



Bolsa de Comercio de Santa Fe

Presidente

Escrib. Martín Vigo Lamas



Director Ejecutivo

Escrib. Martín Vigo Lamas

Directora

Lic. Lucrecia D'Jorge

Coordinador

Mg. Pedro P. Cohan

Integrantes del CES y Presa

<https://www.bcsf.com.ar/ces/quienes-somos.php>

E-mail de contacto:

ces@bcsf.com.ar

Sitio web:

<http://ces.bcsf.com.ar> o

<http://www.bcsf.com.ar>

Redes:



@BCSFOficial



@BCSF_Oficial



@bcsfoficial



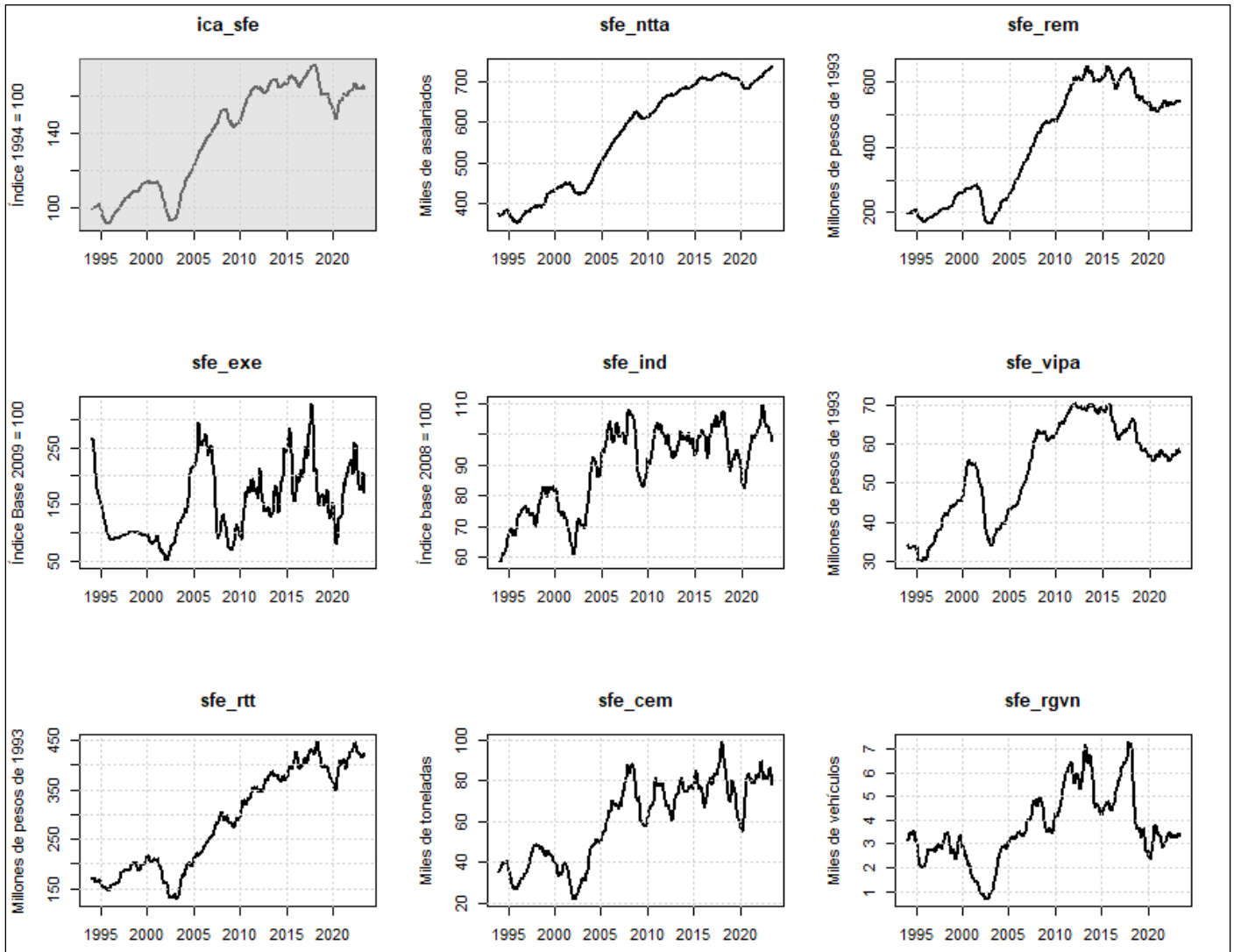
Bolsa de Comercio de Santa Fe

REGRESIÓN LINEAL MÚLTIPLE DEL ICASFE & SUS SERIES COMPONENTES

Noviembre de 2023

ICASFe y sus series componentes en niveles

A continuación, se expone la gráfica del ICASFe y sus ocho componentes en niveles, expresadas en la Unidad de Medida (UM) correspondiente a cada serie. En todos los casos, los datos comienzan en 1994.01 y se extienden hasta el último mes publicado en 2023 a la fecha de este reporte.

Gráfico 1 | ICASFe & sub-indicadores en niveles**Referencias:**

ICASFe | Índice Compuesto Coincidente de Actividad Económica de la provincia de Santa Fe | UM: Índice base 1994 = 100 | Código: ica_sfe

(1) N° de puestos de trabajo registrados en la provincia | UM: Miles de asalariados | Código: sfe_ntta

(2) Expectativas empresarias para incrementar el personal | UM: Índice base 2009 = 100 | Código: sfe_exe

(3) Masa de remuneraciones reales percibida por los asalariados | UM: Millones de pesos de 1993 | Código: sfe_rem

(4) Índice de producción industrial | UM: Índice base 2008 = 100 | Código: sfe_ind

(5) Ventas reales de supermercados | UM: Millones de pesos de 1993 | Código: sfe_vipa

(6) Consumo de cemento Pórtland | UM: Miles de toneladas | Código: sfe_cem

(7) Patentamiento de vehículos nuevos | UM: Miles de vehículos | Código: sfe_rgvn

(8) Recaudación tributaria de la provincia y coparticipación | UM: Millones de pesos de 1993 | Código: sfe_rtt

La metodología de agregación utilizada para construir el índice coincidente¹, implica que cada serie contribuye con 1/n en el algoritmo, pero cada componente aporta en realidad una ponderación implícita variable en términos mensuales que, como criterio de base, penaliza la volatilidad cíclica (debido al proceso de estandarización de las tasas de cambio). En el ejercicio llevado a cabo en este documento el objetivo es considerar de forma complementaria, la importancia relativa de cada sub-indicador e identificar relaciones de largo plazo (hechos estilizados) entre los distintos elementos y el flujo de actividad económica. A tales fines se plantea una regresión lineal múltiple que considera el ICASFe como variable dependiente y a sus componentes como variables explicativas.

Modelo de regresión lineal múltiple con series en niveles

$$ica_{sfe} = \beta_0 + \beta_1 \cdot sfe_{ntta} + \beta_2 \cdot sfe_{rem} + \beta_3 \cdot sfe_{exe} + \beta_4 \cdot sfe_{ind} + \beta_5 \cdot sfe_{vipa} + \beta_6 \cdot sfe_{rtt} + \beta_7 \cdot sfe_{cem} + \beta_8 \cdot sfe_{rgvn} + \varepsilon$$

Código y salida en R

Se realizó la regresión lineal múltiple en el *software* R-Studio, utilizando la librería “dynlm” (código: `dynlm::dynlm(ica_sfe ~ sfe_ntta + sfe_rem + sfe_exe + sfe_ind + sfe_vipa + sfe_rtt + sfe_cem + sfe_rgvn)`). Los resultados obtenidos fueron los siguientes:

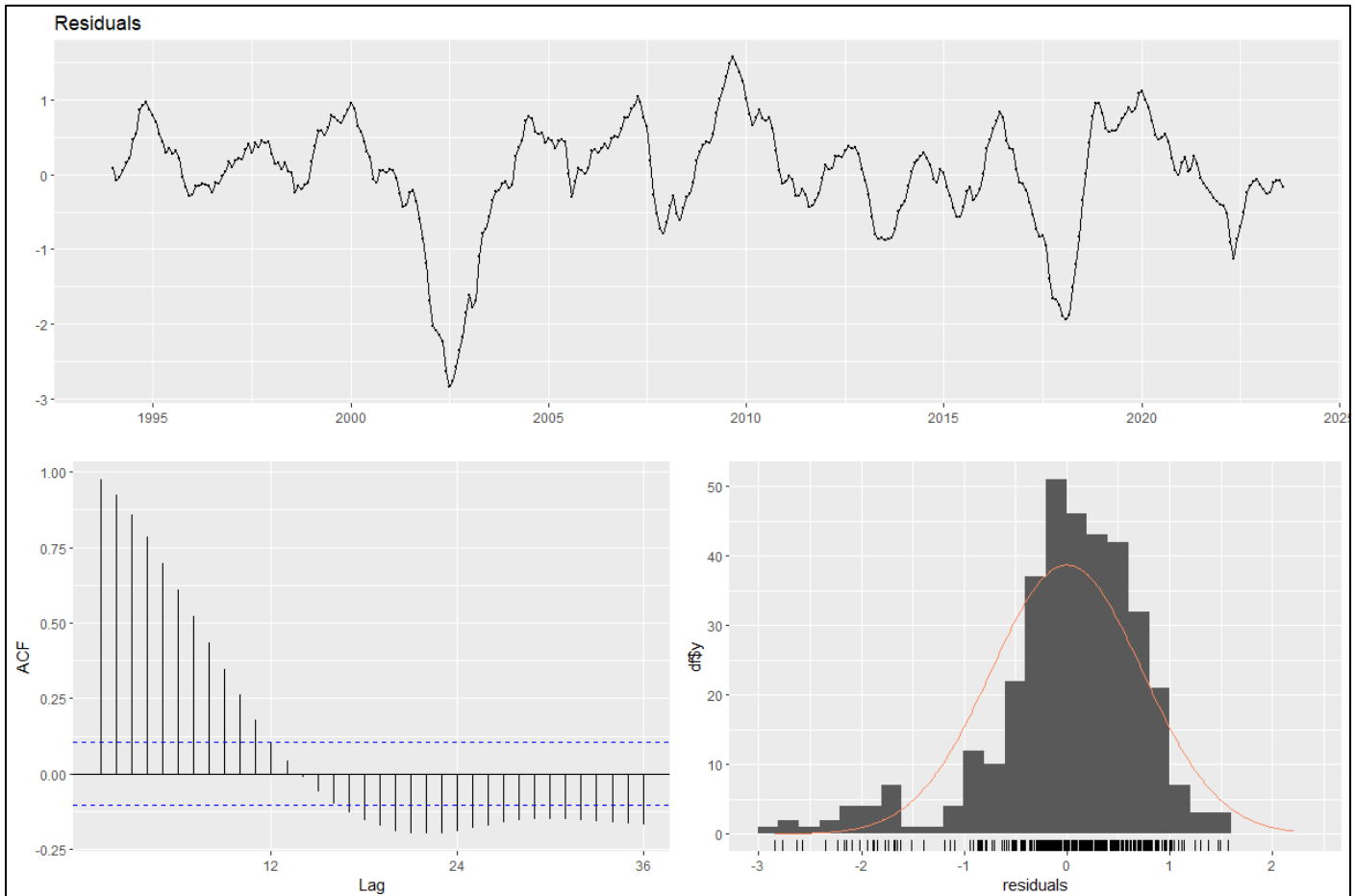
Serie	Coeficiente		Significancia	
			t value	Pr(> t)
Const.	β_0	24.3700	35.3	0.00000
sfe_ntta	β_1	0.0667	27.61	0.00000
sfe_rem	β_2	0.0276	11.28	0.00000
sfe_exe	β_3	0.0204	22.54	0.00000
sfe_ind	β_4	0.1844	21.77	0.00000
sfe_vipa	β_5	0.3677	29.77	0.00000
sfe_rtt	β_6	0.0398	18.9	0.00000
sfe_cem	β_7	0.1441	17.01	0.00000
sfe_rgvn	β_8	1.2660	15.8	0.00000
R ² =0.9993		R ² -ajustado =0.9993		
Durbin-Watson test: DW = 0.047303				

Como era de esperar, la salida indica que todos los elementos tienen coeficientes significativos, con un nivel explicativo conjunto (R² ajustado) de 99,9%. Al mismo tiempo, quedan en evidencia las relaciones de largo plazo entre los sub-indicadores y el ICASFe. Al respecto, 1.000 empleos registrados se asocian con 0,067 puntos de actividad, un millón de pesos de 1993 de remuneraciones se asocia con 0,028 puntos de actividad, un punto índice de expectativas empresarias, con 0,020 puntos de actividad, un punto índice de producción industrial con 0,184, un millón de pesos

¹ Véase metodología completa (archivo “metodología.pdf”) en <https://www.bcsf.com.ar/ces/icasfe.php>

de 1993 de ventas minoristas con 0,368, un millón de pesos de 1993 de recaudación, con 0,040 de actividad, 1.000 toneladas de cemento con 0,144 y 1.000 automóviles patentados con 1,266 puntos de actividad.

Sin embargo, observando el estadístico “Durbin-Watson”, que prueba la autocorrelación de los residuos del modelo, la regresión evidencia autocorrelación positiva de primer orden². Esta situación también queda registrada visualmente al verificar los residuos, donde se observa un patrón, autocorrelación y una distribución que no es normal:



Fuente: elaboración propia en R-Studio (código: “forecast::checkresiduals(modelo_rlm)”).

Una de las fuentes del problema de la autocorrelación de los residuos reside en que las series de tiempo utilizadas no son estacionarias (esto explica la silueta del ACF de la gráfica precedente)³. Por este motivo, se resuelve transformar las mismas, diferenciándolas una vez. Además, el objetivo del presente trabajo se responde de mejor manera a través

² El estadístico tiene la forma: $d = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}$, donde e representa el residuo de la regresión. De esta manera, d es aproximadamente igual a $2(1 - r)$, donde r es el coeficiente de autocorrelación de primer orden de los residuos. Entonces, el valor del estadístico entre 0 y 4, indicando que no existe autocorrelación si $d=2$. Si el mismo es significativamente menor que 2 hay evidencia de correlación positiva.

³ La no estacionariedad, habitual en series económicas, queda en evidencia a simple vista en la Gráfica N° 1. También se efectuaron tests de Dickey-Fuller aumentado, en el cual se verifica que no hay evidencia de estacionariedad en las series en niveles, pero sí en su primera diferencia.

de un cálculo de elasticidades, trabajando en logaritmos. De esta forma, los coeficientes de la regresión se corresponden con el porcentaje en que cambia el ICASFe cuando sus series componentes cambian en un 1%.

En síntesis, para incorporar estas cuestiones, resulta conveniente probar un modelo similar, pero con las series en la primera diferencia de sus logaritmos.

Modelo de regresión lineal múltiple con series en la primera diferencia de sus logaritmos

El modelo propuesto es el siguiente:

$$\Delta \ln(ica_{sfe}) = \beta_0 + \beta_1 \cdot \Delta \ln(sfe_{ntta}) + \beta_2 \cdot \Delta \ln(sfe_{rem}) + \beta_3 \cdot \Delta \ln(sfe_{exe}) + \beta_4 \cdot \Delta \ln(sfe_{ind}) \\ + \beta_5 \cdot \Delta \ln(sfe_{vipa}) + \beta_6 \cdot \Delta \ln(sfe_{rtt}) + \beta_7 \cdot \Delta \ln(sfe_{cem}) + \beta_8 \cdot \Delta \ln(sfe_{rgvn}) + \varepsilon$$

Donde Δ es utilizado para representar la primera diferencia.

Código y salida en R

Nuevamente, se realizó la regresión lineal múltiple con el *software* R-Studio, utilizando la librería “dynlm” (código: “modelo_diflog <- dynlm::dynlm(diff(log(ica_sfe)) ~ diff(log(sfe_ntta)) + diff(log(sfe_rem)) + diff(log(sfe_exe)) + diff(log(sfe_ind)) + diff(log(sfe_vipa)) + diff(log(sfe_rtt)) + diff(log(sfe_cem)) + diff(log(sfe_rgvn)))”).

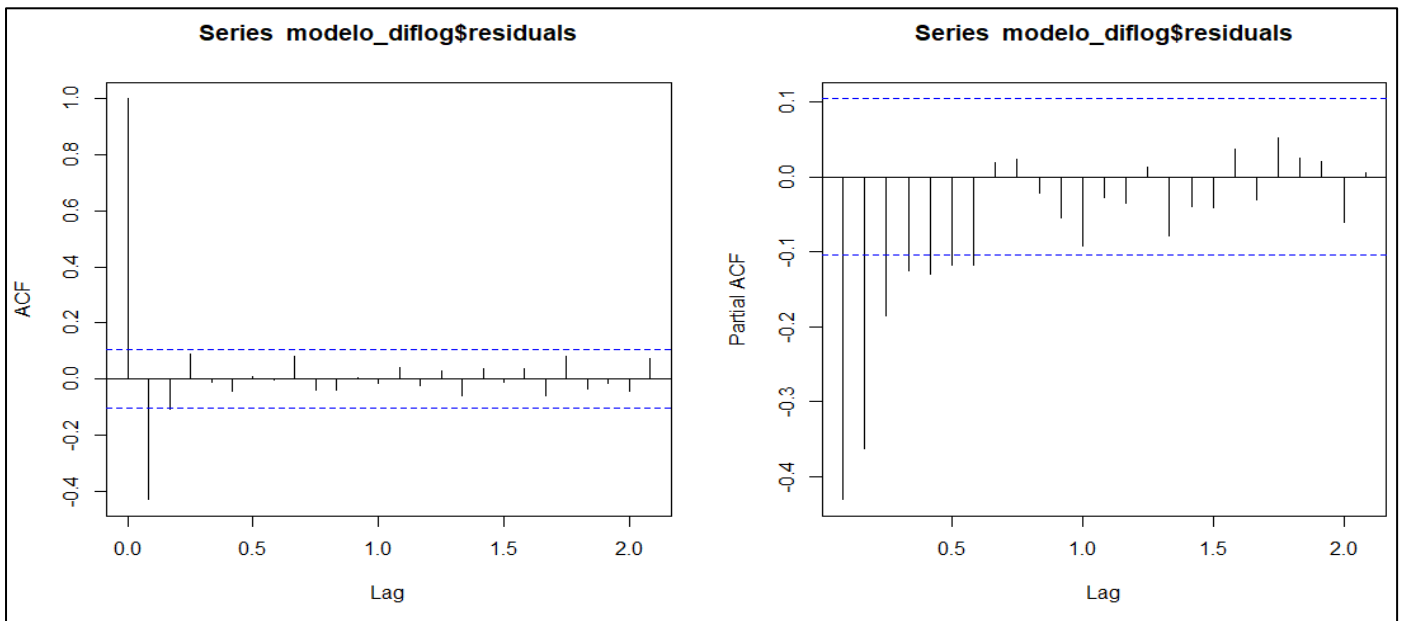
Los resultados obtenidos fueron los siguientes:

Serie	Coeficiente		Significancia	
			t value	Pr(> t)
Const.	β_0	0.0000	-2.25E+12	0.00000
$\Delta \ln(sfe_{ntta})$	β_1	0.3113	6.13E+13	0.00000
$\Delta \ln(sfe_{rem})$	β_2	0.0906	5.69E+13	0.00000
$\Delta \ln(sfe_{exe})$	β_3	0.0230	7.30E+13	0.00000
$\Delta \ln(sfe_{ind})$	β_4	0.0935	6.53E+13	0.00000
$\Delta \ln(sfe_{vipa})$	β_5	0.1091	6.51E+13	0.00000
$\Delta \ln(sfe_{rtt})$	β_6	0.0846	6.40E+13	0.00000
$\Delta \ln(sfe_{cem})$	β_7	0.0494	6.27E+13	0.00000
$\Delta \ln(sfe_{rgvn})$	β_8	0.0332	6.27E+13	0.00000
$R^2=1$		R^2 -ajustado =1		
Durbin-Watson test: DW = 2.8612				

Como puede verse, todos los coeficientes resultan significativos y el poder explicativo del modelo es del 100%, lo cual es de esperarse, dado que los regresores son las series que se agregan para la construcción del ICASFe. Sin embargo, este modelo brinda información acerca de cuánto aporta cada serie a las tasas de cambio del índice compuesto, es decir, los coeficientes indican cuánto varía porcentualmente el ICASFe ante una variación de un 1% en la serie

independiente. Por ejemplo, un cambio del 1% en NTTA, manteniendo todo lo demás constante, implica un cambio del 0,31% en el ICASFe; un cambio del 1% en REM, implicaría un cambio del 0,09% en el ICASFe; un cambio del 1% en VIPA implica un cambio del 0,11% en el ICASFe; y así con el resto de los componentes.

Si bien este modelo mejora los resultados obtenidos previamente, al realizar un análisis sobre los residuos, se sigue observando que los mismos no son ruido blanco. Esto también queda reflejado en el estadístico que evidencia el Durbin-Watson⁴, y en las funciones de autocorrelación y autocorrelación parcial, que se presentan gráficamente a continuación:



Fuente: elaboración propia en R-Studio (funciones utilizadas: “stats::acf(modelo_diflog\$residuals)” y “stats::pacf(modelo_diflog\$residuals)”).

Estas gráficas permiten interpretar que los residuos del modelo responden a un proceso MA(1). Por lo tanto, a continuación, se prueba un modelo de regresión lineal con residuos ARIMA⁵.

Modelo de regresión lineal múltiple con series en la primera diferencia de sus logaritmos y residuos ARIMA

El modelo propuesto es el siguiente:

$$\Delta \ln(ica_{sfe}) = \beta_0 + \beta_1 \cdot \Delta \ln(sfe_{ntta}) + \beta_2 \cdot \Delta \ln(sfe_{rem}) + \beta_3 \cdot \Delta \ln(sfe_{exe}) + \beta_4 \cdot \Delta \ln(sfe_{ind}) \\ + \beta_5 \cdot \Delta \ln(sfe_{vipa}) + \beta_6 \cdot \Delta \ln(sfe_{rtt}) + \beta_7 \cdot \Delta \ln(sfe_{cem}) + \beta_8 \cdot \Delta \ln(sfe_{rgvn}) + \mu_t$$

con: $\mu_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$

Donde Δ es utilizado para representar la primera diferencia.

Código y salida en R

En este caso la librería utilizada también fue “dynlm”. A continuación, se muestra el código:

⁴ Cuando d es significativamente mayor a 2, los residuos sucesivos se encuentran correlacionados de manera negativa.

⁵ Modelo Autorregresivo, integrado de medias móviles o ARIMA (acrónimo del inglés *Auto Regressive Integrated Moving Average*).

```

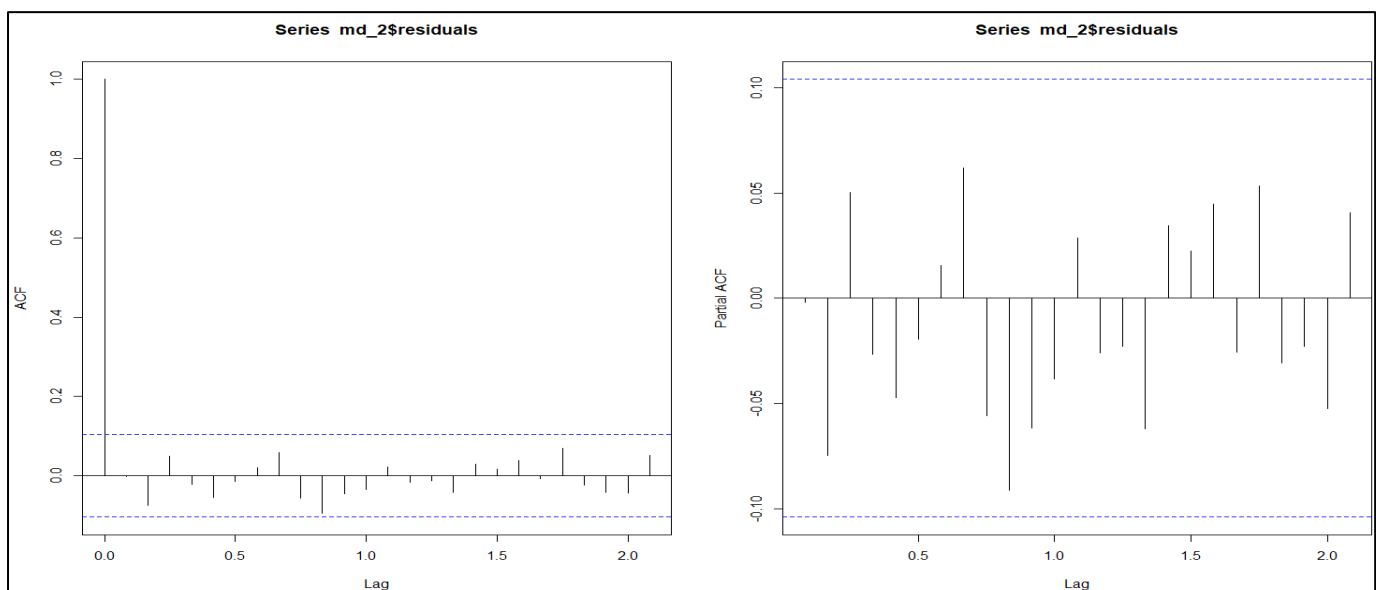
"model.resid <- modelo_diflog$residuals
ma1 <- stats::fitted(forecast::Arima(model.resid, order = c(0, 0, 1)))
modelo_diflog_ma <- stats::update(modelo_diflog, . ~ . + ma1)"

```

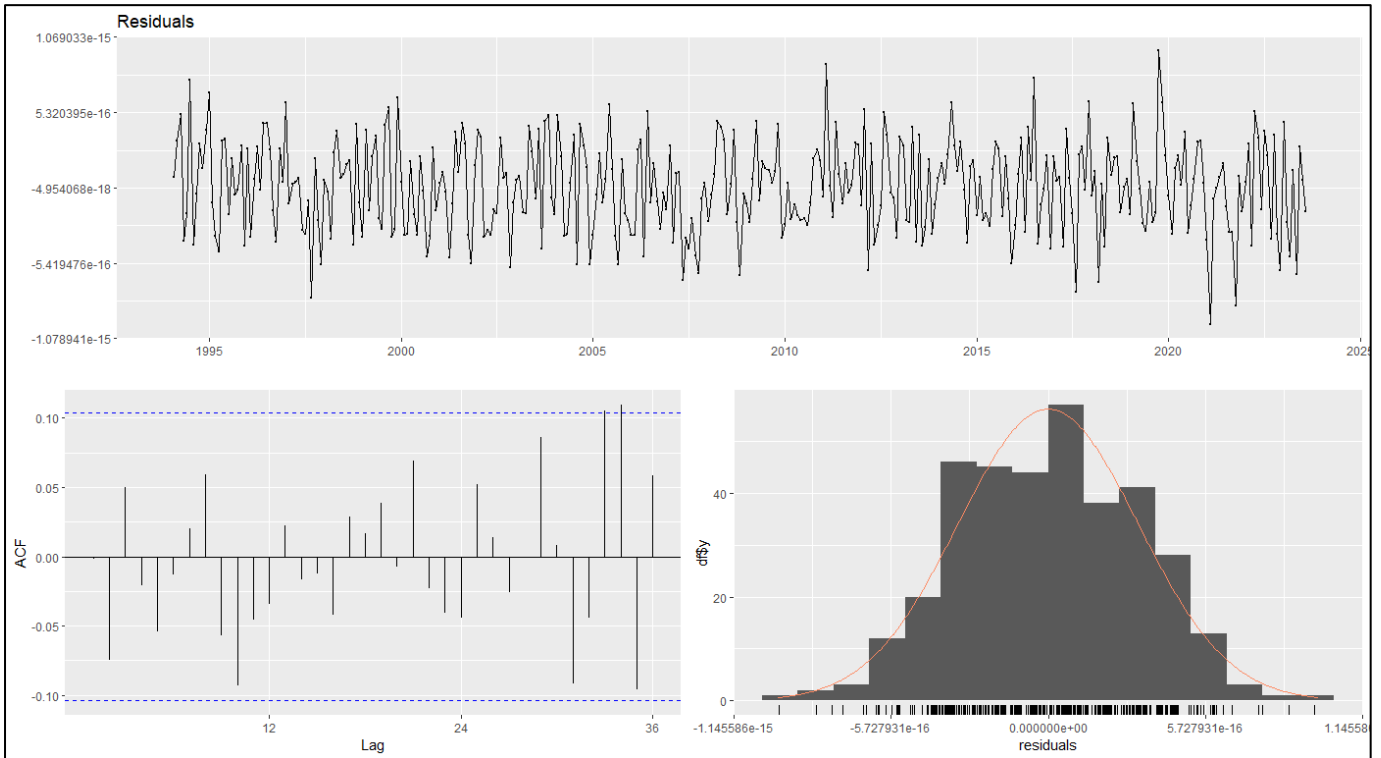
Los resultados obtenidos fueron los siguientes:

Serie	Coeficiente		Significancia	
			t value	Pr(> t)
Const.	β_0	0.0000	-2.80E+12	0.00000
$\Delta \ln(\text{sfe_ntta})$	β_1	0.3113	7.62E+13	0.00000
$\Delta \ln(\text{sfe_rem})$	β_2	0.0906	7.07E+13	0.00000
$\Delta \ln(\text{sfe_exe})$	β_3	0.0230	9.07E+13	0.00000
$\Delta \ln(\text{sfe_ind})$	β_4	0.0937	8.12E+13	0.00000
$\Delta \ln(\text{sfe_vipa})$	β_5	0.1091	8.09E+13	0.00000
$\Delta \ln(\text{sfe_rtt})$	β_6	0.0848	7.96E+13	0.00000
$\Delta \ln(\text{sfe_cem})$	β_7	0.0492	7.79E+13	0.00000
$\Delta \ln(\text{sfe_rgvn})$	β_8	0.0332	7.80E+13	0.00000
ma1	β_9	0.9596	1.38E+01	0.00000
$R^2=1$		$R^2\text{-ajustado}=1$		
Durbin-Watson test: DW = 2.0027				

Como se observa, las estimaciones de los coeficientes permanecen sin variaciones, mientras que los residuos del modelo de regresión son modelados con un ARIMA de orden $pdq=(0, 0, 1)$, de manera que los mismos resultan ruido blanco. En efecto, el test de Durbin-Watson brinda evidencia de que no existe autocorrelación, mientras que el análisis gráfico de los residuos permite llegar a la misma conclusión:



Fuente: elaboración propia en R (funciones utilizadas: "stats::acf(modelo_diflog_ma\$residuals)" y "stats::pacf(modelo_diflog_ma\$residuals)").



Fuente: elaboración propia en R-Studio (código: "forecast::checkresiduals(modelo_diflog_ma)").

Anexos I | Series en la primera diferencia de sus logaritmos y matriz de correlaciones

Gráfico 2 | ICASFe & sub-indicadores en la primera diferencia de sus logaritmos

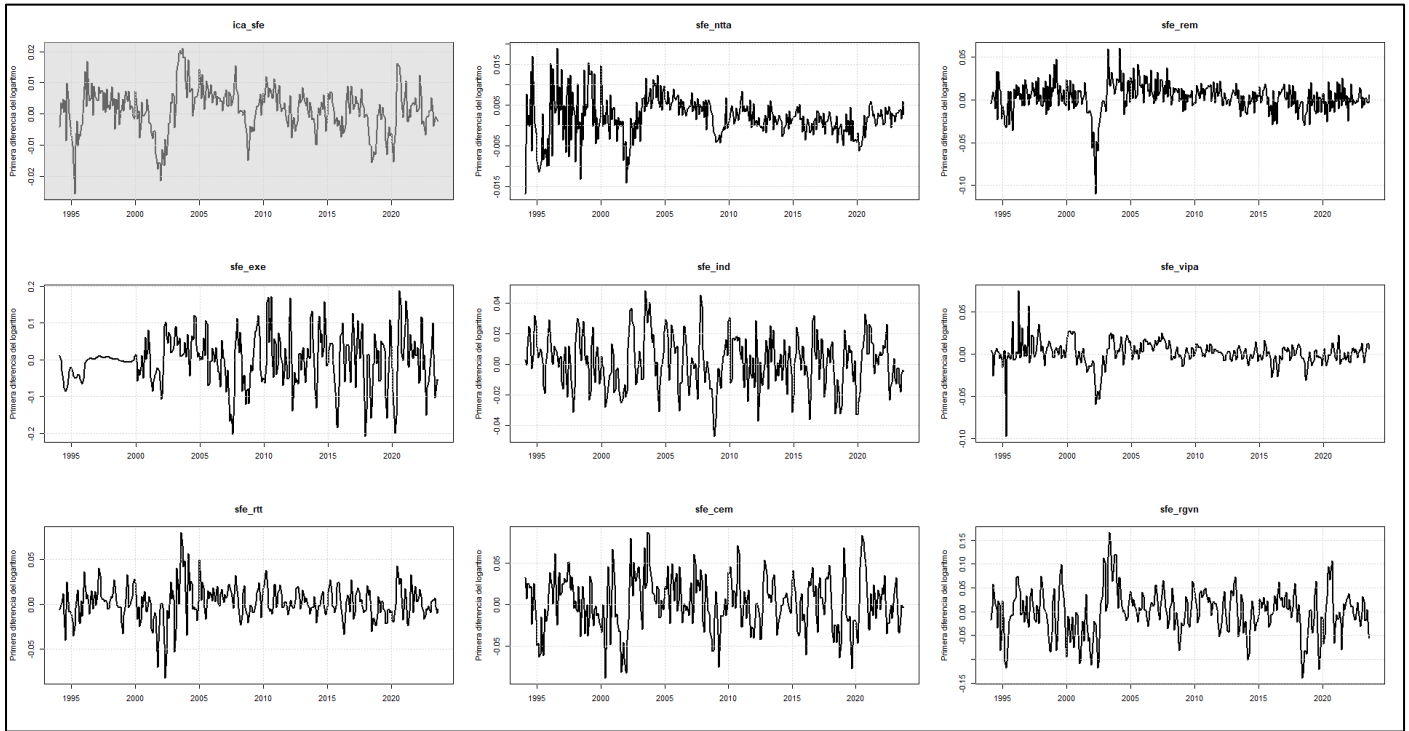
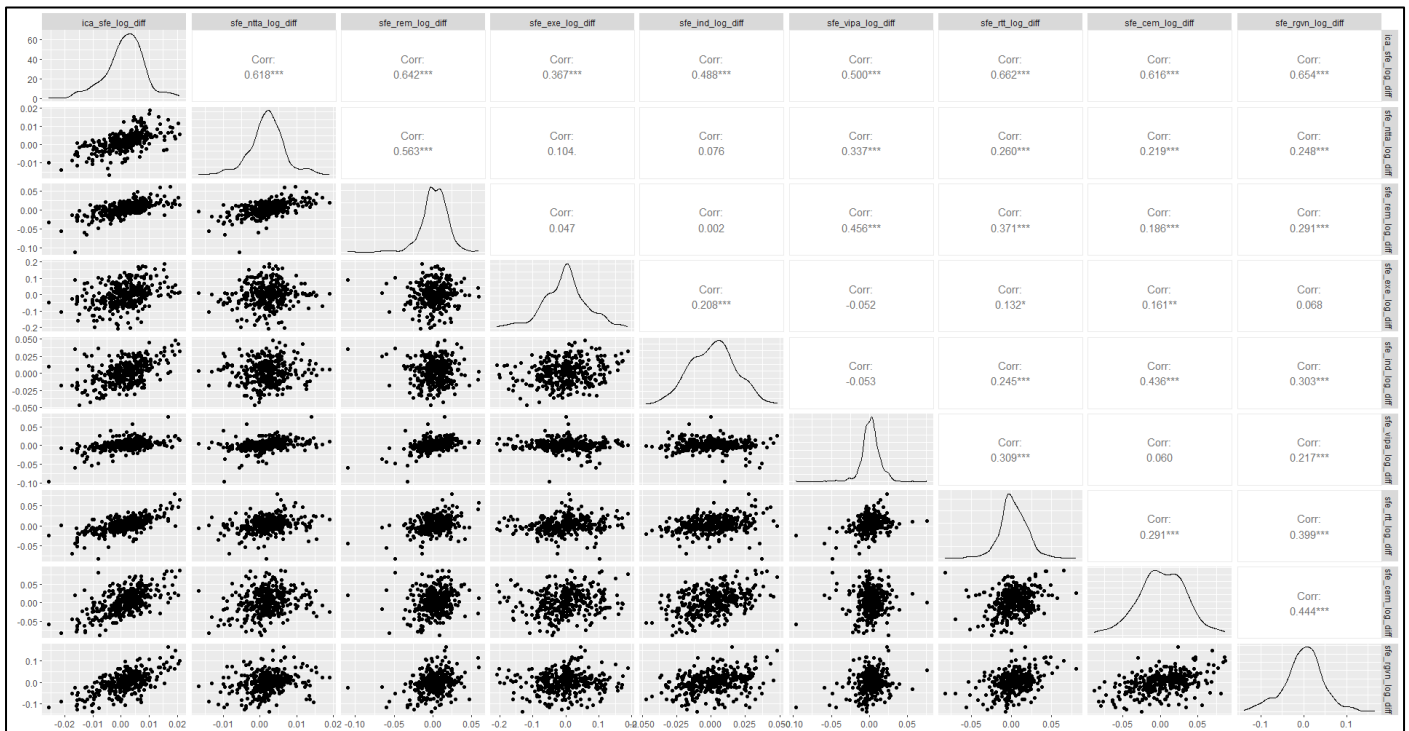


Gráfico 3 | Matriz de correlaciones y gráficas de dispersión | ICASFe & sub-indicadores en la primera diferencia de sus logaritmos



Anexos III | Multicolinealidad

La matriz de correlaciones y gráficas de dispersión de las series en logaritmos dejan en evidencia relaciones de alta asociación lineal entre variables (con coeficientes de correlación mayores a 90% en algunos casos). Aunque esto no resulta un problema a los fines del ejercicio que se realiza en este documento (con fines distintos al modelado propiamente dicho), resulta interesante testear la situación de modo explícito en un anexo.

A tales fines se computa el *Variance Inflation Factor (VIF)* para los dos modelos. Se utiliza la función `vif()` de la librería “car”. A continuación, se presentan los resultados obtenidos para cada uno de los modelos:

> car::vif(modelo_rlm)			
sfe_ntta	sfe_rem	sfe_exe	sfe_ind
60.648	103.538	1.989	7.980
sfe_vipa	sfe_rtt	sfe_cem	sfe_rgvn
14.412	27.902	17.350	8.988
VIF mayor a 10, alta colinealidad.			

> car::vif(modelo_diflog)			
diff(log(sfe_ntta))	diff(log(sfe_rem))	diff(log(sfe_exe))	diff(log(sfe_ind))
1.528	1.778	1.074	1.346
diff(log(sfe_vipa))	diff(log(sfe_rtt))	diff(log(sfe_cem))	diff(log(sfe_rgvn))
1.356	1.398	1.470	1.452
VIF entre 1 y 5, colinealidad aceptable.			

> car::vif(modelo_diflog_ma)			
diff(log(sfe_ntta))	diff(log(sfe_rem))	diff(log(sfe_exe))	diff(log(sfe_ind))
1.534	1.778	1.074	1.347
diff(log(sfe_vipa))	diff(log(sfe_rtt))	diff(log(sfe_cem))	diff(log(sfe_rgvn))
1.356	1.399	1.482	1.453
ma1			
1.018			
VIF entre 1 y 5, colinealidad aceptable.			