724 | 0.040114 * | intercept | 5.6594e+02 | 1.1601e+01 | 48.7854 | < 2.2e-16 *** |
| t2 | 1.6377e-02 | 6.4643e-03 | 2.5335 | 0.012428 * | t | 1.0028e+00 | 5.8510e-01 | 1.7139 | 0.088841 . |
| t3 | -8.1455e-05 | 3.3759e-05 | -2.4128 | 0.017162 * | t2 | 1.6994e-02 | 7.7039e-03 | 2.2059 | 0.029085 * |
| Jan | 2.4609e+01 | 2.0797e+00 | 11.8332 | < 2.2e-16 *** | t3 | -8.4331e-05 | 3.6801e-05 | -2.2916 | 0.023480 * |
| Feb | -1.3608e+01 | 2.4701e+00 | -5.5092 | 1.744e-07 *** | Jan | 2.4224e+01 | 2.7675e+00 | 8.7530 | 7.488e-15 *** |
| Mar | 7.7790e+01 | 2.6687e+00 | 29.1484 | < 2.2e-16 *** | Feb | -1.4266e+01 | 3.2570e+00 | -4.3802 | 2.360e-05 *** |
| Apr | 9.3189e+01 | 2.9588e+00 | 31.4955 | < 2.2e-16 *** | Mar | 7.7465e+01 | 3.5636e+00 | 21.7379 | < 2.2e-16 *** |
| May | 1.5432e+02 | 2.9326e+00 | 52.6222 | < 2.2e-16 *** | Apr | 9.2767e+01 | 3.9880e+00 | 23.2617 | < 2.2e-16 *** |
| Jun | 1.2622e+02 | 3.0799e+00 | 40.9825 | < 2.2e-16 *** | May | 1.5434e+02 | 4.1387e+00 | 37.2910 | < 2.2e-16 *** |
| Jul | 7.6093e+01 | 2.9189e+00 | 26.0694 | < 2.2e-16 *** | Jun | 1.2635e+02 | 4.1640e+00 | 30.3437 | < 2.2e-16 *** |
| Aug | 3.3646e+01 | 2.9359e+00 | 11.4602 | < 2.2e-16 *** | Jul | 7.6224e+01 | 4.1224e+00 | 18.4903 | < 2.2e-16 *** |
| Sep | -8.6685e+00 | 2.6291e+00 | -3.2971 | 0.001246 ** | Aug | 3.3910e+01 | 3.9546e+00 | 8.5748 | 2.045e-14 *** |
| Oct | -5.5399e+00 | 2.4174e+00 | -2.2917 | 0.023459 * | Sep | -8.5258e+00 | 3.5090e+00 | -2.4297 | 0.016426 * |
| Nov | -3.6216e+01 | 2.0161e+00 | -17.9632 | < 2.2e-16 *** | Oct | -5.3385e+00 | 3.1820e+00 | -1.6777 | 0.095710 . |
| | | | | | Nov | -3.6033e+01 | 2.6780e+00 | -13.4551 | < 2.2e-16 *** |



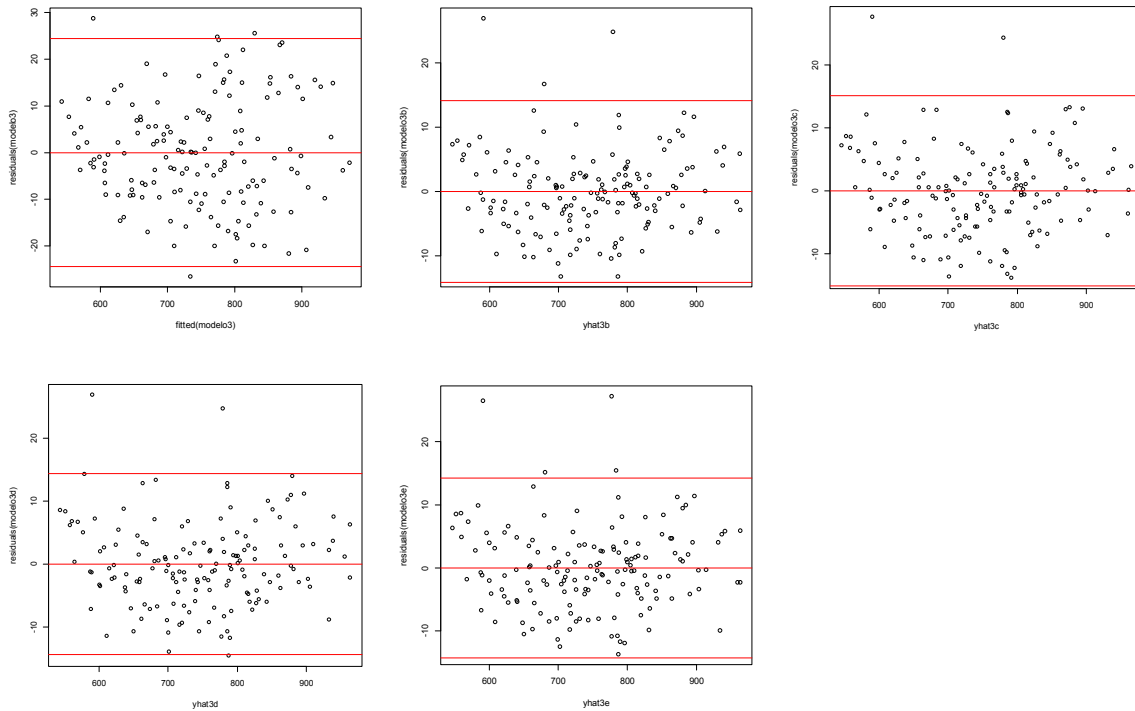
| | p | AIC | BIC |
|----------|----|-----------|-----------|
| modelo3 | 15 | 162.56684 | 217.96720 |
| modelo3b | 29 | 58.87909 | 103.79798 |
| modelo3c | 18 | 59.85563 | 85.10127 |
| modelo3d | 20 | 59.11369 | 87.39776 |
| modelo3e | 21 | 57.64875 | 86.91460 |

| Tabla de pronósticos todos los modelos | | | | | | | | | | |
|--|----------|----------|-----------|----------|--|----------|-----------|-----------|-----------|----------|
| Modelo 3 | | | | | Modelo 3b | | | | | |
| | fit | lwr | upr | | Point | Forecast | Lo 95 | Hi 95 | | |
| Jan 1975 | 845.9031 | 819.6430 | 872.1633 | | Jan 1975 | 840.2519 | 826.4113 | 854.0924 | | |
| Feb 1975 | 807.8973 | 781.5304 | 834.2642 | | Feb 1975 | 801.1289 | 784.5092 | 817.7486 | | |
| Mar 1975 | 899.5479 | 873.0657 | 926.0301 | | Mar 1975 | 900.2454 | 881.7382 | 918.7526 | | |
| Apr 1975 | 915.1627 | 888.5563 | 941.7691 | | Apr 1975 | 920.0779 | 899.6482 | 940.5076 | | |
| May 1975 | 976.4340 | 949.6941 | 1003.1740 | | May 1975 | 980.4213 | 959.1874 | 1001.6552 | | |
| Jun 1975 | 948.5157 | 921.6324 | 975.3990 | | Jun 1975 | 951.8616 | 930.0672 | 973.6561 | | |
| Jul 1975 | 898.4077 | 871.3708 | 925.4445 | | Jul 1975 | 903.8339 | 881.9157 | 925.7521 | | |
| Aug 1975 | 856.1100 | 828.9089 | 883.3111 | | Aug 1975 | 861.9577 | 839.7040 | 884.2115 | | |
| Sep 1975 | 813.6995 | 786.3231 | 841.0760 | | Sep 1975 | 815.7642 | 793.2964 | 838.2320 | | |
| Oct 1975 | 816.9456 | 789.3822 | 844.5090 | | Oct 1975 | 816.6504 | 793.9202 | 839.3805 | | |
| Nov 1975 | 786.0789 | 758.3166 | 813.8412 | | Nov 1975 | 783.5334 | 760.6144 | 806.4524 | | |
| Dec 1975 | 822.2533 | 794.2797 | 850.2269 | | Dec 1975 | 819.7694 | 796.8112 | 842.7277 | | |
| Modelo 3c | | | | | Modelo 3d | | | | | |
| | Point | Forecast | Lo 95 | Hi 95 | | Point | Forecast | Lo 95 | Hi 95 | |
| Jan 1975 | | 838.4882 | 823.6303 | 853.3461 | Jan 1975 | | 836.8151 | 822.6178 | 851.0125 | |
| Feb 1975 | | 800.9379 | 783.1803 | 818.6956 | Feb 1975 | | 800.0944 | 782.9294 | 817.2594 | |
| Mar 1975 | | 893.0502 | 873.1898 | 912.9107 | Mar 1975 | | 891.7382 | 872.8135 | 910.6630 | |
| Apr 1975 | | 909.1231 | 887.9020 | 930.3443 | Apr 1975 | | 907.6820 | 887.0070 | 928.3571 | |
| May 1975 | | 970.9045 | 948.7371 | 993.0718 | May 1975 | | 969.2043 | 947.8222 | 990.5864 | |
| Jun 1975 | | 943.1437 | 920.3124 | 965.9751 | Jun 1975 | | 941.2943 | 919.0249 | 963.5637 | |
| Jul 1975 | | 892.8225 | 869.5187 | 916.1262 | Jul 1975 | | 891.4170 | 868.7667 | 914.0674 | |
| Aug 1975 | | 850.3419 | 826.6997 | 873.9841 | Aug 1975 | | 848.9882 | 825.9049 | 872.0714 | |
| Sep 1975 | | 807.6441 | 783.7582 | 831.5300 | Sep 1975 | | 806.7375 | 783.4318 | 830.0431 | |
| Oct 1975 | | 810.9852 | 786.9231 | 835.0473 | Oct 1975 | | 809.7607 | 786.2410 | 833.2803 | |
| Nov 1975 | | 780.0797 | 755.8408 | 804.3185 | Nov 1975 | | 778.9732 | 755.3240 | 802.6225 | |
| Dec 1975 | | 816.2326 | 791.8745 | 840.5908 | Dec 1975 | | 814.9645 | 791.2062 | 838.7229 | |
| Modelo 3e | | | | | AmplitudI.P Cobertura RMSE MAE MAPE | | | | | |
| | Point | Forecast | Lo 95 | Hi 95 | modelo3 | 54.04210 | 1.0000000 | 12.68132 | 10.731773 | 1.265102 |
| Jan 1975 | | 835.6604 | 821.6782 | 849.6426 | modelo3b | 41.27881 | 0.8333333 | 13.22929 | 11.671769 | 1.356171 |
| Feb 1975 | | 793.4301 | 776.6669 | 810.1933 | modelo3c | 43.69779 | 0.8333333 | 12.56558 | 10.128478 | 1.212182 |
| Mar 1975 | | 891.4786 | 872.6262 | 910.3310 | modelo3d | 42.43001 | 0.8333333 | 12.91548 | 10.042572 | 1.206607 |
| Apr 1975 | | 906.0485 | 885.0871 | 927.0099 | modelo3e | 43.27902 | 0.8333333 | 13.07821 | 9.362639 | 1.126413 |
| May 1975 | | 968.6836 | 946.6302 | 990.7369 | | | | | | |
| Jun 1975 | | 941.7349 | 918.8796 | 964.5902 | | | | | | |
| Jul 1975 | | 894.4568 | 871.0223 | 917.8913 | | | | | | |
| Aug 1975 | | 852.1744 | 828.3824 | 875.9664 | | | | | | |
| Sep 1975 | | 806.6321 | 782.5889 | 830.6752 | | | | | | |
| Oct 1975 | | 808.4343 | 784.2215 | 832.6471 | | | | | | |
| Nov 1975 | | 775.5383 | 751.2144 | 799.8623 | | | | | | |
| Dec 1975 | | 812.4912 | 788.0913 | 836.8912 | | | | | | |
| | | | | |  | | | | | |

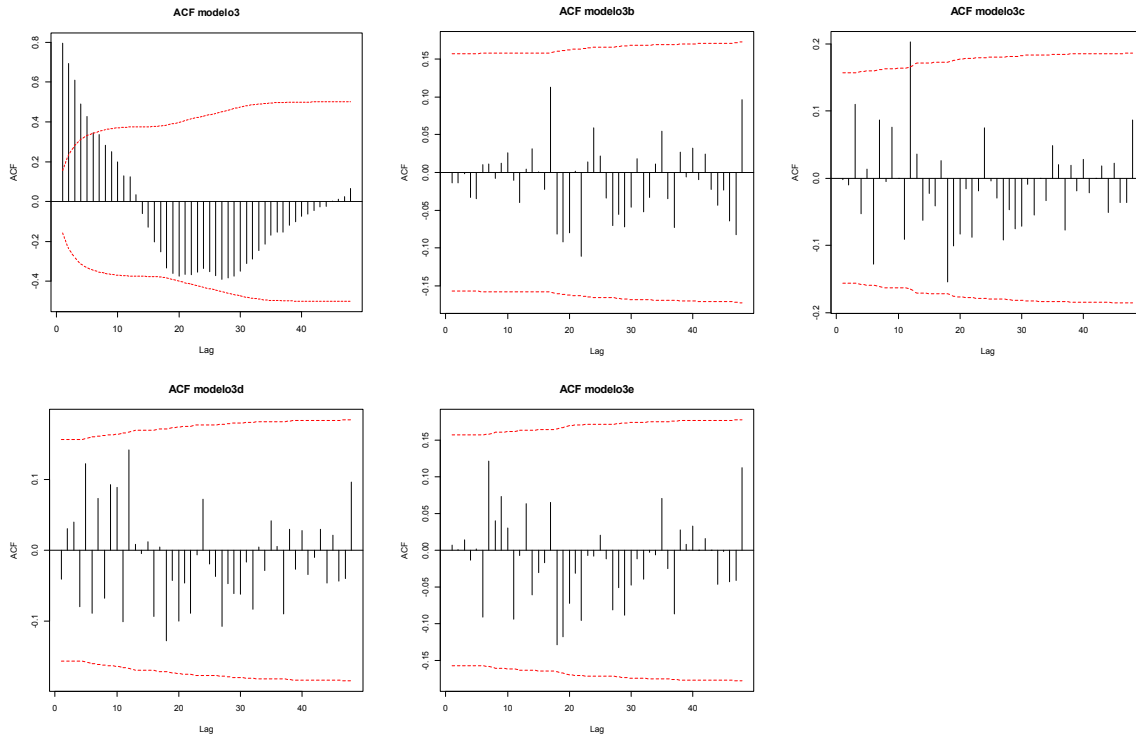
Gráficos de residuales vs. tiempo (todos los modelos)



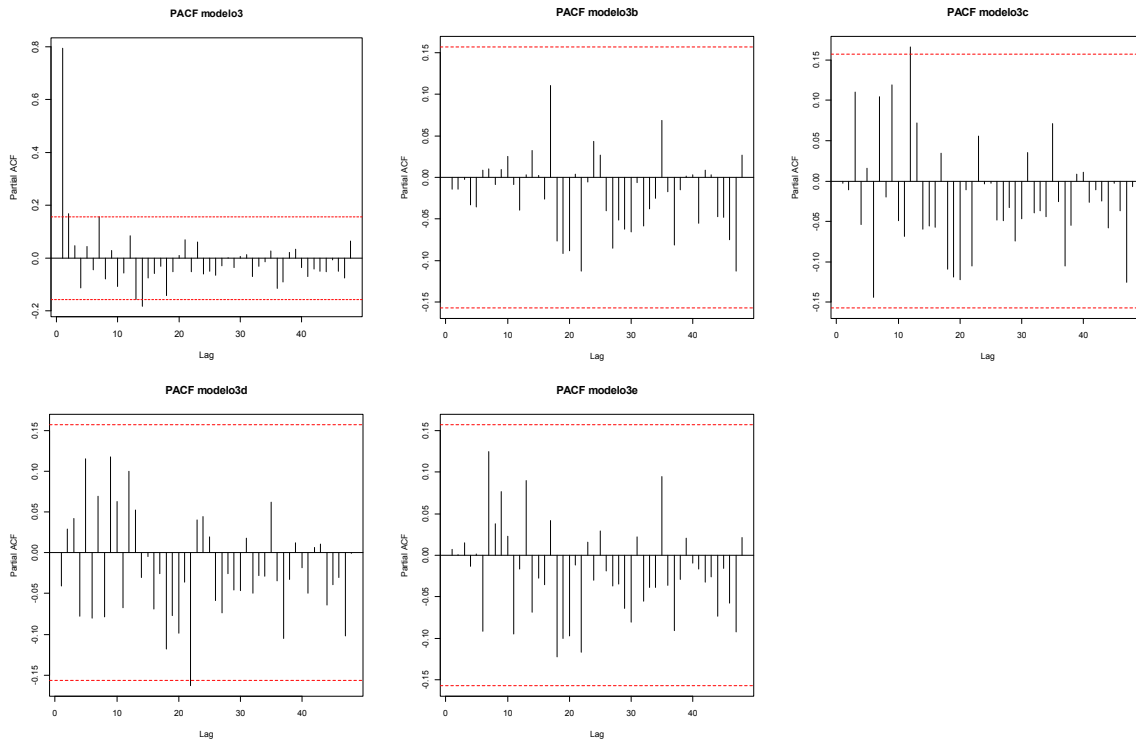
Gráficos de residuales vs. valores ajustados (todos los modelos)



ACF muestrales residuales de ajuste (todos los modelos)



PACF muestrales residuales de ajuste (todos los modelos)



Box-Pierce modelo3b

| m | X.squared | df | p.value |
|----|------------|----|-----------|
| 6 | 0.4317868 | 6 | 0.9985723 |
| 12 | 0.8528392 | 12 | 0.9999942 |
| 18 | 4.1034754 | 18 | 0.9997145 |
| 24 | 8.8816559 | 24 | 0.9978375 |
| 30 | 11.5091053 | 30 | 0.9990644 |
| 36 | 12.8279334 | 36 | 0.9998711 |
| 42 | 14.0396762 | 42 | 0.9999849 |
| 48 | 17.6275366 | 48 | 0.9999823 |

Ljung-box modelo3b

| m | X.squared | df | p.value |
|----|------------|----|-----------|
| 6 | 0.4492973 | 6 | 0.9984019 |
| 12 | 0.9081210 | 12 | 0.9999917 |
| 18 | 4.6073535 | 18 | 0.9993512 |
| 24 | 10.1999147 | 24 | 0.9936722 |
| 30 | 13.4444065 | 30 | 0.9959975 |
| 36 | 15.1493524 | 36 | 0.9991185 |
| 42 | 16.7711318 | 42 | 0.9998238 |
| 48 | 21.9652459 | 48 | 0.9995461 |

Box-Pierce modelo3c

| m | X.squared | df | p.value |
|----|-----------|----|-----------|
| 6 | 4.95372 | 6 | 0.5497625 |
| 12 | 14.69983 | 12 | 0.2582635 |
| 18 | 19.70462 | 18 | 0.3496925 |
| 24 | 24.56857 | 24 | 0.4294988 |
| 30 | 28.05633 | 30 | 0.5674517 |
| 36 | 29.16024 | 36 | 0.7833832 |
| 42 | 30.41068 | 42 | 0.9081487 |
| 48 | 32.53402 | 48 | 0.9572159 |

Ljung-box modelo3c

| m | X.squared | df | p.value |
|----|-----------|----|-----------|
| 6 | 5.171883 | 6 | 0.5219646 |
| 12 | 15.792777 | 12 | 0.2009127 |
| 18 | 21.487514 | 18 | 0.2555342 |
| 24 | 27.178825 | 24 | 0.2961868 |
| 30 | 31.493108 | 30 | 0.3914652 |
| 36 | 32.917738 | 36 | 0.6159739 |
| 42 | 34.587524 | 42 | 0.7846493 |
| 48 | 37.657945 | 48 | 0.8584011 |

Box-Pierce modelo3d

| m | X.squared | df | p.value |
|----|-----------|----|-----------|
| 6 | 1.964941 | 6 | 0.9228946 |
| 12 | 8.950700 | 12 | 0.7071355 |
| 18 | 14.473724 | 18 | 0.6977289 |
| 24 | 19.397214 | 24 | 0.7304875 |
| 30 | 22.865001 | 30 | 0.8207687 |
| 36 | 23.798026 | 36 | 0.9408229 |
| 42 | 25.162433 | 42 | 0.9815863 |
| 48 | 26.879234 | 48 | 0.9941275 |

Ljung-box modelo3d

| m | X.squared | df | p.value |
|----|-----------|----|-----------|
| 6 | 2.056298 | 6 | 0.9144487 |
| 12 | 9.693957 | 12 | 0.6427897 |
| 18 | 15.969469 | 18 | 0.5946782 |
| 24 | 21.724849 | 24 | 0.5956855 |
| 30 | 26.014502 | 30 | 0.6743911 |
| 36 | 27.218886 | 36 | 0.8538491 |
| 42 | 29.044471 | 42 | 0.9354344 |
| 48 | 31.540148 | 48 | 0.9680257 |

Box-Pierce modelo3e

| m | X.squared | df | p.value |
|----|-----------|----|-----------|
| 6 | 1.374005 | 6 | 0.9674206 |
| 12 | 6.297414 | 12 | 0.9003535 |
| 18 | 10.914389 | 18 | 0.8979542 |
| 24 | 15.518876 | 24 | 0.9046547 |
| 30 | 18.607506 | 30 | 0.9478472 |
| 36 | 19.760559 | 36 | 0.9871876 |
| 42 | 21.271070 | 42 | 0.9967794 |
| 48 | 24.128804 | 48 | 0.9984227 |

Ljung-box modelo3e

| m | X.squared | df | p.value |
|----|-----------|----|-----------|
| 6 | 1.445978 | 6 | 0.9630019 |
| 12 | 6.724044 | 12 | 0.8752982 |
| 18 | 11.956723 | 18 | 0.8494638 |
| 24 | 17.314087 | 24 | 0.8349606 |
| 30 | 21.133161 | 30 | 0.8836779 |
| 36 | 22.631171 | 36 | 0.9596557 |
| 42 | 24.646500 | 42 | 0.9849185 |
| 48 | 28.798654 | 48 | 0.9873214 |

Gráfico de normalidad residuos modelo3b

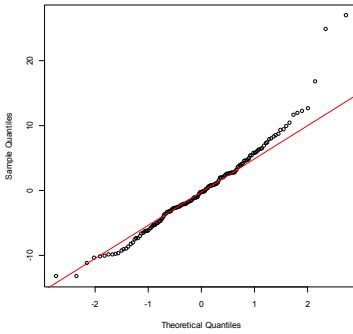


Gráfico de normalidad residuos modelo3c

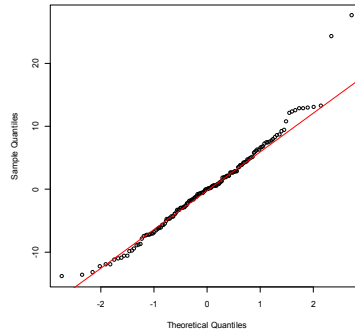


Gráfico de normalidad residuos modelo3d

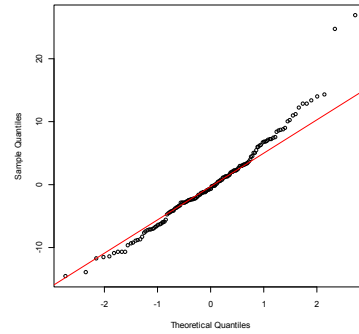
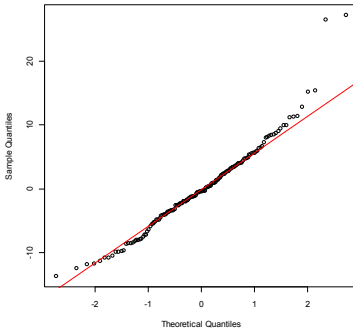


Gráfico de normalidad residuos modelo3e



Shapiro-Wilk normality test

data: residuals(modelo3b)

W = 0.95766, p-value = 0.0001082

Shapiro-Wilk normality test

data: residuals(modelo3c)

W = 0.97173, p-value = 0.002704

Shapiro-Wilk normality test

data: residuals(modelo3d)

W = 0.96768, p-value = 0.001017

Shapiro-Wilk normality test

data: residuals(modelo3e)

W = 0.95938, p-value = 0.0001562

Programación R usada

NOTA: La programación dada aquí para los modelos de regresión con errores arma sólo puede correrla con la versión R 3.3.3 lanzada el 06-marzo-2017 o con versiones posteriores. Con versiones anteriores no podrá obtener los valores ajustados sobre estos modelos usando la programación de este taller, por tanto deberá actualizar su paquete R.

1. Definiendo librerías a usar y funciones de usuario

```
library(forecast)
library(car)
library(TSA)
library(FitAR)
library(lmtest)

#Creando función usuario para obtener test Box-Pierce y Ljung-Box
BP.LB.test=function(serie,maxlag,type){
  aux=floor(maxlag/6)
  X.squared=c(rep(NA,aux))
  df=c(rep(NA,aux))
  p.value=c(rep(NA,aux))
  for(i in 1:aux){
    test=Box.test(serie,lag=(6*i),type=type)
    X.squared[i]=test[[1]]
    df[i]=test[[2]]
    p.value[i]=test[[3]]
  }
  lag=6*c(1:aux)
  teste=as.data.frame(cbind(X.squared,df,p.value))
  rownames(teste)=lag
  teste
}

#Creando una función usuario para obtener test Durbin-Watson
#para autocorrelación de orden 1, positiva y negativa
pruebaDW1=function(modelo){
  dwneg=durbinWatsonTest(modelo,max.lag=1,method="normal",alternative="negative")
  dwpos=durbinWatsonTest(modelo,max.lag=1,method="normal",alternative="positive")
  res=data.frame(1,dwneg$r,dwneg$dw,dwpos$p,dwneg$p)
  names(res)=c("lag","rho estimado","Estadístico D-W","VP rho>0","VP rho<0")
  res
}

#Creando función usuario crit.inf.resid() para calcular  $C_n^*(p)$ 
crit.inf.resid=function(residuales,n.par,AIC="TRUE"){
  if(AIC=="TRUE"){
    #Calcula AIC
    CI=log(mean(residuales^2))+2*n.par/length(residuales)
  }
  if(AIC=="FALSE"){
    #Calcula BIC
    CI=log(mean(residuales^2))+n.par*log(length(residuales)/length(residuales))
  }
  CI
}

#Función para calcular la amplitud de los I.P
amplitud=function(LIP,LSP){
  a=LSP-LIP
  am=mean(a)
  am
}

#Función para calcular la cobertura de los I.P
cobertura=function(real,LIP,LSP){
  I=ifelse(real>=LIP & real<=LSP,1,0)
  p=mean(I)
  p
}
```

2. Lectura datos y gráficos descriptivos de la serie

```
yt=scan()
589 561 640 656 727 697 640 599 568 577 553 582
600 566 653 673 742 716 660 617 583 587 565 598
628 618 688 705 770 736 678 639 604 611 594 634
658 622 709 722 782 756 702 653 615 621 602 635
677 635 736 755 811 798 735 697 661 667 645 688
713 667 762 784 837 817 767 722 681 687 660 698
717 696 775 796 858 826 783 740 701 706 677 711
734 690 785 805 871 845 801 764 725 723 690 734
750 707 807 824 886 859 819 783 740 747 711 751
804 756 860 878 942 913 869 834 790 800 763 800
826 799 890 900 961 935 894 855 809 810 766 805
821 773 883 898 957 924 881 837 784 791 760 802
828 778 889 902 969 947 908 867 815 812 773 813
834 782 892 903 966 937 896 858 817 827 797 843

yt=ts(yt,frequency=12, start=c(1962,1)) #serie con todos los datos

#Grafico la serie de tiempo
plot(yt,main="Produccion de leche mensual 1962.1-1975.12")

#Grafico de Descomposición en estacionalidad, tendencia y error
plot(decompose(yt))

#Boxplots según meses año calendario y periodograma
boxplot(yt~cycle(yt),names=month.abb)
periodogram(diff(yt),lwd=3;abline(v=c(1:6)/12,col=2,lty=2))

#ACF sobre la serie de producción de leche
win.graph(width=4.8,height=4.8,pointsize=8)
acf(as.numeric(yt),ci.type="ma",lag.max=48,main="ACF de la serie de producción mensual de leche")
```

3. Definiendo variables para el ajuste con validación cruzada

```
n=length(yt)-12
t=1:n; t2=t^2;t3=t^3
yt2=ts(yt[t],frequency=12, start=c(1962,1)) #serie recortada
mes=seasonaldummy(yt2)
I1=mes[,1]
I2=mes[,2]
I3=mes[,3]
I4=mes[,4]
I5=mes[,5]
I6=mes[,6]
I7=mes[,7]
I8=mes[,8]
I9=mes[,9]
I10=mes[,10]
I11=mes[,11]

X3=cbind(t,t2,t3,mes) #matriz de regresión que será usada en modelos de regresión
#con errores ARMA, en los periodos de ajuste
```

4. Definiendo variables para pronósticos ex – post

```
tnuevo=(n+1):length(yt)
t2nuevo=tnuevo^2
t3nuevo=tnuevo^3
mesnuevo=seasonaldummy(yt2,h=12)
#Separando una a una las 11 indicadoras para los tiempos de pronóstico
I1n=mesnuevo[,1]
I2n=mesnuevo[,2]
I3n=mesnuevo[,3]
I4n=mesnuevo[,4]
I5n=mesnuevo[,5]
I6n=mesnuevo[,6]
I7n=mesnuevo[,7]
I8n=mesnuevo[,8]
I9n=mesnuevo[,9]
I10n=mesnuevo[,10]
I11n=mesnuevo[,11]

X3nuevo=cbind(t=tnuevo,t2=t2nuevo,t3=t3nuevo,mes=mesnuevo) #matriz de regresión con valores de predictores
#para pronósticos en modelos con errores ARMA
ytnuevo=ts(yt[tnuevo],freq=12,start=c(1975,1)) #Serie de los últimos 12 valores observados
#para realizar la validación cruzada
```

5. Ajuste, análisis supuestos y pronósticos modelo 3 con tendencia cúbica y errores estructurales R.B.

```

modelo3=lm(yt2~t+t2+t3+I1+I2+I3+I4+I5+I6+I7+I8+I9+I10+I11); summary(modelo3)

yhat3=ts(fitted(modelo3),freq=12,start=c(1962,1))

plot(yt)
lines(yhat3,col=2)
legend("topleft",legend=c("Real","Ajuste modelo3"),lwd=2,col=c(1,2))

#Medidas de ajuste para modelo3 con exp(Cn*(p))
aic3=exp(crit.inf.resid(residuales=residuals(modelo3),n.par=15));aic3
bic3=exp(crit.inf.resid(residuales=residuals(modelo3),n.par=15,AIC="FALSE"));bic3

#Pronósticos modelo 3
pred3=predict(modelo3,newdata=data.frame(t=tNuevo,t2=t2Nuevo,t3=t3Nuevo,I1=I1n,I2=I2n,I3=I3n,I4=I4n,I5=I5n,I6=I6n,I7=I7n,I8=I8n,I9=I9n,I10=I10n,I11=I11n),interval="prediction")
pred3=ts(pred3,freq=12,start=c(1975,1))
pred3

#Amplitud media de los I.P
ampmod3=amplitud(LIP=pred3[,2],LSP=pred3[,3])
ampmod3

#cobertura de los I.P
cobmod3=cobertura(real=ytnuevo,LIP=pred3[,2],LSP=pred3[,3])
cobmod3

#precisión pronósticos puntuales
accuracy(pred3[,1],ytnuevo)

#Gráfico de residuales vs tiempo modelo3
win.graph(width=4.8,height=4.8,pointsize=8)
plot(t,residuals(modelo3),type="o"); abline(h=0,col=2)
abline(h=c(-2*summary(modelo3)$sigma,2*summary(modelo3)$sigma),col=2)

#Gráfico de residuales vs predichos modelo3
win.graph(width=4.8,height=4.8,pointsize=8)
plot(fitted(modelo3), residuals(modelo3), type="p"); abline(h=0,col=2)
abline(h=c(-2*summary(modelo3)$sigma,2*summary(modelo3)$sigma),col=2)

#ACF o FAC residuales modelo3
win.graph(width=4.8,height=4.8,pointsize=8)
acf(residuals(modelo3),ci.type="ma",lag.max=48,main="ACF modelo3",ci.col=2)

#PACF sobre residuales modelo3
win.graph(width=4.8,height=4.8,pointsize=8)
pacf(residuals(modelo3),lag.max=48,main="PACF modelo3",ci.col=2)

#Usando la función usuario BP.LB.test()para obtener test B-P y L-B de ruido blanco
#Para los residuales modelo 3
BP.LB.test(residuals(modelo3),maxlag=48,type="Box")
BP.LB.test(residuals(modelo3),maxlag=48,type="Ljung")

#Usando la función usuario pruebaDW1 sobre residuales modelo 3
pruebaDW1(modelo3)

```

6. Identificación modelos ARMA para errores estructurales con métodos automáticos

```

#Identificación con criterio AIC, auto.arima()
auto.arima(residuals(modelo3),ic="aic")

#Lo siguiente puede arrojar un modelo SARMA
et3=ts(residuals(modelo3),frequency=12, start=c(1962,1)) #serie de tiempo residuos modelo 3
auto.arima(et3,ic="aic") #Identificación com critério AIC, auto.arima()

#Identificación com AIC y BIC, SelectModel()
SelectModel(et3, lag.max = 36, Criterion="AIC",ARModel = "AR")
SelectModel(et3, lag.max = 36, Criterion="BIC",ARModel = "AR")

#Identificación con criterio BIC, armasubsets()
plot(armasubsets(et3,nar=12,nma=12,y.name='AR',ar.method='ml'))

#Identificación con EACF
eacf(residuals(modelo3),ar.max=24,ma.max=24)

```

7. Modelos de regresión con errores ARMA

7.1 Modelo con errores AR(14): $p=29$ parámetros

```

modelo3b=Arima(yt2,order=c(14,0,0),xreg=X3,method="ML")
dfb=n-29 #n-Total de parámetros
coeftest(modelo3b,df=dfb)
yhat3b=modelo3b$fitted #para obtener valores ajustados
plot(yt)
lines(yhat3b,col=2)
legend("topleft",legend=c("Real","ajustada modelo3b"),col=c(1,2),lty=1)
#Medidas de ajuste para modelo3b con  $\exp(C_n^*(p))$ 
aic3b=exp(crit.inf.resid(residuales=residuals(modelo3b),n.par=29));aic3b
bic3b=exp(crit.inf.resid(residuales=residuals(modelo3b),n.par=29,AIC="FALSE"));bic3b

#Gráfico de residuales de ajuste vs tiempo
win.graph(width=4.8,height=4.8,points=8)
plot(t,residuals(modelo3b),type="o"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3b$sigma2),2*sqrt(modelo3b$sigma2)),col=2)

#Gráfico de residuales de ajuste vs ajustados
win.graph(width=4.8,height=4.8,points=8)
plot(yhat3b, residuals(modelo3b), type="p"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3b$sigma2),2*sqrt(modelo3b$sigma2)),col=2)

#ACF sobre residuales de ajuste modelo3b
win.graph(width=4.8,height=4.8,points=8)
acf(as.numeric(residuals(modelo3b)),ci.type="ma",lag.max=48,main="ACF modelo3b",ci.col=2)

#PACF sobre residuales de ajuste modelo3b
win.graph(width=4.8,height=4.8,points=8)
pacf(as.numeric(residuals(modelo3b)),lag.max=48,main="PACF modelo3b",ci.col=2)

#Box-pierce & Ljung-Box
BP.LB.test(residuals(modelo3b),maxlag=48,type="Box")
BP.LB.test(residuals(modelo3b),maxlag=48,type="Ljung")

shapiro.test(residuals(modelo3b))
win.graph(width=4.8,height=4.8,points=8)
qqnorm(residuals(modelo3b),main="Gráfico de normalidad residuos modelo3b")
qqline(residuals(modelo3b),col=2)

#Pronósticos e I.P del 95%
pred3b=ts(as.data.frame(forecast(modelo3b,xreg=X3nuevo,level=95)),freq=12,start=c(1975,1)) #Matriz de regresión
#para pronósticos es
#X3nuevo

pred3b

#Precisión pronósticos puntuales
accuracy(pred3b[,1],ytnuevo)

#Precisión pronósticos por I.P
ampmod3b=amplitud(LIP=pred3b[,2],LSP=pred3b[,3])
ampmod3b

cobmod3b=cobertura(real=ytnuevo,LIP=pred3b[,2],LSP=pred3b[,3])
cobmod3b

```

7.2 Modelo con errores ARMA(2,10) pero sólo con $\phi_1, \phi_2, \theta_{10}$: $p=18$ parámetros

```

modelo3c=Arima(yt2,order=c(2,0,10),fixed=c(NA,NA,rep(0,9),NA,rep(NA,15)),xreg=X3,method="ML")
dfc=n-18 #n-Total parámetros
coeftest(modelo3c,df=dfc)

#Calculando valores ajustados
yhat3c=modelo3c$fitted

plot(yt)
lines(yhat3c,col=2)
legend("topleft",legend=c("Real","ajustada modelo3c"),col=c(1,2),lty=1)
#Medidas de ajuste para modelo3c con  $\exp(C_n^*(p))$ 
aic3c=exp(crit.inf.resid(residuales=residuals(modelo3c),n.par=18));aic3c
bic3c=exp(crit.inf.resid(residuales=residuals(modelo3c),n.par=18,AIC="FALSE"));bic3c

#Gráfico de residuales de ajuste vs tiempo
win.graph(width=4.8,height=4.8,points=8)
plot(t,residuals(modelo3c),type="o"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3c$sigma2),2*sqrt(modelo3c$sigma2)),col=2)

```

```

#Gráfico de residuales de ajuste vs ajustados
win.graph(width=4.8,height=4.8,points=8)
plot(yhat3c, residuals(modelo3c), type="p"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3c$sigma2),2*sqrt(modelo3c$sigma2)),col=2)

#ACF sobre residuales de ajuste modelo3c
win.graph(width=4.8,height=4.8,points=8)
acf(as.numeric(residuals(modelo3c)),ci.type="ma",lag.max=48,main="ACF modelo3c",ci.col=2)

#PACF sobre residuales de ajuste modelo3c
win.graph(width=4.8,height=4.8,points=8)
pacf(as.numeric(residuals(modelo3c)),lag.max=48,main="PACF modelo3c",ci.col=2)

#Box-pierce & Ljung-Box
BP.LB.test(residuals(modelo3c),maxlag=48,type="Box")
BP.LB.test(residuals(modelo3c),maxlag=48,type="Ljung")

shapiro.test(residuals(modelo3c))
win.graph(width=4.8,height=4.8,points=8)
qqnorm(residuals(modelo3c),main="Gráfico de normalidad residuos modelo3c")
qqline(residuals(modelo3c),col=2)

#Pronósticos e I.P del 95%
pred3c=ts(as.data.frame(forecast(modelo3c,xreg=X3nuevo,level=95)),freq=12,start=c(1975,1)) #matriz de regresión
#en pronósticos es
#X3nuevo

pred3c

#Precisión pronósticos puntuales
accuracy(pred3c[,1],ytnuevo)

#Precisión pronósticos por I.P
ampmod3c=amplitud(LIP=pred3c[,2],LSP=pred3c[,3])
ampmod3c
cobmod3c=cobertura(real=ytnuevo,LIP=pred3c[,2],LSP=pred3c[,3])
cobmod3c

```

7.3 Modelo con errores ARMA(3,2): p=20 parámetros

```

modelo3d=Arima(yt2,order=c(3,0,2),xreg=X3,method="ML")
dfd=n-20 #n-Total parámetros
coeftest(modelo3d,df=dfd)

#Cálculo valores ajustados
yhat3d=modelo3d$fitted
plot(yt)
lines(yhat3d,col=2)
legend("topleft",legend=c("Real","ajustada modelo3d"),col=c(1,2),lty=1)

#Medidas de ajuste para modelo3d con exp(Cn(p))
aic3d=exp(crit.inf.resid(residuales=residuals(modelo3d),n.par=20));aic3d
bic3d=exp(crit.inf.resid(residuales=residuals(modelo3d),n.par=20,AIC="FALSE"));bic3d

#Gráfico de residuales de ajuste vs tiempo
win.graph(width=4.8,height=4.8,points=8)
plot(t,residuals(modelo3d),type="o"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3d$sigma2),2*sqrt(modelo3d$sigma2)),col=2)

#Gráfico de residuales de ajuste vs ajustados
win.graph(width=4.8,height=4.8,points=8)
plot(yhat3d, residuals(modelo3d), type="p"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3d$sigma2),2*sqrt(modelo3d$sigma2)),col=2)

#ACF sobre residuales de ajuste modelo3d
win.graph(width=4.8,height=4.8,points=8)
acf(as.numeric(residuals(modelo3d)),ci.type="ma",lag.max=48,main="ACF modelo3d",ci.col=2)

#PACF sobre residuales de ajuste modelo3d
win.graph(width=4.8,height=4.8,points=8)
pacf(as.numeric(residuals(modelo3d)),lag.max=48,main="PACF modelo3d",ci.col=2)

#Box-pierce & Ljung-Box
BP.LB.test(residuals(modelo3d),maxlag=48,type="Box")
BP.LB.test(residuals(modelo3d),maxlag=48,type="Ljung")

shapiro.test(residuals(modelo3d))
win.graph(width=4.8,height=4.8,points=8)
qqnorm(residuals(modelo3d),main="Gráfico de normalidad residuos modelo3d")

```

```

qqline(residuals(modelo3d),col=2)
#Pronósticos e I.P del 95%
pred3d=ts(as.data.frame(forecast(modelo3d,xreg=X3nuevo,level=95)),freq=12,start=c(1975,1)) #matriz de regresión
#en pronósticos es
#X3nuevo

pred3d

#Precisión pronósticos puntuales
accuracy(pred3d[,1],ytnuevo)

#Precisión pronósticos por I.P
ampmod3d=amplitud(LIP=pred3d[,2],LSP=pred3d[,3])
ampmod3d

cobmod3d=cobertura(real=ytnuevo,LIP=pred3d[,2],LSP=pred3d[,3])
cobmod3d

```

7.4 Modelo con errores ARMA(4,0)xARMA(2,0)[12]: p=21 parámetros

```

modelo3e=Arima(yt2,order=c(4,0,0),seasonal=list(order=c(2,0,0)),xreg=X3,method="ML")
dfe=n-21 #n-Total parámetros
coeftest(modelo3e, df=dfe)

#Cálculo valores ajustados
yhat3e=modelo3e$fitted
plot(yt)
lines(yhat3e,col=2)
legend("topleft",legend=c("Real","ajustada modelo3e"),col=c(1,2),lty=1)

#Medidas de ajuste para modelo3e con exp(Cn*(p))
aic3e=exp(crit.inf.resid(residuales=residuals(modelo3e),n.par=21));aic3e
bic3e=exp(crit.inf.resid(residuales=residuals(modelo3e),n.par=21,AIC="FALSE"));bic3e

#Gráfico de residuales de ajuste vs tiempo
win.graph(width=4.8,height=4.8,pointsize=8)
plot(t,residuals(modelo3e),type="o"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3e$sigma2),2*sqrt(modelo3e$sigma2)),col=2)

#Gráfico de residuales de ajuste vs ajustados
win.graph(width=4.8,height=4.8,pointsize=8)
plot(yhat3e, residuals(modelo3e), type="p"); abline(h=0,col=2)
abline(h=c(-2*sqrt(modelo3e$sigma2),2*sqrt(modelo3e$sigma2)),col=2)

#ACF sobre residuales de ajuste modelo3e
win.graph(width=4.8,height=4.8,pointsize=8)
acf(as.numeric(residuals(modelo3e)),ci.type="ma",lag.max=48,main="ACF modelo3e",ci.col=2)

#PACF sobre residuales de ajuste modelo3e
win.graph(width=4.8,height=4.8,pointsize=8)
pacf(as.numeric(residuals(modelo3e)),lag.max=48,main="PACF modelo3e",ci.col=2)

#Box-pierce & Ljung-Box
BP.LB.test(residuals(modelo3e),maxlag=48,type="Box")
BP.LB.test(residuals(modelo3e),maxlag=48,type="Ljung")

shapiro.test(residuals(modelo3e))
win.graph(width=4.8,height=4.8,pointsize=8)
qqnorm(residuals(modelo3e),main="Gráfico de normalidad residuos modelo3e")
qqline(residuals(modelo3e),col=2)

#Pronósticos e I.P del 95%
pred3e=ts(as.data.frame(forecast(modelo3e,xreg=X3nuevo,level=95)),freq=12,start=c(1975,1)) #matriz de regresión
#en pronósticos es
#X3nuevo

pred3e

#Precisión pronósticos puntuales
accuracy(pred3e[,1],ytnuevo)

#Precisión pronósticos por I.P
ampmod3e=amplitud(LIP=pred3e[,2],LSP=pred3e[,3])
ampmod3e

cobmod3e=cobertura(real=ytnuevo,LIP=pred3e[,2],LSP=pred3e[,3])
cobmod3e

```

8. Comparaciones pronósticos y de las medidas de ajuste y de precisión de pronósticos

```
#Gráfico comparativo de pronósticos ex-post
plot(ytnuevo,type="b",pch=19,col=1,lwd=2,xaxt="n",ylim=c(750,1000))
axis(1,at=time(ytnuevo),labels=c("I-75","II-75","III-75","IV-75","V-75","VI-75","VII-75","VIII-75","IX-75","X-75","XI-75","XII-75"),cex.axis=0.7)
lines(pred3[,1],col=2,type="b",pch=2,lwd=2)
lines(pred3b[,1],col=4,type="b",pch=4,lwd=2)
lines(pred3c[,1],col=5,type="b",pch=5,lwd=2)
lines(pred3d[,1],col=6,type="b",pch=6,lwd=2)
lines(pred3e[,1],col=7,type="b",pch=7,lwd=2)
legend("topright",legend=c("Real","pron. modelo3","pron. modelo3b","pron. modelo3c","pron. modelo3d","pron. modelo3e"),col=c(1,2,4,5,6,7),pch=c(19,2,4,5,6,7),lwd=2)

#Tabla con todas las medidas de ajuste
tablamedidas=data.frame(p=c(15,29,18,20,21),AIC=c(aic3,aic3b,aic3c,aic3d,aic3e),
BIC=c(bic3,bic3b,bic3c,bic3d,bic3e),row.names=c("modelo3","modelo3b","modelo3c","modelo3d","modelo3e"))

tablamedidas

#Tabla con todas las medidas de precisión de pronósticos
tablaprec=data.frame(AmplitudI.P=c(ampmod3,ampmod3b,ampmod3c,ampmod3d,ampmod3e),
Cobertura=c(cobmod3,cobmod3b,cobmod3c,cobmod3d,cobmod3d),
RMSE=c(accuracy(pred3[,1],ytnuevo)[2],accuracy(pred3b[,1],ytnuevo)[2],accuracy(pred3c[,1],ytnuevo)[2],
accuracy(pred3d[,1],ytnuevo)[2],accuracy(pred3e[,1],ytnuevo)[2]),
MAE=c(accuracy(pred3[,1],ytnuevo)[3],accuracy(pred3b[,1],ytnuevo)[3],accuracy(pred3c[,1],ytnuevo)[3],
accuracy(pred3d[,1],ytnuevo)[3],accuracy(pred3e[,1],ytnuevo)[3]),
MAPE=c(accuracy(pred3[,1],ytnuevo)[5],accuracy(pred3b[,1],ytnuevo)[5],accuracy(pred3c[,1],ytnuevo)[5],
accuracy(pred3d[,1],ytnuevo)[5],accuracy(pred3e[,1],ytnuevo)[5]),
row.names=c("modelo3","modelo3b","modelo3c","modelo3d","modelo3e"))

tablaprec
```


Algunos resultados Taller 9

1. Análisis de componentes

- La tendencia es no lineal y debe ser modelada con función creciente dado lo que representan los datos. Se propone un polinomio de grado 3 en la modelación global.
- Hay patrón que se repite anualmente en la serie; boxplots indican que $E[Y_t | t \in \text{mes } i]$ no es igual para todo i (el mes del año); periodograma muestra que la serie tiene asociación con las ondas sinusoidales armónicas que tiene frecuencias $F_j = j/12$, $j = 1, 2, \dots, 6$. Por todo lo anterior, sí hay patrón estacional.

2. Análisis ACF de Y_t :

- La estimación de $\text{corr}(Y_t, Y_{t+k})$ indicia que la serie no proviene de un proceso ergódico, es decir, no se cumple que $\lim_{k \rightarrow 0} \text{corr}(Y_t, Y_{t+k}) = 0$ rápidamente.
- en $\widehat{\text{corr}}(Y_t, Y_{t+k})$ se observa patrón que se repite cada k múltiplo de 12 y decae lentamente y esto indica presencia de patrón periódico exacto en la serie. En los primeros $k = 1, 2, \dots, 6$ pareciera decaer rápido pero si observa la escala vertical del gráfico, el cambio en la estimación realmente no decae muy rápido, además en $k = 1$ inicia con un valor cercano a 1, esto puede indicar la presencia de tendencia en la serie.
- Sí, pues los patrones descritos en ACF suelen presentarse cuando hay tendencia y componente estacional periódica exacta.

- Modelo 3 es $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \dots + \delta_{11} I_{11,t} + E_t$, con $\{E_t\}_{t \in \mathbb{Z}^+}$ un $RB \sim N(0, \sigma^2)$. La ecuación ajustada es,

$$\hat{Y}_t = 564.07 + 1.149t + 0.01425t^2 - 7.07 \times 10^{-5}t^3 + 25.27I_{1,t} - 13.08I_{2,t} + 78.26I_{3,t} + 93.60I_{4,t} + 154.6I_{5,t} + 126.53I_{6,t} + 76.28I_{7,t} + 33.87I_{8,t} - 8.61I_{9,t} - 5.38I_{10,t} - 36.23I_{11,t}.$$

- Desde que el modelo es de componentes aditivas, entonces para los δ_i tales que $H_0 : \delta_i = 0$ se rechaza en favor de $H_1 : \delta_i \neq 0$, el valor estimado se interpreta como la siguiente diferencia de medias estimadas: $\hat{\delta}_i = \hat{E}[Y_t | t \in \text{mes } i] - \hat{E}[Y_t | t \in \text{mes } 12]$ correspondiendo a la estimación de la diferencia que en promedio en cada año se da entre la producción de leche (Lb/vaca) del mes i con respecto a diciembre. Para los δ_i tales que $H_0 : \delta_i = 0$ no se rechaza, se concluye que en promedio, en cada año la producción de ese mes es igual a la producción de diciembre.

- En el modelo 3: $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \dots + \delta_{11} I_{11,t} + E_t$, con $\{E_t\}_{t \in \mathbb{Z}^+}$ un $RB \sim N(0, \sigma^2)$ no se cumple que el error estructural E_t sea un RB pues:

- Aunque según gráficos de residuales no hay evidencia fuerte contra supuestos $E[E_t] = 0$, y $\text{Var}[E_t] = \sigma^2 \forall t > 0$, Hay ciclos en la serie de residuos $\hat{E}_t = Y_t - \hat{Y}_t$, lo cual implica que $\rho(k) = \text{corr}(E_t, E_{t+k}) \neq 0$ para al menos un $k \neq 0$, es más, dado que los ciclos son evidentes a primera vista esto implica que $\rho(1) = \text{corr}(E_t, E_{t+1}) > 0$.

- La conclusión anterior se verifica en el test de la ACF para $k = 1, 2, \dots, 48$:

$$H_0 : \rho(k) = \text{corr}(E_t, E_{t+k}) = 0 \text{ vs. } H_1 : \rho(k) \neq 0,$$

$$\text{estadístico de prueba } \hat{\rho}(k) = \widehat{\text{corr}}(E_t, E_{t+k}) = \frac{\sum_{t=1}^{156-k} \hat{E}_t \hat{E}_{t+k}}{\sum_{t=1}^{156} \hat{E}_t^2} \overset{\text{aprox}}{\sim} N(0, \frac{1}{156}) \text{ y criterio de}$$

$$\text{rechazo } |\hat{\rho}(k)| \geq 2/\sqrt{156},$$

detecta que $\rho(k) \neq 0$ para $k = 1, 2, 3, 4, 5, 6$, y como $\hat{\rho}(1) > 0$, entonces se infiere que

$$\rho(1) > 0.$$

Por otra parte el test de la PACF para $k = 1, 2, \dots, 48$:

$$H_0 : \phi_{kk} = \text{corr}(E_t, E_{t+k} | E_{t+1}, \dots, E_{t+k-1}) = 0 \text{ vs. } H_1 : \phi_{kk} \neq 0,$$

estadístico de prueba $\hat{\phi}_{kk} = \widehat{\text{corr}}(E_t, E_{t+k} | E_{t+1}, \dots, E_{t+k-1}) \stackrel{\text{aprox}}{\sim} N(0, \frac{1}{156})$ y criterio de rechazo $|\hat{\phi}_{kk}| \geq 2/\sqrt{156}$,

muestra que $\phi_{kk} \neq 0$ para $k = 1, 2, 7, 13, 14$ corroborando una vez más que E_t no es un ruido blanco.

También se obtiene evidencia contra supuesto de ruido blanco con los tests Ljung-Box & Box-Pierce, cada uno evaluado con $m = 6, 12, 18, 24, 30, 36, 42$:

$H_0 : \rho(1) = \rho(2) = \dots = \rho(m) = 0$ vs. $H_1 : \rho(k) \neq 0$ para al menos un $k, k = 1, 2, \dots, m$, estadístico de prueba Box-Pierce, $Q_{BP} = 156 \sum_{k=1}^m [\hat{\rho}(k)]^2 \stackrel{\text{aprox}}{\sim} \chi_m^2$ y estadístico de prueba

Ljung-Box, $Q_{LB} = 156 \times 158 \sum_{k=1}^m \frac{[\hat{\rho}(k)]^2}{156-k} \stackrel{\text{aprox}}{\sim} \chi_m^2$ y criterios de rechazo VP = $P(\chi_m^2 \geq$

$Q_{BP})$, VP = $P(\chi_m^2 \geq Q_{LB})$ pequeños,

y se rechaza H_0 en las siete pruebas presentadas, pues en todas ellas se observan valores P pequeños, lo cual conduce nuevamente a concluir que no es cierto que E_t en el modelo 3 sea un ruido blanco.

Con el test Durbin Watson también se obtiene evidencia en contra de E_t un RB: Bajo este test el modelo de regresión es (misma estructura de regresión propuesta pero el error estructural lo asume un AR(1) estacionario con media cero)

$$Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \dots + \delta_{11} I_{11,t} + E_t,$$

donde

$$E_t = \phi_1 E_{t-1} + a_t, \{a_t\}_{t \in \mathbb{Z}^+} \text{ un RB } \sim N(0, \sigma_a^2) \text{ y } |\phi_1| < 1$$

entonces $\rho(1) = \phi_1 = \text{corr}(E_t, E_{t+1})$ y se prueba que $H_0 : \rho(1) = 0$ vs. $H_1 : \rho(1) > 0$. Se elige contrastar la hipótesis nula contra la hipótesis $H_1 : \rho(1) > 0$, ya que el estadístico de la prueba $d_1 = 0.4068884 < 2$ lo que sólo es posible si $\rho(1) > 0$. El valor P correspondiente, $P(DW_1 < 0.4068884) \approx 0.00$ implica rechazo de H_0 y por tanto evidencia a favor de H_1 , es decir, que se detecta que en los errores estructurales E_t hay autocorrelación positiva de orden 1, lo cual invalida supuesto de ruido blanco.

- c) Aunque no se cumple el supuesto de ruido blanco, sin embargo, para E_t se cumple que su media es constante en cero y se puede asumir con varianza aprox. constante (ver gráfico de residuales del modelo 3). Como además los ciclos observados son ergódicos según ACF, es decir, podemos concluir con ACF muestral que se cumple $\lim_{k \rightarrow \infty} \rho(k) = \text{corr}(E_t, E_{t+k}) = 0$ rápidamente, entonces podemos decir que el error estructural E_t del modelo global es estacionario en covarianza y por tanto se procede a identificar qué tipo de ARMA estacionario.

6) Identificación:

- Con ACF-PAF: Se identifica patrón cola exponencial sinusoidal amortiguado en ACF y posible patrón corte en PACF, luego E_t es un AR(p) estacionario con $p \geq 2$, ya que el tipo de ACF observado sólo ocurre cuando el polinomio AR tiene raíces complejas y por tanto como mínimo p es igual a 2. Evaluando PACF el último corte ocurre en $k = 14$ entonces tomamos $p = 14$, es decir, un AR(14).
- Con auto.arima criterio AIC se identifican según salidas R presentadas que E_t puede ser un ARMA(3,2) estacionario de media cero, o un ARMA(4,0)(2,0)[12] también estacionario con media cero.
- Con armasubsets tablero presentado, ignoramos el modelo del primer renglón porque en éste la parte AR no cumple $p \geq 2$, así que buscando el modelo de menor BIC satisfaciendo esta condición, encontramos ese modelo en segundo renglón: E_t una ARMA(2,10) pero sólo con parámetros ϕ_1, ϕ_2 y θ_{10} .

- Con SelectModel Criterio AIC el mejor modelo es un AR(2) y el segundo mejor un AR(14); con BIC ignoramos el primer modelo pues es un AR(1) pues necesitamos $p \geq 2$, tomamos entonces el segundo modelo un AR(2) o el tercero que es un AR(3).
- Con EACF, de acuerdo a la regla de identificación con esta gráfica, el triángulo de ceros de mayor área y satisfaciendo parte AR de orden $p \geq 2$, se identifica un ARMA(6,5).

En resumen los modelos identificados para el error estructural son:

Tabla 1: Modelos ARMA(p,q) de media cero

| Nombre | Ecuación | Polinomio AR: $\phi_p(B)$ | Polinomio MA: $\theta_q(B)$ |
|---|---|---|---|
| AR(14) | $E_t = \sum_{j=1}^{14} \phi_j E_{t-j} + a_t, a_t \text{ RB} \sim N(0, \sigma_a^2)$ | $\phi_{14}(B) = 1 - \sum_{j=1}^{14} \phi_j B^j$ | NA |
| ARMA(3,2) | $E_t = \sum_{j=1}^3 \phi_j E_{t-j} + a_t + \sum_{i=1}^2 \theta_i a_{t-i}, a_t \text{ RB} \sim N(0, \sigma_a^2)$ | $\phi_3(B) = 1 - \sum_{j=1}^3 \phi_j B^j$ | $\theta_2(B) = 1 + \sum_{i=1}^2 \theta_i B^i$ |
| ARMA(6,5) | $E_t = \sum_{j=1}^6 \phi_j E_{t-j} + a_t + \sum_{i=1}^5 \theta_i a_{t-i}, a_t \text{ RB} \sim N(0, \sigma_a^2)$ | $\phi_6(B) = 1 - \sum_{j=1}^6 \phi_j B^j$ | $\theta_5(B) = 1 + \sum_{i=1}^5 \theta_i B^i$ |
| ARMA(2,10) sólo con ϕ_1, ϕ_2 y θ_{10} | $E_t = \phi_1 E_{t-1} + \phi_2 E_{t-2} + a_t + \theta_{10} a_{t-10}, a_t \text{ RB} \sim N(0, \sigma_a^2)$ | $\phi_2(B) = 1 - \sum_{j=1}^2 \phi_j B^j$ | $\theta_{10}(B) = 1 + \theta_{10} B^{10}$ |
| AR(2) | $E_t = \sum_{j=1}^2 \phi_j E_{t-j} + a_t, a_t \text{ RB} \sim N(0, \sigma_a^2)$ | $\phi_2(B) = 1 - \sum_{j=1}^2 \phi_j B^j$ | NA |
| AR(3) | $E_t = \sum_{j=1}^3 \phi_j E_{t-j} + a_t, a_t \text{ RB} \sim N(0, \sigma_a^2)$ | $\phi_3(B) = 1 - \sum_{j=1}^3 \phi_j B^j$ | NA |

Los modelos AR(p) de la Tabla anterior son estacionarios sólo si sus respectivos polinomios $\phi_p(B)$ tienen sus p raíces de módulo mayor a 1 y son siempre invertibles; los demás modelos ARMA(p,q) en Tabla 1 son estacionarios sólo si sus respectivos polinomios $\phi_p(B)$ tienen sus p raíces de módulo mayor a 1 y serán invertibles si sus respectivos polinomios $\theta_q(B)$ tienen raíces de módulo mayor a 1.

ARMA(4,0)(2,0)[12]: $\phi_4(B)\Phi_2(B^{12})E_t = a_t$, con $\phi_4(B) = 1 - \sum_{j=1}^4 \phi_j B^j$ y $\Phi_2(B^{12}) = 1 - \Phi_1 B^{12} - \Phi_2 B^{24}$, entonces su ecuación final es

$$E_t = \sum_{j=1}^4 \phi_j E_{t-j} + \Phi_1 E_{t-12} - \sum_{j=1}^4 \phi_j \Phi_1 E_{t-j-12} + \Phi_2 E_{t-24} - \sum_{j=1}^4 \phi_j \Phi_2 E_{t-j-24} + a_t, a_t \text{ RB} \sim N(0, \sigma_a^2)$$

Como este modelo sólo tiene parte AR, será invertible, pero para estacionario se requiere que los polinomios $\phi_4(B)$ y $\Phi_2(B^{12})$ tengan cada uno todas sus raíces (4 raíces en el primero y 24 raíces en el segundo) con módulo mayor a 1.

7) No se usaron todos los modelos ARMA para el error estructural previamente identificados. Los modelos que se presentan en este taller son los siguientes:

- **Modelo 3b: MRLM con errores AR(14):**
Ecuación estructural: $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \dots + \delta_{11} I_{11,t} + E_t$,
Ecuación ARMA: $E_t = \sum_{j=1}^{14} \phi_j E_{t-j} + a_t, \{a_t\}_{t \in \mathbb{Z}^+}$ un RB $\sim N(0, \sigma_a^2)$
- **Modelo 3c: MRLM con errores ARMA(2,10) pero sólo con parámetros ϕ_1, ϕ_2 y θ_{10}**
Ecuación estructural: $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \dots + \delta_{11} I_{11,t} + E_t$,
Ecuación ARMA: $E_t = \phi_1 E_{t-1} + \phi_2 E_{t-2} + a_t + \theta_{10} a_{t-10}, \{a_t\}_{t \in \mathbb{Z}^+}$ un RB $\sim N(0, \sigma_a^2)$
- **Modelo 3d: MRLM con errores ARMA(3,2):** Ecuación estructural: $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \dots + \delta_{11} I_{11,t} + E_t$,
Ecuación ARMA: $E_t = \sum_{j=1}^3 \phi_j E_{t-j} + a_t + \sum_{i=1}^2 \theta_i a_{t-i}, \{a_t\}_{t \in \mathbb{Z}^+}$ un RB $\sim N(0, \sigma_a^2)$
- **Modelo 3e: MRLM con errores ARMA(4,0)(2,0)[12]** Ecuación estructural: $Y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \delta_1 I_{1,t} + \delta_2 I_{2,t} + \dots + \delta_{11} I_{11,t} + E_t$,
Ecuación ARMA: $E_t = \sum_{j=1}^4 \phi_j E_{t-j} + \Phi_1 E_{t-12} - \sum_{j=1}^4 \phi_j \Phi_1 E_{t-j-12} + \Phi_2 E_{t-24} - \sum_{j=1}^4 \phi_j \Phi_2 E_{t-j-24} + a_t, \{a_t\}_{t \in \mathbb{Z}^+}$ un RB $\sim N(0, \sigma_a^2)$

- a) Comparación de ajustes por gráficas, no se distingue diferencias, todos aparentan buen ajuste. Por AIC y BIC, son modelos 3c (menor BIC) y 3e (segundo menor BIC y menor AIC).

| Modelo | p | AIC | BIC |
|----------|----|-----------|-----------|
| modelo3 | 15 | 162.56684 | 217.96720 |
| modelo3b | 29 | 58.87909 | 103.79798 |
| modelo3c | 18 | 59.85563 | 85.10127 |
| modelo3d | 20 | 59.11369 | 87.39776 |
| modelo3e | 21 | 57.64875 | 86.91460 |

Ecuación ajustada del modelo 3e:

ec. ajustada estructural:

$$\hat{Y}_t = 565.94 + 1.0028t + 0.01699t^2 - 8.4331 \times 10^{-5}t^3 + 24.22I_{1,t} - 14.27I_{2,t} + 77.47I_{3,t} + 92.77I_{4,t} + 154.34I_{5,t} + 126.35I_{6,t} + 76.22I_{7,t} + 33.91I_{8,t} - 8.53I_{9,t} - 5.34I_{10,t} - 36.03I_{11,t} + \hat{E}_t$$

ec. ajustada del ARMA: $\hat{E}_t = 0.6613\hat{E}_{t-1} + 0.1796\hat{E}_{t-2} + 0.1286\hat{E}_{t-3} - 0.1391\hat{E}_{t-4} + 0.2218\hat{E}_{t-12} - 0.6613 \times 0.2218\hat{E}_{t-13} - 0.1796 \times 0.2218\hat{E}_{t-14} - 0.1286 \times 0.2218\hat{E}_{t-15} + 0.1391 \times 0.2218\hat{E}_{t-16} + 0.0736\hat{E}_{t-24} - 0.6613 \times 0.0736\hat{E}_{t-25} - 0.1796 \times 0.0736\hat{E}_{t-26} - 0.1286 \times 0.0736\hat{E}_{t-27} + 0.1391 \times 0.0736\hat{E}_{t-28}$,

donde \hat{E}_t es el ARMA ajustado o ciclos ajustados y \hat{E}_t es el residuo estructural en el tiempo t , con

$$\hat{E}_t = Y_t - (565.94 + 1.0028t + 0.01699t^2 - 8.4331 \times 10^{-5}t^3 + 24.22I_{1,t} - 14.27I_{2,t} + 77.47I_{3,t} + 92.77I_{4,t} + 154.34I_{5,t} + 126.35I_{6,t} + 76.22I_{7,t} + 33.91I_{8,t} - 8.53I_{9,t} - 5.34I_{10,t} - 36.03I_{11,t})$$

- b) Después de la identificación de procesos estacionarios para los errores estructurales E_t y luego de ajustar los modelos de regresión con E_t modelado según un ARMA, se valida si para sus respectivos errores de ajuste a_t se cumplen supuesto de RB y normalidad.

- En todas las gráficas de residuales de ajuste y que ahora corresponden a $\hat{a}_t = Y_t - \hat{Y}_t$, no se encuentra evidencia contra media en cero, varianza constante y no parecen haber ciclos (no se observa a primera vista este tipo de patrón o cualquier otro que indique correlaciones).

- Con tests ACF, donde definimos $\rho(k) = \text{corr}(a_t, a_{t+k})$ y PACF con $\phi_{kk} = \text{corr}(a_t, a_{t+k} | a_{t+1}, \dots, a_{t+k-1})$ con $k = 1, 2, \dots, 48$, construidos con los residuales del ajuste:

$$H_0 : \rho(k) = \text{corr}(a_t, a_{t+k}) = 0 \text{ vs. } H_1 : \rho(k) \neq 0,$$

$$\text{estadístico de prueba } \hat{\rho}(k) = \widehat{\text{corr}}(a_t, a_{t+k}) = \frac{\sum_{t=1}^{156-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=1}^{156} \hat{a}_t^2} \overset{\text{aprox}}{\sim} N(0, \frac{1}{156}) \text{ y criterio de}$$

$$\text{rechazo } |\hat{\rho}(k)| \geq 2/\sqrt{156},$$

Con test PACF, $H_0 : \phi_{kk} = \text{corr}(a_t, a_{t+k} | a_{t+1}, \dots, a_{t+k-1}) = 0 \text{ vs. } H_1 : \phi_{kk} \neq 0$,
estadístico de prueba $\hat{\phi}_{kk} = \widehat{\text{corr}}(a_t, a_{t+k} | a_{t+1}, \dots, a_{t+k-1}) \overset{\text{aprox}}{\sim} N(0, \frac{1}{156})$ y criterio de rechazo $|\hat{\phi}_{kk}| \geq 2/\sqrt{156}$,

sólo en modelo 3c, se rechaza en ACF y PACF para $k = 12$, en los otros tres modelos no hay rechazos en estos tests (aunque en modelo 3d la PACF muestra un corte en $k=22$, se desprecia, el valor estimado es pequeño, ocurre en un k alto y éste no es múltiplo de 12). Luego, puede estar ocurriendo que en modelo 3c para los errores de ajuste a_t existen autocorrelaciones entre las variables separadas 12 períodos y eso implica que el proceso no es RB, o bien, por ser un k alto pudiera ser debido a la influencia de observaciones extremas en la estimación de ACF y PACF en ese k , y éstas observaciones claramente se ven en los gráficos de residuos y entre ellas hay dos que están bastante alejadas del resto.

- Tests Ljung-Box & Box-Pierce, con $\rho(k) = \text{corr}(a_t, a_{t+k})$, cada uno evaluado con $m = 6, 12, 18, 24, 30, 36, 42, 48$:

$$H_0 = \rho(1) = \rho(2) = \dots = \rho(m) = 0 \text{ vs. } H_1 : \rho(k) \neq 0 \text{ para al menos un } k, k = 1, 2, \dots, m,$$

estadístico de prueba Box-Pierce, $Q_{BP} = 156 \sum_{k=1}^m [\hat{\rho}(k)]^2 \stackrel{\text{aprox}}{\sim} \chi_m^2$ y estadístico de

prueba Ljung-Box, $Q_{LB} = 156 \times 158 \sum_{k=1}^m \frac{[\hat{\rho}(k)]^2}{156-k} \stackrel{\text{aprox}}{\sim} \chi_m^2$ y criterios de rechazo VP =

$P(\chi_m^2 \geq Q_{BP})$, VP = $P(\chi_m^2 \geq Q_{LB})$ pequeños, en todas las pruebas los valores P no son pequeños, por tanto para ninguno de los modelos 3b, 3c, 3d y 3e se encuentra evidencia en contra de a_t un ruido blanco, aunque para el modelo 3c los valores P son más pequeños que para el resto de modelos.

Conclusión sobre evaluación RB: Los modelos 3b, 3d, y 3e son válidos en este supuesto. Para modelo 3c aunque rechaza con ACF y PACF para $k=12$, vemos que en valor absoluto, las estimaciones de estas funciones en ese k son valores menores a 0.2 (no son grandes) y que tests Ljung-Box y Box-Pierce no rechazaron y además es muy probable que las observaciones más extremas o atípicas en los residuos son las que inflan estas dos estimaciones, luego se decide no rechazar RB en este modelo.

- **Normalidad:** $H_0 : a_t \sim \text{normal}$ vs. $H_1 : a_t \not\sim \text{normal}$, mediante test Shapiro Wilk y gráficos de probabilidad normal construidos con los residuales del ajuste \hat{a}_t , el test Shapiro produce valores P muy pequeños para los cuatro modelos, sin embargo como este test es sensible a presencia de observaciones atípicas (rechazar normalidad por la influencia de una o dos observaciones extremas) procedemos a evaluar gráficos de probabilidad y en estos, *excluyendo las dos observaciones más extremas* (esquina superior derecha en estas gráficas) concluimos que efectivamente se debe rechazar el supuesto de normalidad en errores de ajuste de modelos 3b y 3d; en los otros dos modelos hay algo de desviación con respecto a la recta de probabilidad normal, pero no es grave.

En conclusión, consideramos válidos los modelos 3c y 3e.

- c) Comparando pronósticos, según medidas de precisión de pronósticos puntuales y por I.P, los mejores en pronósticos son los modelo 3d (segundo mejor MAE, segundo mejor MAPE, segunda mejor amplitud media I.P), modelo 3e (mejor MAE, mejor MAPE, tercer mejor RMSE, tercer mejor en amplitud media I.P), aunque en todos los modelos con errores ARMA se disminuyó la cobertura a un 83.33% comparados con el modelo inicial.

| Modelo | Amplitud I.P | Cobertura I.P (%) | RMSE | MAE | MAPE |
|----------|--------------|-------------------|----------|----------|---------|
| modelo3 | 54.04210 | 100.00 | 12.68132 | 10.73177 | 1.26510 |
| modelo3b | 41.27881 | 83.33 | 13.22929 | 11.67177 | 1.35617 |
| modelo3c | 43.69779 | 83.33 | 12.56558 | 10.12848 | 1.21218 |
| modelo3d | 42.43001 | 83.33 | 12.91548 | 10.04257 | 1.20661 |
| modelo3e | 43.27902 | 83.33 | 13.07821 | 9.36264 | 1.12641 |

Comparando gráficamente, los pronósticos de modelos 3d y 3e se acercan más a los valores reales de la validación cruzada.

Ecuación de pronóstico del modelo 3e:

Ecuación estructural de pronóstico:

$$\hat{Y}_{156}(L) = 565.94 + 1.0028(156+L) + 0.01699(156+L)^2 - 8.4331 \times 10^{-5}(156+L)^3 + 24.22I_{1,156+L} - 14.27I_{2,156+L} + 77.47I_{3,156+L} + 92.77I_{4,156+L} + 154.34I_{5,156+L} + 126.35I_{6,156+L} + 76.22I_{7,156+L} + 33.91I_{8,156+L} - 8.53I_{9,156+L} - 5.34I_{10,156+L} - 36.03I_{11,156+L} + \hat{E}_{156}(L)$$

Ecuación de pronóstico del ARMA:

$$\hat{E}_{156}(L) = 0.6613\hat{E}_{156}(L-1) + 0.1796\hat{E}_{156}(L-2) + 0.1286\hat{E}_{156}(L-3) - 0.1391\hat{E}_{156}(L-4) + 0.2218\hat{E}_{156}(L-12) - 0.6613 \times 0.2218\hat{E}_{156}(L-13) - 0.1796 \times 0.2218\hat{E}_{156}(L-14) - 0.1286 \times 0.2218\hat{E}_{156}(L-15) + 0.1391 \times 0.2218\hat{E}_{156}(L-16) + 0.0736\hat{E}_{156}(L-24) - 0.6613 \times 0.0736\hat{E}_{156}(L-25) - 0.1796 \times 0.0736\hat{E}_{156}(L-26) - 0.1286 \times 0.0736\hat{E}_{156}(L-27) + 0.1391 \times 0.0736\hat{E}_{156}(L-28),$$

donde

$$\hat{E}_{156}(L-j) = \begin{cases} \text{residuos estructural } \hat{E}_{156+L-j} & \text{si } L-j \leq 0 \\ \text{Pronóstico de los ciclos, } L-j \text{ períodos después de } t=156 & \text{si } L-j > 0, \end{cases}$$

para $j = 1, 2, 3, 4, 12, 13, 14, 15, 16, 24, 25, 26, 27, 28$.

- d) Se recomienda modelo 3e por buen ajuste entre los modelos comparados, de los mejores en pronóstico y es un modelo válido, por tanto puede recomendarse.