

# APLICACIÓN E IMPLEMENTACIÓN DE MODELOS VAR Y DEEP LEARNING

ANÁLISIS DE SERIES TEMPORALES

Profesores: Mg. Rodrigo Del Rosso – Dr. Martín E. Masci

**Alumno: Fernando Gabriel Martínez**

## Consigna Final

### Aplicación e implementación de modelos VAR y deep learning

#### Resumen Ejecutivo:

En el presente trabajo analizamos tres series temporales correspondientes al precio del petróleo, el oro y la plata entre los años 1980 a 2020. Estudiamos la estacionariedad y la relación entre las tres variables. Planteamos un modelo de vectores autorregresivos (VAR) el cual permite realizar predicciones teniendo en cuenta la dependencia entre todas las variables y su pasado. Además permite examinar como la perturbación de una variable afecta a las otras. Verificamos los supuestos del modelo y luego realizamos predicciones para el período que va de agosto 2019 a julio 2020. Comparamos las predicciones anteriores con otros modelos: modelo ingenuo con estacionalidad, modelo autorregresivo de redes neuronales (NNAR) y modelo de redes neuronales Long Short-Term Memory (LSTM).

#### • I) Introducción:

Se analizaron las siguientes tres series temporales con ocurrencia de enero de 1980 a julio de 2020:

- Precio mensual del Oro de acuerdo a la London Bullion Market Association (LBMA) medido en: USD / onza troy
- Precio mensual de la Plata de acuerdo a la London Bullion Market Association (LBMA) medido en: USD / onza troy
- Precio mensual del barril de Petróleo crudo de acuerdo a la WTI (West Texas Intermediate) medido en USD / barril

Dichos datos fueron extraídos de la página del [Fondo Monetario Internacional \(FMI\)](#). Existe una relación entre dichas series. Por un lado el aumento del precio del petróleo, generalmente está acompañado por un aumento de la inflación debido a su influencia en el precio de los bienes y servicios. Por otro lado, uno de los usos del oro es como reserva de valor. Por lo que al aumentar la inflación y ante la pérdida de valor del USD, muchos inversores buscan refugio en él. Esto hace que aumente su precio.

En cuanto a la plata y el oro, ambos actúan de manera similar como reserva de valor. Sin embargo hay que destacar que la plata tiene mayor uso en el ámbito industrial (en la electrónica preferentemente), que como reserva de valor. De allí que no sólo el efecto de reserva de valor actúe en su comportamiento, sino también su uso industrial. En cambio el oro lo hace en el sentido inverso.

Estudiamos lo expuesto anteriormente, mediante el análisis de dichas series temporales. Planteamos un modelo VAR (vectores autorregresivos) entre las variables elegidas, examinando sus supuestos, para finalmente poder usar el modelo y realizar predicciones. Luego comparamos dichas predicciones con aquellas efectuadas por un modelo ingenuo con estacionalidad, un modelo autorregresivo de redes neuronales (NNAR) y finalmente un modelo de redes neuronales Long Short-Term Memory (LSTM).

#### • II) Marco Teórico:

El análisis de los Vectores Autorregresivos (VAR) surge en la década de los 80 y se hicieron populares gracias a la introducción de Sims (1980) para la estimación de ecuaciones simultáneas que pretendía facilitar el análisis sobre la influencia de las políticas fiscales respecto del PBI.

Los modelos VAR constituyen una extensión directa de los modelos ARIMA cuando tenemos más de una serie temporal y se quiere captar las dependencias que puede haber entre estas series. Un caso de ejemplo es la dependencia entre las tasas de crecimiento interanuales del consumo y del PIB. Lógicamente, ambas series estarán correlacionadas; pero esta dependencia no solo será contemporánea (en el mismo instante) sino que también el pasado del PIB puede influir sobre el consumo actual y viceversa. Por supuesto, si mirásemos la función de autocorrelación, observaríamos también que cada serie depende asimismo de sus propios retardos (es decir, tanto el PIB como el consumo dependen de su propio pasado y no solo del pasado de la otra serie).

De acuerdo con Sims los principales supuestos de un modelo VAR son:

- Es un modelo de ecuaciones simultáneas donde cada variable crea sus propios rezagos.
- Es una herramienta para generar pronósticos confiables en el corto plazo.
- Es capaz de separar los efectos pasados que explican al vector de las variables endógenas a través de su pasado o mediante variables autorregresivas.
- La estimación del modelo es sencilla, ya que es posible utilizar el método de los Mínimos Cuadrados Ordinarios (MCO).
- Los valores retardados de todas las ecuaciones aparecen como variables explicativas en todas las ecuaciones.

Una de las principales utilidades de los modelos VAR, además de la realización de predicciones hacia el futuro, es el estudio de la respuesta dinámica de una serie cuando se produce un shock inesperado en la otra serie. Esta respuesta dinámica viene dada por la función de respuesta al impulso, la cual muestra el efecto en los instantes posteriores del shock.

---

Una red neuronal, explicada de una forma simplista, es un conjunto de nodos o unidades de procesamiento que procesa y envía información entre sí. Las unidades de procesamiento se organizan en capas. Hay tres partes normalmente en una red neuronal: una capa de entrada, con unidades que representan los campos de entrada; una o varias capas ocultas; y una capa de salida, con una unidad o unidades que representa el campo o los campos de destino. Las unidades se conectan con fuerzas de conexión variables (o ponderaciones). Los datos de entrada se presentan en la primera capa, y los valores se propagan desde cada neurona hasta cada neurona de la capa siguiente. Al final, se envía un resultado desde la capa de salida. La red aprende examinando los registros individuales, generando una predicción para cada registro y realizando ajustes a las ponderaciones cuando realiza una predicción incorrecta. Este proceso se repite muchas veces y la red sigue mejorando sus predicciones hasta haber alcanzado uno o varios criterios de parada.

Un primer acercamiento al uso de redes neuronales para el problema del pronóstico de valores a futuro es utilizar las conocidas como redes neuronales autorregresivas (NNAR). La clave para el uso de este tipo de redes es considerar los datos ordenados de la serie temporal como entradas de la red, del mismo modo que usamos los valores retrasados en un modelo de autorregresión lineal. Cuando se trata de pronósticos utilizando este algoritmo, la red se aplica de forma iterativa. Para pronosticar un paso adelante, simplemente utilizamos las entradas históricas disponibles. Para pronosticar dos pasos adelante, usamos el pronóstico de un paso como entrada, junto con los datos históricos. Este proceso continúa hasta que hayamos computado todos los pronósticos requeridos.

Las redes neuronales recurrentes (RNN), son un tipo de redes capaces de reconocer y predecir secuencias de datos a lo largo del tiempo. Este tipo de redes se fundamentan en bucles que permiten que la salida de la red o de una parte de ella en un momento dado sirva como entrada de la propia red en el siguiente momento. Para entender el funcionamiento de las RNN, podemos considerar un perceptrón con una sola capa oculta de manera que la salida del perceptrón es utilizada como entrada en la siguiente evaluación. Este bucle en la arquitectura de la red es precisamente lo que permite a la red “recordar” información a lo largo del tiempo. En la medida en la que le añadimos capas, su capacidad de modelado irá creciendo de manera que será capaz de reconocer mayores secuencias cada vez con menor error.

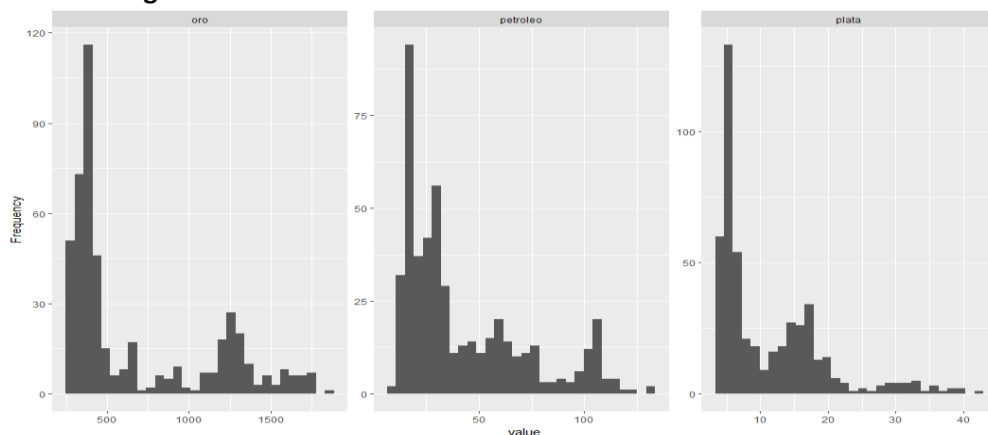
Las RNN tienen un contratiempo importante conocido como problema del desvanecimiento del gradiente: tienen dificultades para aprender dependencias de largo alcance. Cuando se realiza la propagación hacia atrás, es decir, nos movemos hacia atrás en la red y calculamos los gradientes de pérdida (error) con respecto a los pesos, los gradientes tienden a ser cada vez más pequeños a medida que nos movemos hacia atrás en la red. Esto significa que las neuronas en las capas anteriores aprenden muy lentamente en comparación con las neuronas en las capas posteriores en la jerarquía. Las capas anteriores de la red son las más lentas de entrenar. Este es un problema en todos los tipos de redes neuronales, pero particularmente es nocivo para redes en donde lo que se pretende es tener la componente de memoria necesaria para pronóstico de series temporales. Afortunadamente, este problema fue resuelto por (Hochreiter and Schmidhuber 1997) mediante la creación de las LSTM. Las redes de memoria a corto/largo plazo, LSTM, son un tipo especial de RNN, capaz de aprender dependencias a largo plazo.

## • II) Análisis de Resultados:

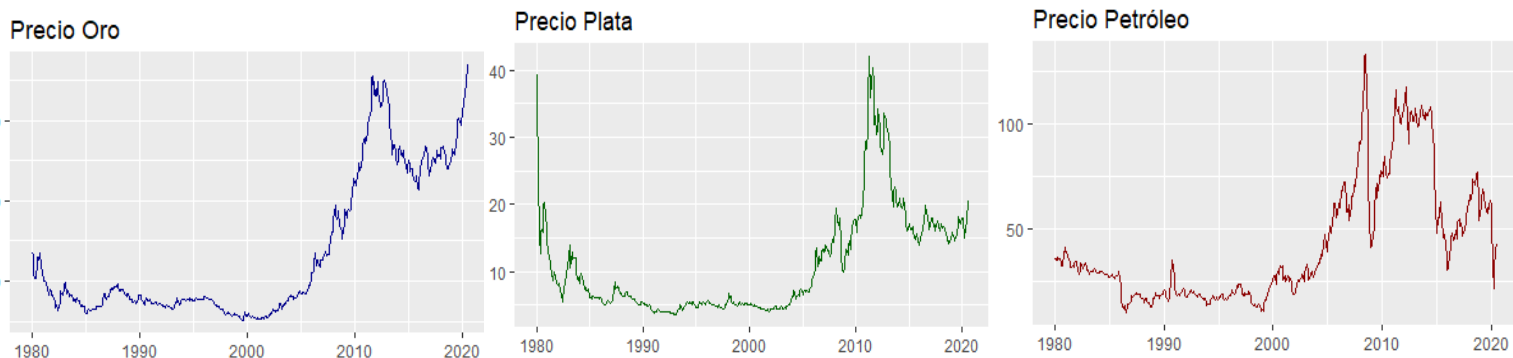
Hicimos un primer análisis exploratorio de las medidas estadísticas descriptivas de las variables:

	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
<b>oro</b>	487	673.97	445.38	413.41	612.32	152.90	256.69	1843.31	1586.62	0.99	-0.52	20.18
<b>plata</b>	487	11.00	7.81	7.22	9.66	4.18	3.65	41.97	38.32	1.50	2.06	0.35
<b>petróleo</b>	487	42.22	29.03	29.99	37.78	19.44	9.88	133.01	123.13	1.11	0.16	1.32

Graficamos los histogramas de cada una:



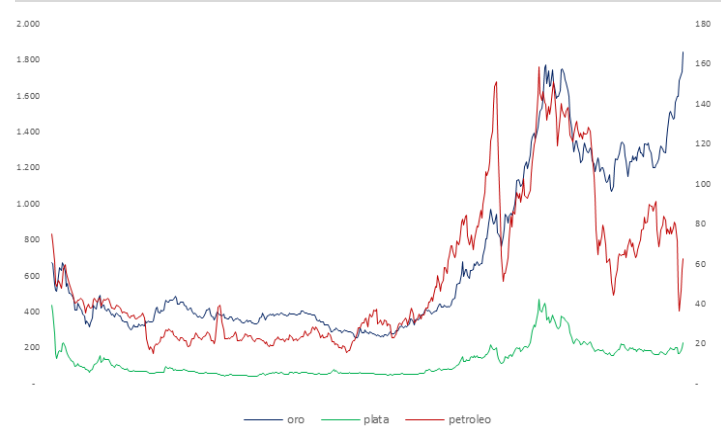
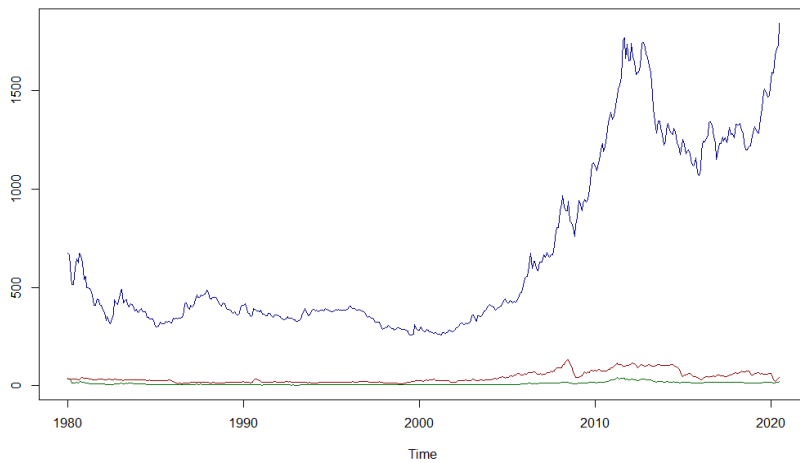
Graficamos las tres series temporales:



Nota: azul → oro, verde → plata, rojo → petróleo

Graficamos las tres series en la misma escala y en dos escalas (petróleo y plata en una misma escala):

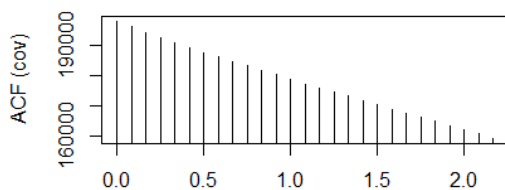
Evolución del Precio de los Commodities



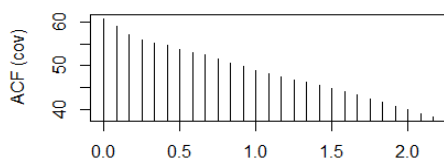
De acuerdo la metodología Box-Jenkins, hicimos un análisis visual para determinar si las series son estacionarias. Los supuestos de estacionariedad débil se cumplen si la media y varianza son constantes en el tiempo y si el valor de la covarianza entre dos periodos depende sólo de la distancia o rezago entre los tiempos. Como vemos en los gráficos anteriores, las medias no parecen ser constantes a lo largo de la serie lo que ya nos sugiere que no son estacionarias. Por otro lado vemos un aumento significativo de los precios a partir del año 2008, seguido de una caída y luego una tendencia alcista hasta el año 2011 para luego descender hasta el año 2016.

Graficamos la FAS (Función de Autocovarianzas), FAC (Función de Autocorrelación) y FACP (Función de Autocorrelación Parcial) de cada serie temporal:

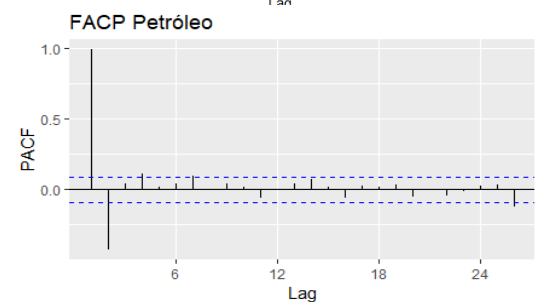
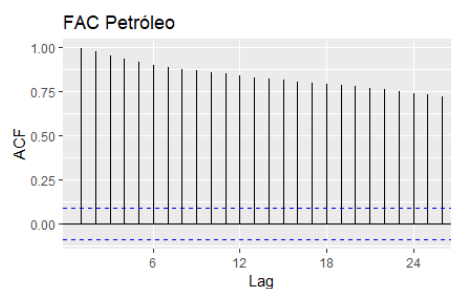
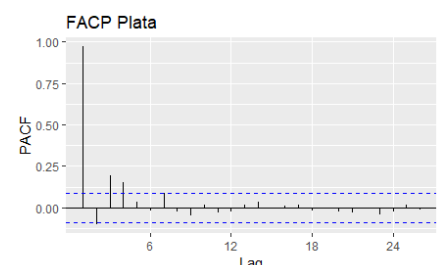
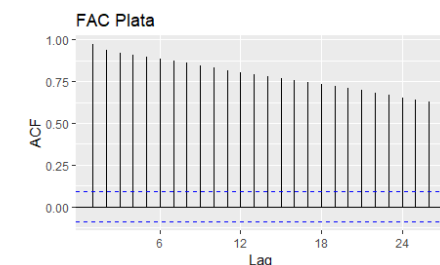
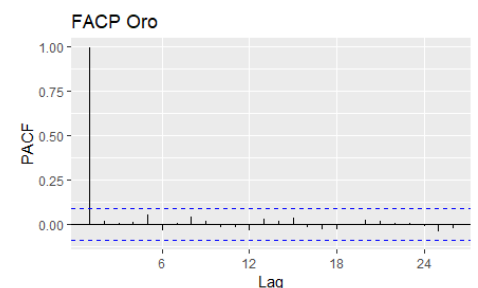
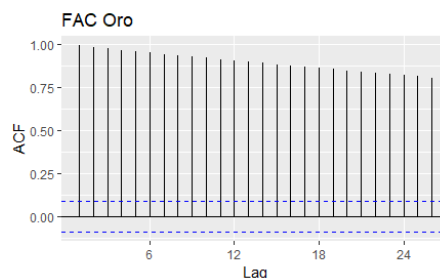
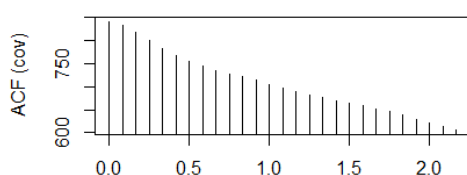
➤ Variable: Oro  
oro



➤ Variable: Plata  
plata



➤ Variable: Petróleo  
petroleo



A partir de la FAC de las distintas series, inferimos que las series no son estacionarias dado que decrecen de manera lineal. Asimismo, vemos en las FACP que los valores actuales están explicados por componentes autorregresivos. En el caso del oro, el 1er lag es el más significativo; para la plata, del 1ro al 4to lag y finalmente para el petróleo, el 1ro, 2do y 4to lag.

### Análisis de la estacionariedad:

#### ❖ Ljung-Box:

Para analizar la estacionariedad de las series verificamos si todos los coeficientes de autocorrelación son iguales a cero. Para ello, realizamos un test de Ljung-Box para los distintos lags. Si rechazamos la hipótesis nula significa que hay coeficientes de autocorrelación distintos a cero.

Dado que los p-value son menores a 0.05, rechazamos la hipótesis nula, por lo que podemos afirmar que hay autocorrelación. Esto implica no estacionariedad y dependencia con respecto a períodos anteriores.

#### ❖ Dickey-Fuller:

Realizamos un test de raíz unitaria de Dickey-Fuller en sus variantes trend (tendencia y constante), drift (constante) y none (sin constante ni tendencia). Para seleccionar el lag más significativo utilizamos el criterio de información de Akaike (AIC).

#### ➤ Variable: Oro

none	1pct	5pct	10pct	Valor Estadístico	
tau1	-2.58	-1.95	-1.62	1.7506	
drift	1pct	5pct	10pct	Valor Estadístico	
tau2	-3.44	-2.87	-2.57	0.9967	
phi1	6.47	4.61	3.79	1.5299	Coef. no significativo
trend	1pct	5pct	10pct	Valor Estadístico	
tau3	-3.98	-3.42	-3.13	-1.4198	
phi2	6.15	4.71	4.05	3.4337	Coef. no significativo
phi3	8.34	6.30	5.36	4.1042	Coef. no significativo

#### ➤ Variable: Plata

none	1pct	5pct	10pct	Valor Estadístico	
tau1	-2.58	-1.95	-1.62	-2.1014	
drift	1pct	5pct	10pct	Valor Estadístico	
tau2	-3.44	-2.87	-2.57	-3.25	
phi1	6.47	4.61	3.79	5.3242	Coef. no significativo
trend	1pct	5pct	10pct	Valor Estadístico	
tau3	-3.98	-3.42	-3.13	-4.9914	
phi2	6.15	4.71	4.05	9.0903	Coef. no significativo
phi3	8.34	6.30	5.36	13.591	Coef. no significativo

#### ➤ Variable: Petróleo

none	1pct	5pct	10pct	Valor Estadístico	
tau1	-2.58	-1.95	-1.62	-1.323	
drift	1pct	5pct	10pct	Valor Estadístico	
tau2	-3.44	-2.87	-2.57	-2.4413	
phi1	6.47	4.61	3.79	2.9821	Coef. no significativo
trend	1pct	5pct	10pct	Valor Estadístico	
tau3	-3.98	-3.42	-3.13	-3.2198	
phi2	6.15	4.71	4.05	3.4589	Coef. no significativo
phi3	8.34	6.30	5.36	5.1861	Coef. no significativo

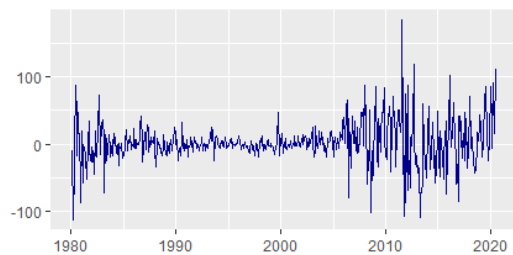
Para el caso del oro y el petróleo, en las tres variantes del test tau1, tau2 y tau3 son mayores a los valores críticos del estadístico para los distintos intervalos de confianza (99%, 95% y 90%), por lo que no se puede rechazar la hipótesis nula (la serie no es estacionaria). Además, los coeficientes  $\phi_1$ ,  $\phi_2$  y  $\phi_3$  no son significativos.

Para el caso de la plata, según el intervalo de confianza se rechaza la hipótesis nula. Los coeficientes  $\phi_1$ ,  $\phi_2$  y  $\phi_3$  no son significativos.

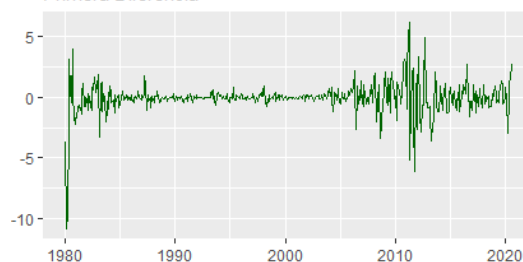
Por todo lo expuesto anteriormente tenemos evidencia suficiente para asumir la **no estacionariedad** de las series. Para hacer estacionarias nuestras series, las diferenciaremos una vez ( $d=1$ ).

**Graficamos las primeras diferencias de las tres series temporales:**

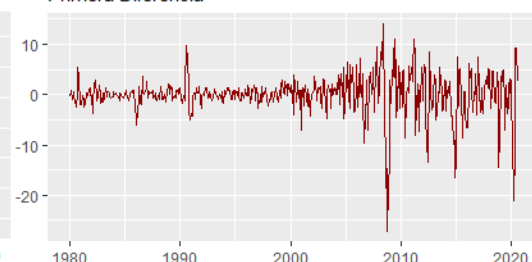
Precio Oro  
Primera Diferencia



Precio Plata  
Primera Diferencia



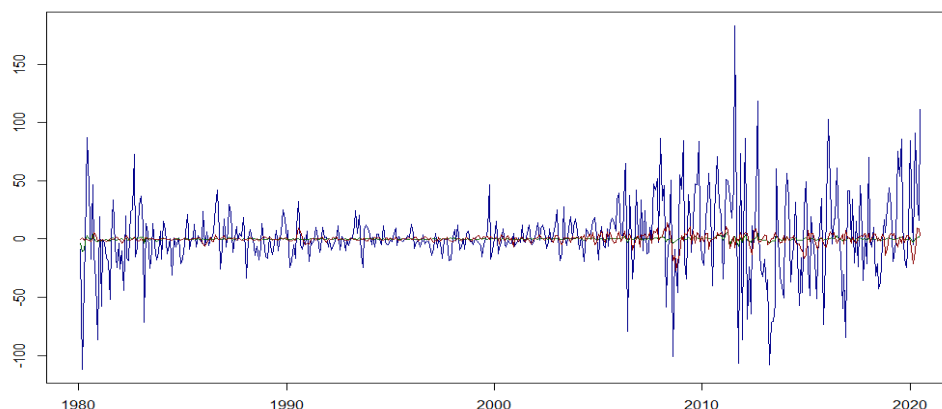
Precio Petróleo  
Primera Diferencia



Del análisis visual, ahora sí parecen cumplirse las condiciones de estacionariedad. A partir del año 2008, parece haber más varianza en las series diferenciadas.

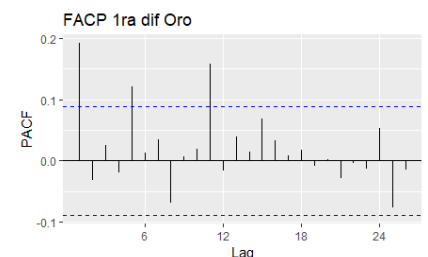
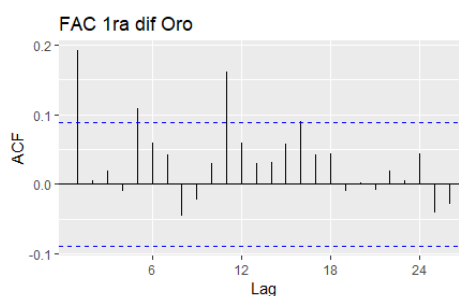
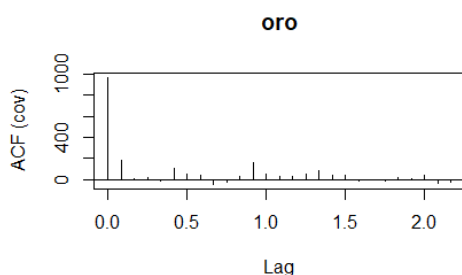
**Graficamos las tres series diferenciadas juntas en la misma escala:**

Evolución del Precio Diferenciado de los Commodities

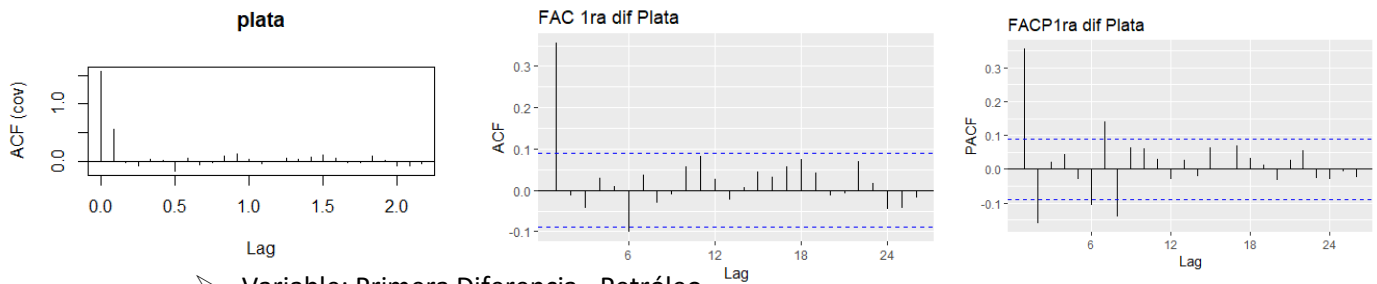


**Graficamos la FAS (Función de Autocovarianzas), FAC (Función de Autocorrelación) y FACP (Función de Autocorrelación Parcial) de cada serie temporal diferenciada:**

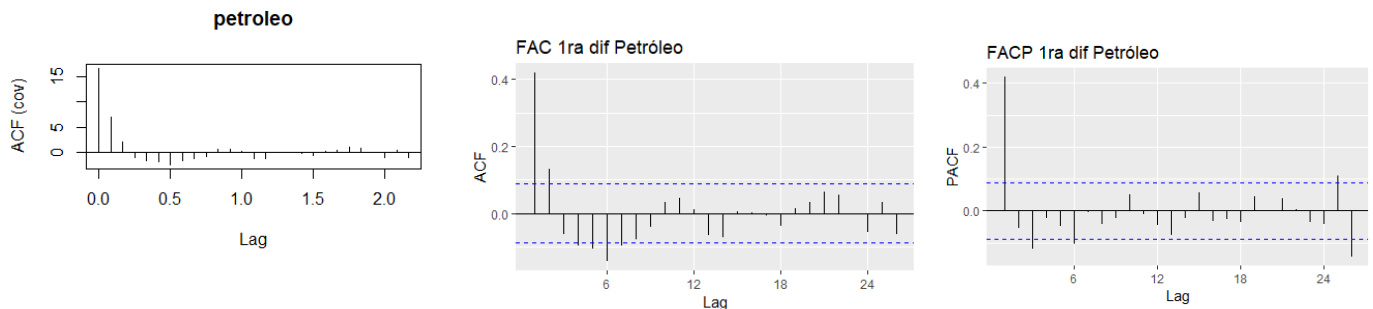
➤ Variable: Primera Diferencia - Oro



➤ Variable: Primera Diferencia – Plata



➤ Variable: Primera Diferencia - Petróleo



De los gráficos de FAS y FAC vemos que decrecen de forma exponencial lo que indica que podemos llegar a estar en presencia de un proceso estacionario. Por otro lado en el caso de las FACP del oro y la plata, parecen haber algunos lags que son significativos, lo que puede significar que aún existe alguna componente autorregresiva importante.

**Análisis de la estacionariedad:**

❖ Ljung-Box:

Al igual que con las series originales, verificamos si existe autocorrelación mediante un test de Ljung-Box para los distintos lags. El p-value se ubicó por encima del cero, pero sigue siendo menor a 0.05. Esto implicaría no estacionariedad, lo cual está en contra del resultado que esperábamos. Pasaremos a confirmarlo mediante un test de Dickey-Fuller

❖ Dickey-Fuller:

Realizamos un test de raíz unitaria de Dickey-Fuller en sus variantes trend (tendencia y constante), drift (constante) y none (sin constante ni tendencia). Para seleccionar el lag más significativo utilizamos el criterio de información de Akaike (AIC).

➤ Variable: Primera Diferencia - Oro

none	1pct	5pct	10pct	Valor Estadístico
tau1	-2.58	-1.95	-1.62	-14.2874
drift	1pct	5pct	10pct	Valor Estadístico
tau2	-3.44	-2.87	-2.57	-14.406
phi1	6.47	4.61	3.79	103.8277
trend	1pct	5pct	10pct	Valor Estadístico
tau3	-3.98	-3.42	-3.13	-14.6564
phi2	6.15	4.71	4.05	71.6442
phi3	8.34	6.30	5.36	107.4045

Coef. no significativo

Coef. no significativo

Coef. no significativo



➤ Variable: Primera Diferencia – Plata

none	1pct	5pct	10pct	Valor Estadístico
tau1	-2.58	-1.95	-1.62	-16.4937

drift	1pct	5pct	10pct	Valor Estadístico
tau2	-3.44	-2.87	-2.57	-16.4646
phi1	6.47	4.61	3.79	135.7413

Coef. no significativo

trend	1pct	5pct	10pct	Valor Estadístico
tau3	-3.98	-3.42	-3.13	-16.5087
phi2	6.15	4.71	4.05	91.115
phi3	8.34	6.30	5.36	136.4726

Coef. no significativo

Coef. no significativo

➤ Variable: Primera Diferencia - Petróleo

none	1pct	5pct	10pct	Valor Estadístico
tau1	-2.58	-1.95	-1.62	-12.4373

drift	1pct	5pct	10pct	Valor Estadístico
tau2	-3.44	-2.87	-2.57	-12.4244
phi1	6.47	4.61	3.79	77.1842

Coef. no significativo

trend	1pct	5pct	10pct	Valor Estadístico
tau3	-3.98	-3.42	-3.13	-12.4111
phi2	6.15	4.71	4.05	51.351
phi3	8.34	6.30	5.36	77.0252

Coef. no significativo

Coef. no significativo

En todos los casos tau1, tau2 y tau3 son menores a los valores críticos del estadístico para los distintos intervalos de confianza (99%, 95% y 90%), por lo que se puede rechazar la hipótesis nula de que las series no son estacionarias.

### Modelo VAR (vectores autorregresivos):

Planteamos un modelo VAR con las tres series diferenciadas. A partir de distintos criterios de información (Akaike, Hannan Quinn, Schwarz y error de predicción final de Akaike) definimos la cantidad de lags convenientes en 10 lags.

### Causalidad de Granger:

Se dice que una variable X causa a Z en el sentido de Granger cuando en presencia del pasado de Z, el pasado de X ayuda a explicar Z. Analizamos la causalidad de Granger para el caso del modelo con tres variables (entre cada variable y par de variables). Respecto a la cantidad de lags a elegir, con muy pocos rezagos (lags), podemos encontrar autocorrelación en los residuos, lo que nos da un test sesgado. Con demasiados, podríamos rechazar incorrectamente la hipótesis nula debido a una correlación espuria. Se elige el lag en función del criterio de información AIC (Akaike) y se prueba la causalidad en ambos sentidos (bidireccional). La hipótesis nula del test es que no existe causalidad en el sentido de Granger.

❖ Causalidad dos variables:

p-value	oro	plata	petróleo
oro		0.00001033	0.001684
plata	0.003468		0.0003095
petróleo	0.008025	0.3542	

Nota: Se lee "causalidad" de la columna hacia la "fila"

Un p-value menor a 0.05 denota la presencia de causalidad de Granger. Todos los p-value nos dieron menor a 0.05, excepto la causalidad de la plata sobre el petróleo.

❖ Causalidad tres variables:

p-value	oro, plata	plata , petróleo	oro, petróleo
oro		0.0002109	
plata			0.00001717
petróleo	0.001509		

p-value	oro	plata	petróleo
oro, plata			0.002061
plata , petróleo	0.00000004053		
oro, petróleo		0.000001313	

Nota: Se lee "causalidad" de la columna hacia la "fila"

Un p-value menor a 0.05 denota la presencia de causalidad de Granger. Todos los p-value nos dieron menor a 0.05.

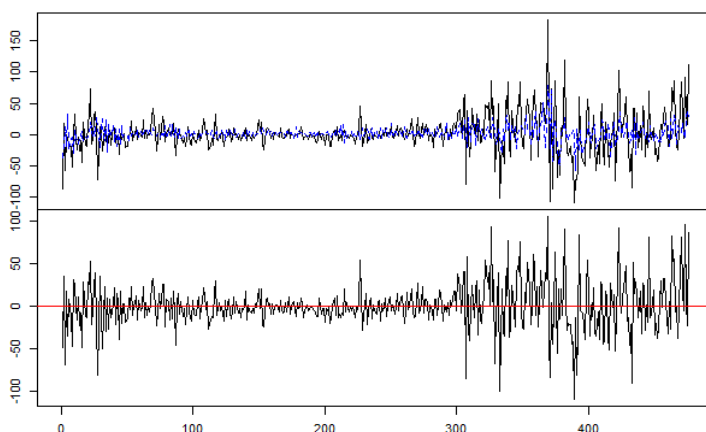
Analizamos la matriz de correlación de los residuos:

	oro	plata	petróleo
oro	1	0.7169	0.08645
plata	0.71687	1	0.27748
petróleo	0.08645	0.2775	1

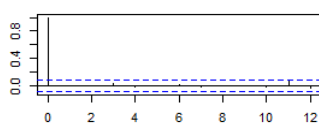
Como vemos, ningún valor es cero. Existe una correlación significativa entre los residuos de la plata y el oro. Una menor correlación entre los residuos del petróleo y la plata y, una muy baja entre los residuos del oro y del petróleo.

Graficamos los residuos y su correspondiente FAC, FACP:

Diagram of fit and residuals for oro



ACF Residuals



PACF Residuals

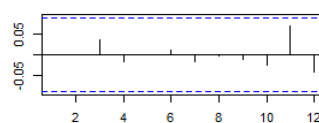
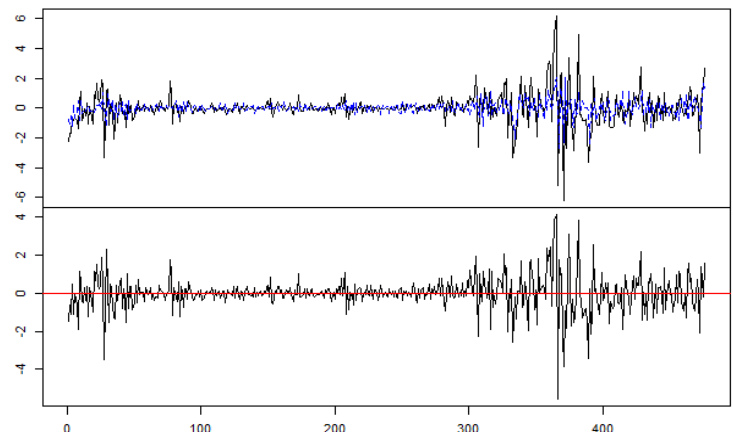
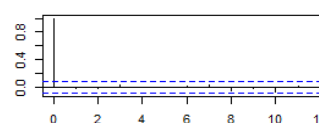


Diagram of fit and residuals for plata



ACF Residuals



PACF Residuals

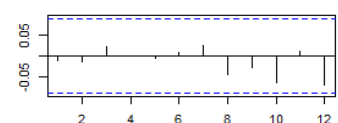
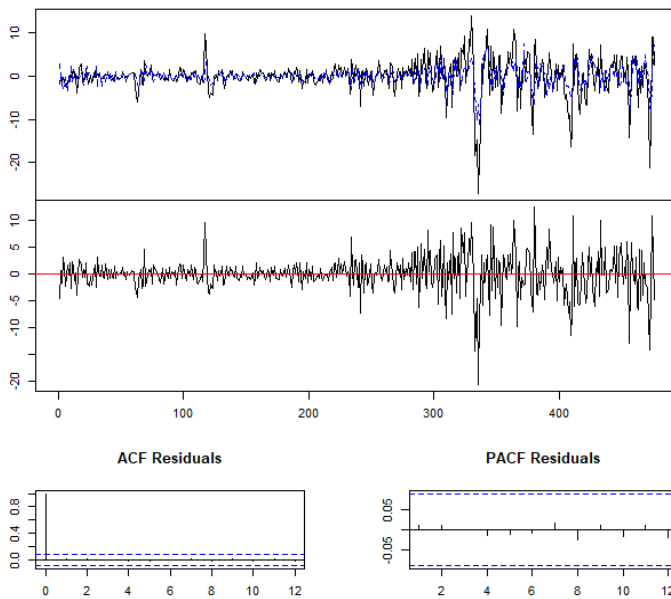


Diagram of fit and residuals for petroleo



Según los gráficos, los residuos parecen comportarse como un ruido blanco. De las gráficas de FAC, se observa que se mantienen dentro del intervalo de confianza y a partir de las FACP no parece haber presencia del algún componente autorregresivo significativo.

#### ❖ Autocorrelación de los residuos:

Verificamos la autocorrelación de los residuos del modelo mediante un test de Portmanteau. Siendo el p-value menor a 0.05, rechazamos la hipótesis nula de incorrelación (los coeficiente de autocorrelación son iguales a cero) por lo que no estaríamos cumpliendo el supuesto de incorrelación.

#### ❖ Normalidad de los residuos:

Analizamos la normalidad de los residuos mediante un test de Jarque-Vera. Siendo el p-value menor a 0.05, rechazo el supuesto de normalidad. En este caso no se cumple con el supuesto de normalidad de los residuos para el modelo elegido.

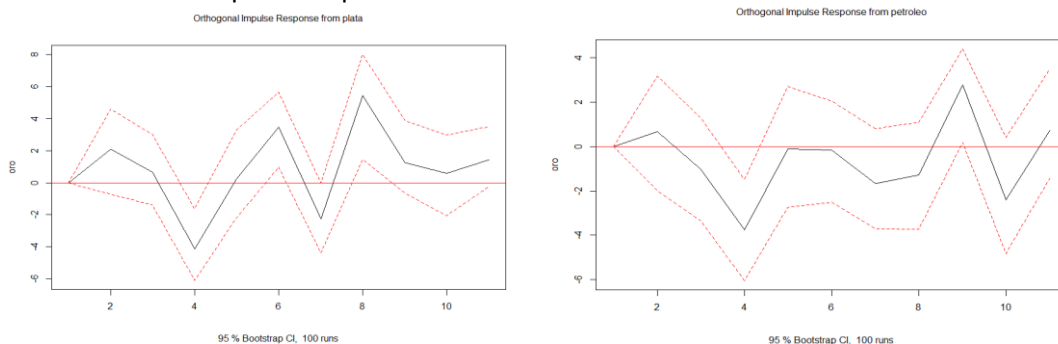
#### ❖ Homocedasticidad de los residuos:

Mediante un test Arch, comprobamos que los residuos se comportan de manera heterocedástica. Por lo que tampoco cumple con la condición de homocedasticidad de los residuos.

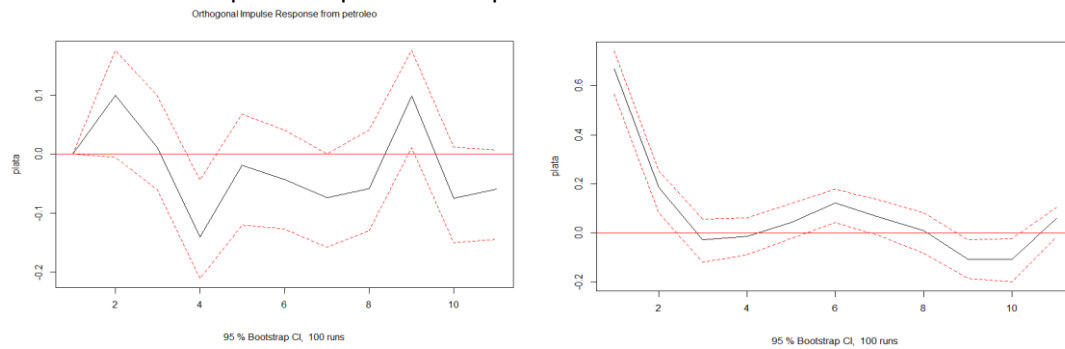
### **Función impulso-respuesta:**

Mediante las siguientes gráficas, vemos como reaccionan cada una de las variables ante una perturbación de las otras:

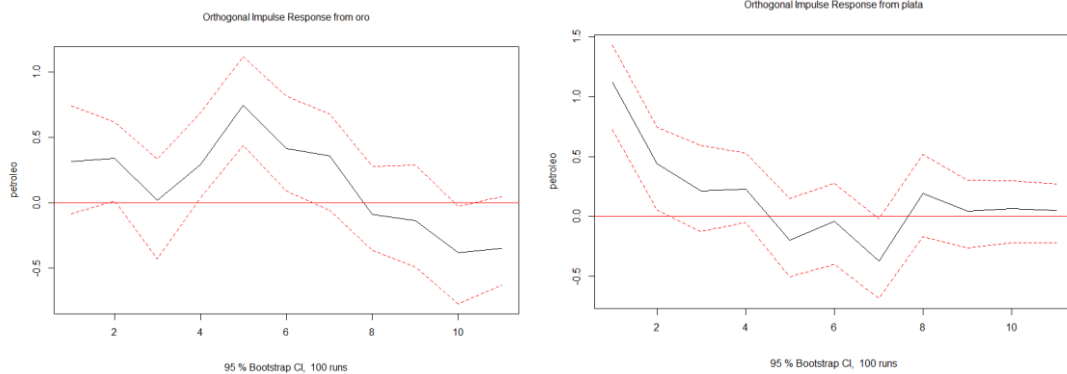
#### ➤ Función impulso-respuesta del oro:



➤ Función impulso-respuesta de la plata:

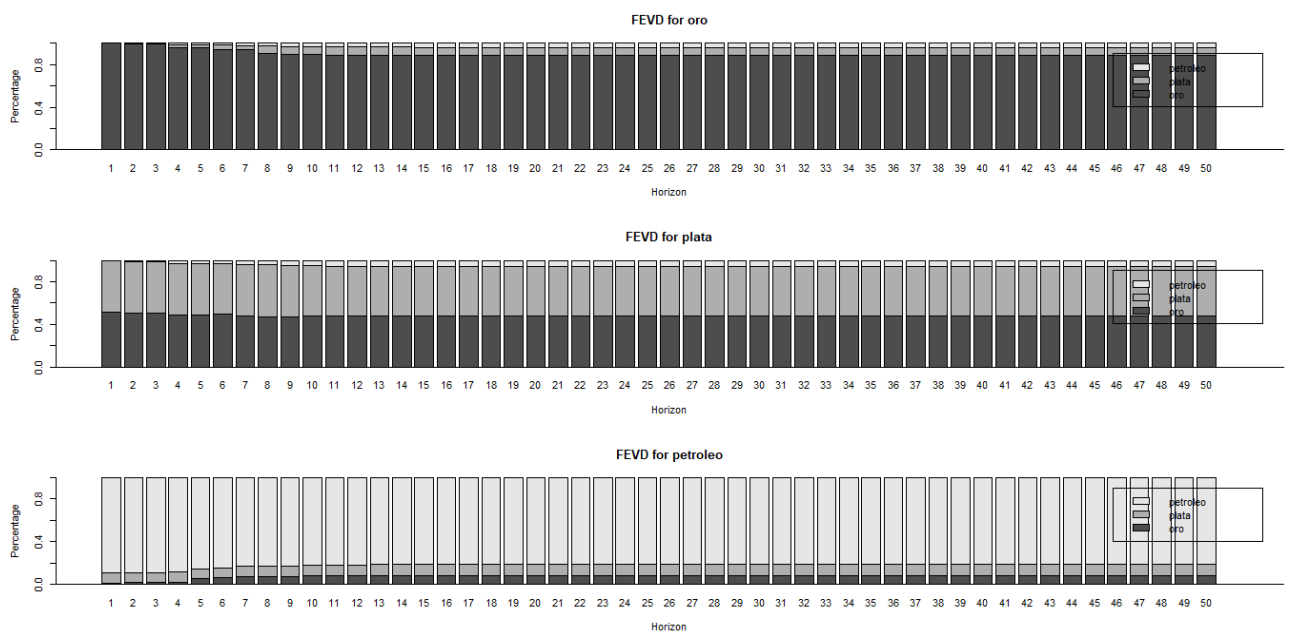


➤ Función impulso-respuesta del petróleo:



Como vemos de las gráficas, ante una perturbación de una variable, no se estabilizan rápidamente en el tiempo. Además, esto lo observamos en que las bandas de confianza no tienden a achicarse con el tiempo.

Descomposición de las varianzas:



Nota: orden de gris claro a gris oscuro → petróleo, plata y oro

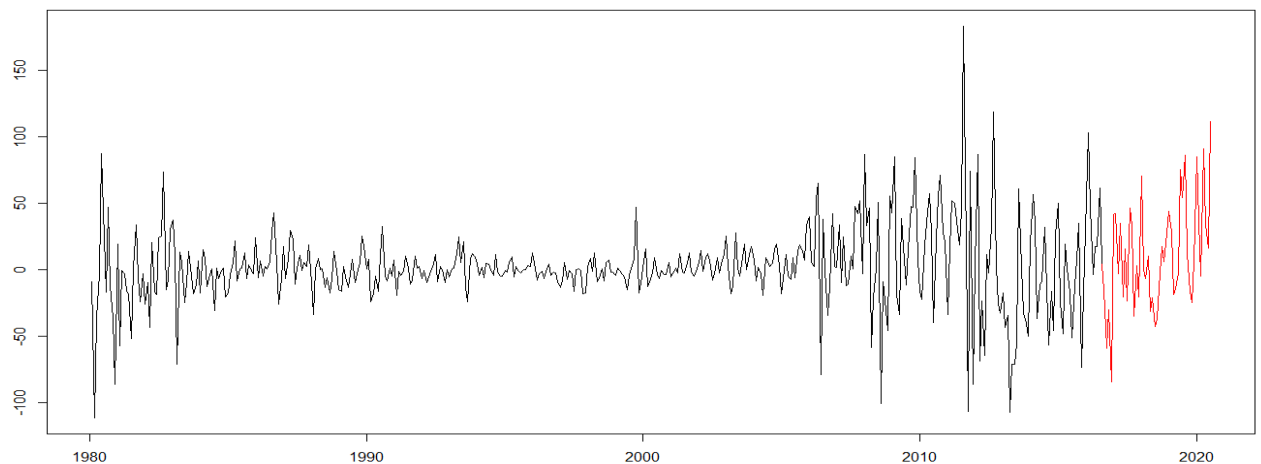
El gráfico me indica el porcentaje explicativo de cada variable sobre la variabilidad. Siendo el porcentaje aproximado para cada caso el siguiente:

- Oro Oro: 88% Plata: 8% Petróleo: 4%
- Plata Oro: 47% Plata: 46% Petróleo: 7%
- Petróleo Oro: 8% Plata: 10% Petróleo: 82%

### Predicciones:

A partir de la **serie diferenciada del oro**, realizaremos predicciones para sus últimos cuatro años (aproximadamente un 10% del dataset) mediante las siguientes técnicas: modelo VAR, modelo ingenuo con estacionalidad, modelo autorregresivo de redes neuronales (NNAR) y modelo de redes neuronales Long Short-Term Memory (LSTM). Luego compararemos la precisión de todos los métodos.

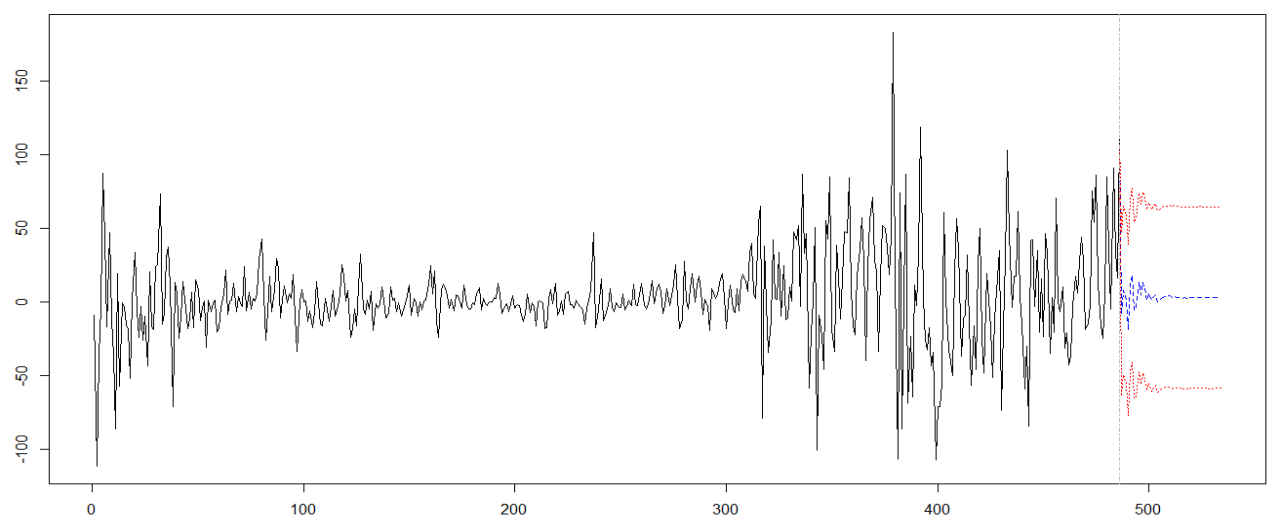
**Graficamos el conjunto de entrenamiento (en negro) y los valores a ser predichos (en rojo):**



#### ➤ Modelo VAR:

**Graficamos las predicciones del modelo (azul):**

Forecast of series oro



Vemos que las predicciones tienen una cierta variabilidad al principio y luego se estabilizan con el tiempo en torno al cero.

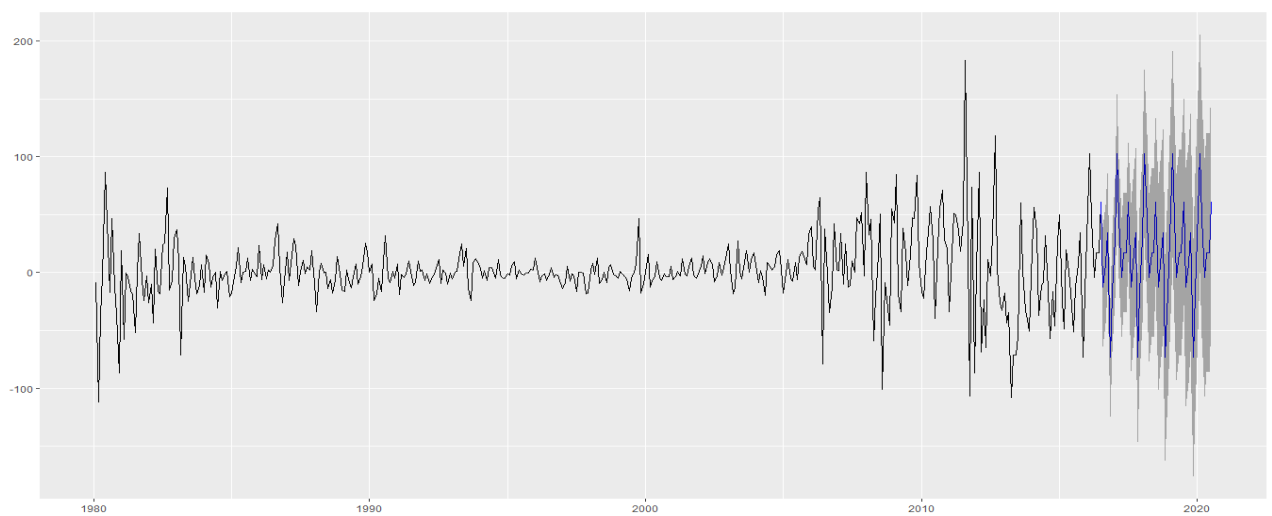
### Precisión Modelo VAR:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training	0,29165	34,57367	21,92402	100,5378	304,9059	1,161152	NA	NA
Test	7,653966	41,52969	31,55671	116,9219	118,7989	1,671324	0,215189	1,172256

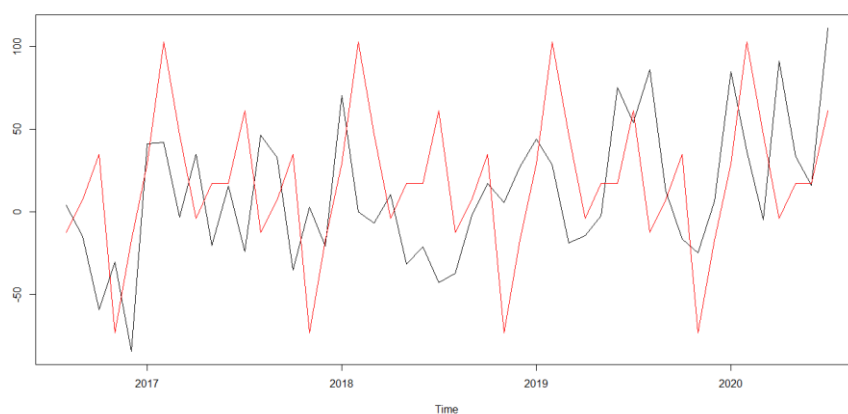
#### ➤ Modelo Ingenuo con Estacionalidad:

En el método ingenuo, las predicciones que se producen son iguales al último valor observado. Si se cree que la serie de tiempo pueda tener estacionalidad, el enfoque ingenuo con estacionalidad puede ser el más apropiado, donde las predicciones son iguales al valor de la temporada pasada

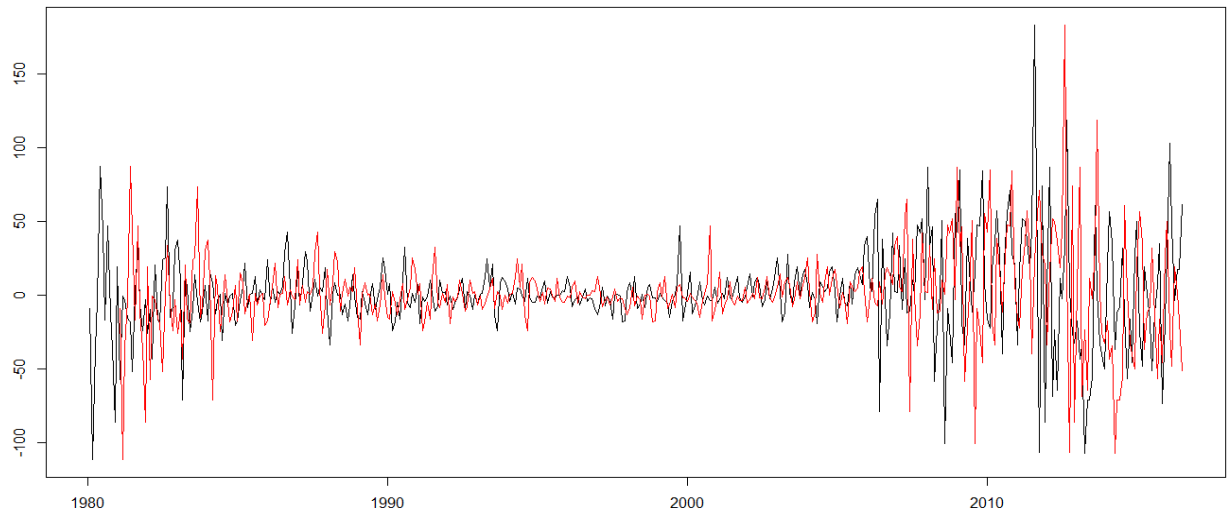
#### Graficamos las predicciones del modelo (azul):



#### Graficamos el conjunto de testeo y las predicciones (en rojo):

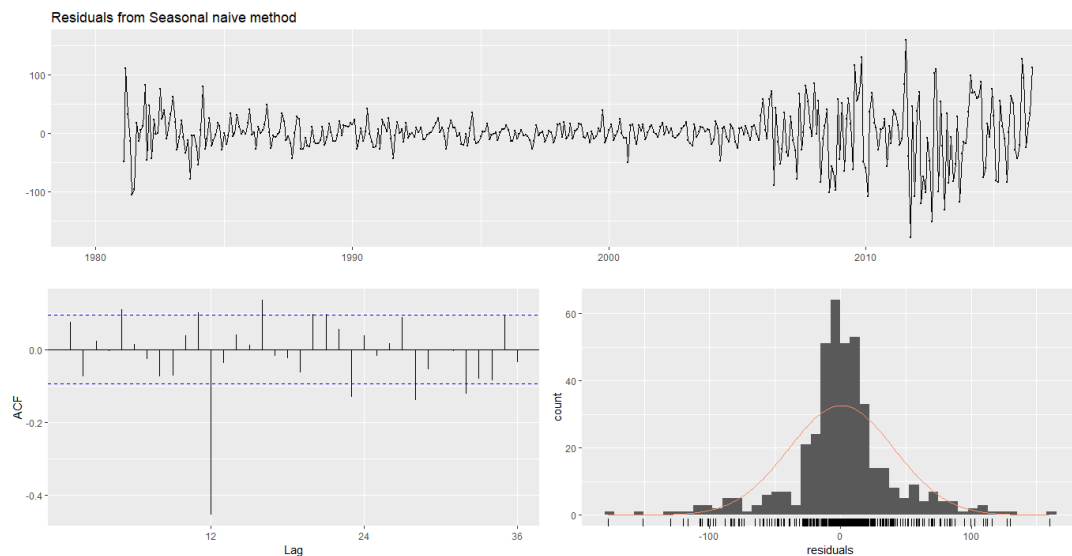


Graficamos el conjunto de entrenamiento y las predicciones (en rojo):



Supuestos del modelo:

Comprobamos la normalidad de los residuos con distintos tests (JB, AD SW). Con los tres test rechacé el supuesto de normalidad, no cumpliendo con el supuesto del modelo. De la FAC vemos una componente importante para el lag 12.



Planteamos un test de Ljung-Box para verificar la incorrelación de los residuos. Rechazamos dicho test, lo que indica correlación y falta de independencia, no cumpliendo con otro de los supuestos teóricos del modelo.

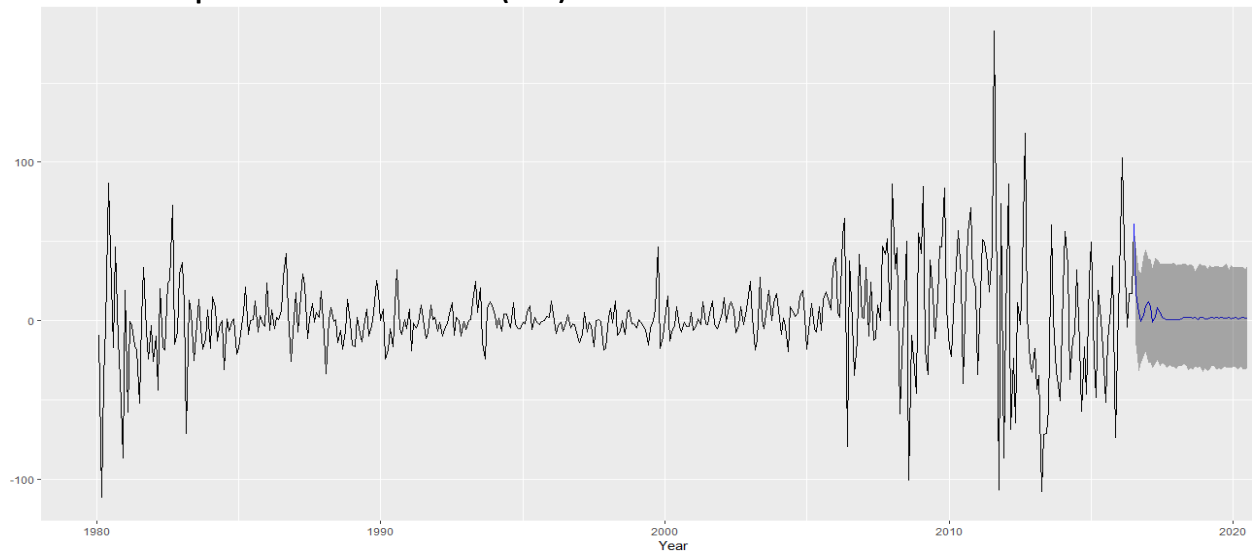
Precisión Modelo Ingenuo con Estacionalidad:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training	0,761199	40,18983	26,44731	198,3647	553,3026	1	0,07544	NA
Test	-6,73271	53,54939	44,87709	1770,594	1814,9	1,696849	-0,07574	5,868574

➤ **Modelo Autorregresivo de Redes Neuronales (NNAR)**

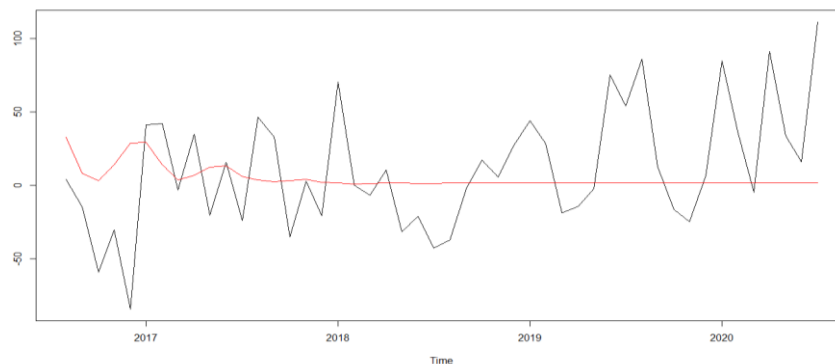
Definimos una red neuronal con una capa oculta y diez nodos de inputs. Elegimos dicho valor para el input, debido a que fue la cantidad de lags recomendado en el modelo VAR.

**Graficamos las predicciones del modelo (azul):**

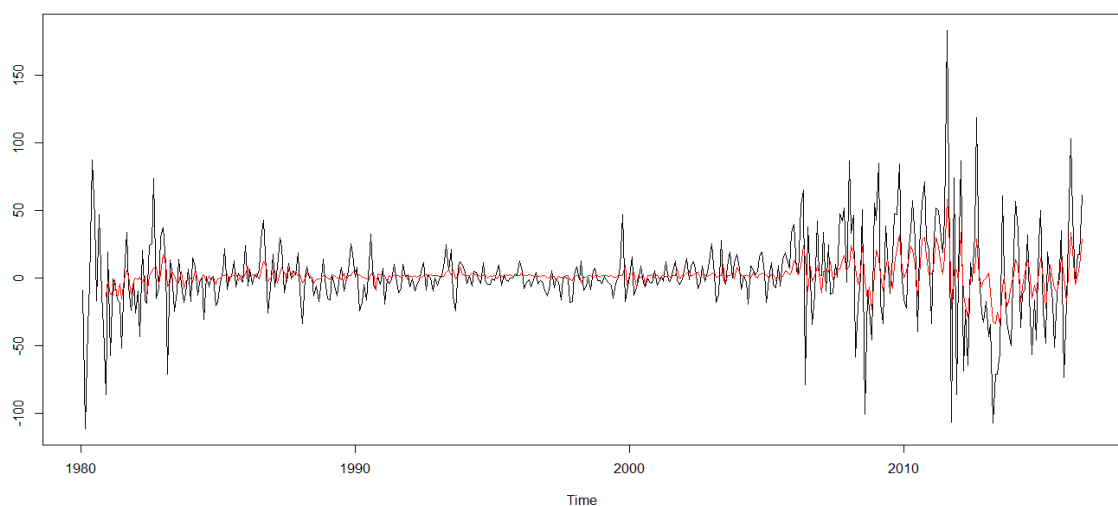


Al igual que el modelo VAR, existe una variabilidad que luego se estabiliza con el tiempo en torno al cero.

**Graficamos el conjunto de testeo y las predicciones (en rojo):**



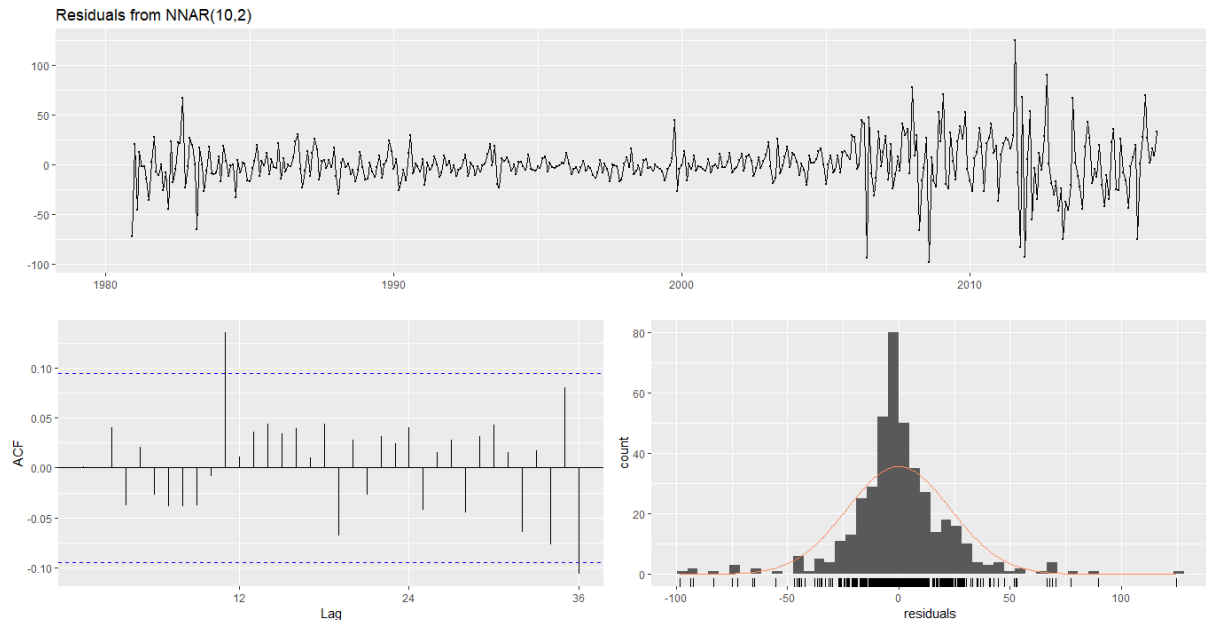
**Graficamos el conjunto de entrenamiento y las predicciones (en rojo):**





### Supuestos del modelo:

Comprobamos la normalidad de los residuos con distintos tests (JB, AD SW). Con los tres test rechazé el supuesto de normalidad, no cumpliendo con el supuesto del modelo. De la gráfica de FAC vemos una componente importante para el lag 11.



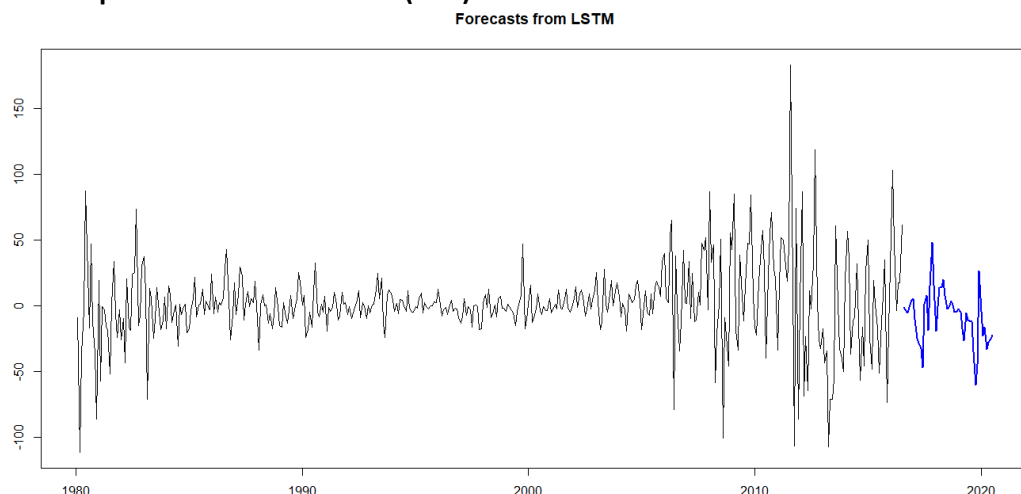
### Precisión Modelo Autorregresivo de Redes Neuronales (NNAR)

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training	-0,01473	23,23464	15,16066	97,80391	186,7499	0,57324	0,000887	NA
Test	5,898569	43,00248	32,79226	95,06839	128,5211	1,239909	0,292903	1,087577

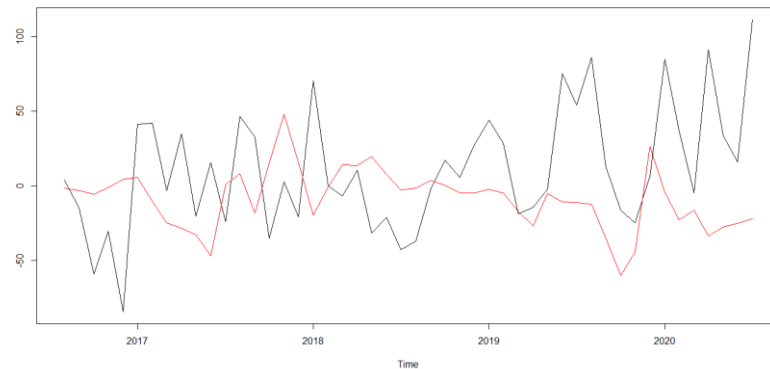
### ➤ Modelo de redes neuronales Long Short-Term Memory (LSTM):

Las redes neuronales Long Short-Term Memory (LSTM) son una extensión de las redes neuronales recurrentes, que básicamente amplían su memoria para aprender de sucesos del pasado. Las LSTM permiten a las redes neuronales recordar sus inputs durante un largo período de tiempo dado que las LSTM contienen su información en la memoria

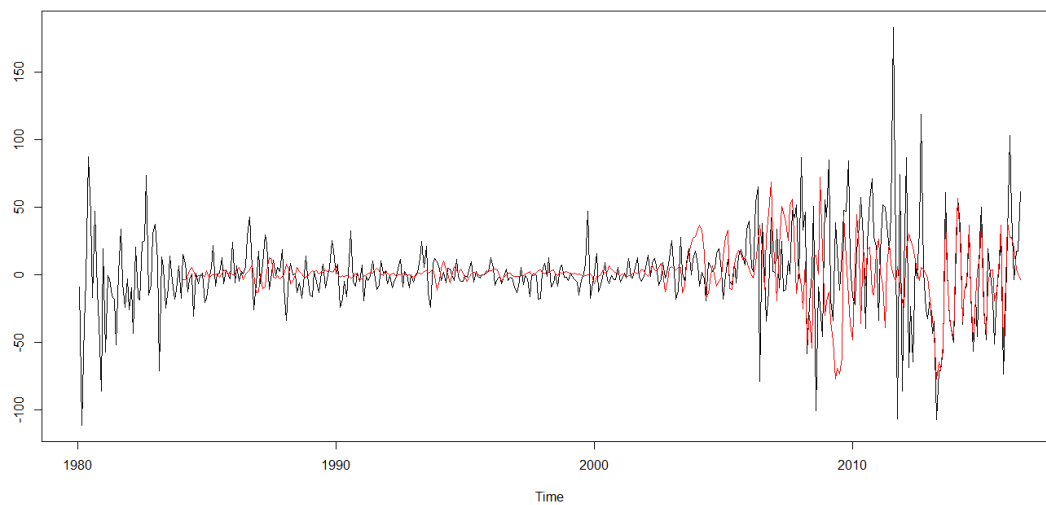
### Graficamos las predicciones del modelo (azul):



Graficamos el conjunto de testeo y las predicciones (en rojo):

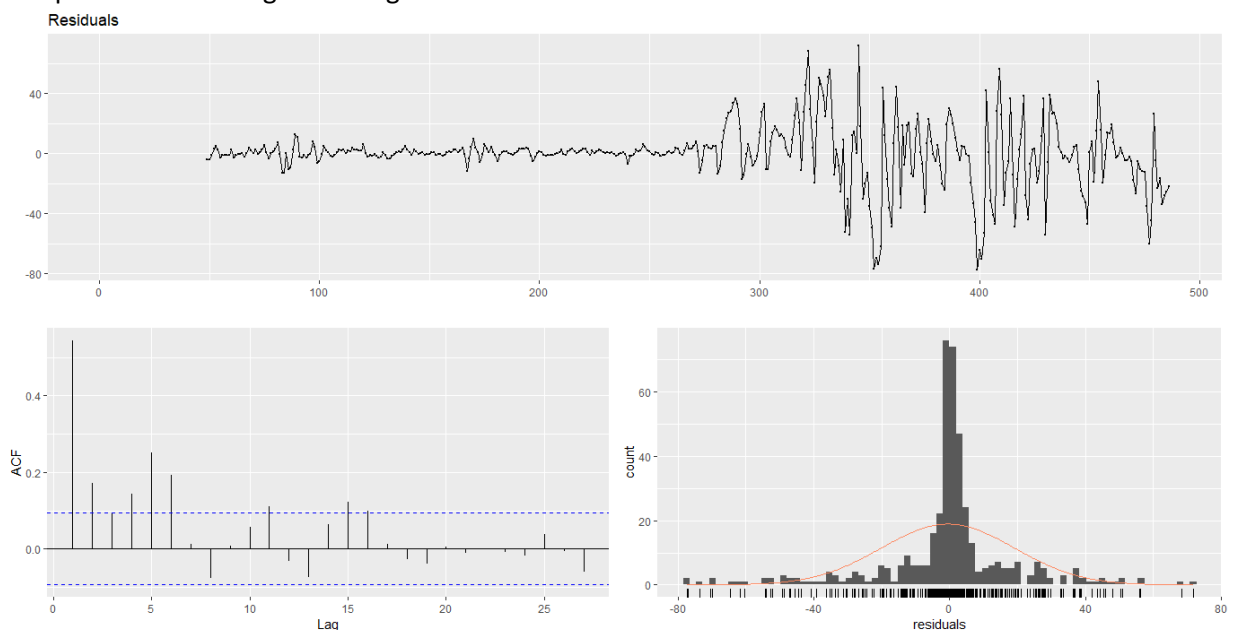


Graficamos el conjunto de entrenamiento y las predicciones (en rojo):



Supuestos del modelo:

Comprobamos la normalidad de los residuos con distintos tests (JB, AD SW). Con los tres test rechacé el supuesto de normalidad, no cumpliendo con el supuesto del modelo. De la FAC vemos varios componentes autorregresivos significativos.



#### Precisión Modelo LSTM:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training	16,52525	30,27871	17,90782	-93,5236	14638,7	0,791191		
Test	19,21584	52,33828	42,04549	29,40268	3312,27	1,065383		

#### • IV) Conclusiones:

En el presente trabajo hemos analizado tres series temporales correspondiente al precio de distintos commodities (oro, plata y petróleo). A priori, conocíamos la relación que pudiera existir entre las mismas, las cuales comprobamos al plantear un modelo del tipo VAR que tiene en cuenta como el pasado de las variables influyen entre ellas. Mediante dicho modelo constatamos la causalidad en el sentido de Granger entre las variables y además examinamos como el shock en una variable repercute en las otras con la función impulso – respuesta. Luego realizamos predicciones sobre la primera diferencia del oro. Un punto a destacar aquí es que al trabajar con predicciones de la primera diferencia perdemos información sobre la relatividad de dicha variación. No es lo mismo tener una variación absoluta en 10 unidades, partiendo de 1000 que partiendo de 10. Esto igualmente puede sortearse, realizando la anti transformación de la primera diferencia en la serie original.

Posteriormente, comparamos dicha predicción con otros modelos como ser modelo ingenuo con estacionalidad y modelo autorregresivo de redes neuronales (NNAR). La particularidad de estos métodos es que uno pierde poder explicativo del modelo, comparado con el modelo VAR, el cual se basa en distintos supuestos estadísticos como ser la relación entre las variables.

Las medidas de precisión obtenidas en el conjunto de testeo para todos los modelos fueron las siguientes:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
VAR	7,653966	41,52969	31,55671	116,9219	118,7989	1,671324	0,215189	1,172256
NAIVE	-6,73271	53,54939	44,87709	1770,594	1814,9	1,696849	-0,07574	5,868574
NNAR	5,898569	43,00248	32,79226	95,06839	128,5211	1,239909	0,292903	1,087577
LSTM	19,21584	52,33828	42,04549	29,40268	3312,27	1,065383		

De acuerdo al cuadro anterior, el mejor modelo sería el VAR. No obstante según el gráfico con las predicciones de dicho método, no parece ajustarse de forma satisfactoria a los datos reales del conjunto de testeo. Además, hemos detectado que no se cumplen ciertos supuestos teóricos como ser la normalidad, incorrelación y homocedasticidad de los residuos.

En cuanto al comportamiento de las series, cabe destacar que a partir de la mitad de la década de 1980 las series parecen estabilizarse. Luego en el año 2008 comienzan a aumentar su variabilidad. Dado que realizamos predicciones para los últimos 12 meses, quizás obtendríamos una mejor precisión si tomáramos un recorte más actual de la serie y no la totalidad de la historia desde 1980.

#### • V) Bibliografía:

Peña, D. (2005). Análisis de Series Temporales. Alianza Editorial.  
Uriel, E. (1984). Series Temporales. Paraninfo

#### Links:

<https://stats.stackexchange.com/questions/131261/granger-causality-interpretation-using-r/132527>  
<https://stackoverflow.com/questions/49744962/multiple-variable-names-in-var-causality-function>

<https://boostedml.com/2019/11/testing-predictive-value-in-time-series-granger-causality-in-r.html>  
<https://stats.stackexchange.com/questions/107954/lag-order-for-granger-causality-test>  
<https://www.r-bloggers.com/granger-causality-test/>  
<https://www.economodel.com/time-series-analysis>  
<https://kevin-kotze.gitlab.io/tsm/ts-10-tut/>  
<http://repositorio-digital.cide.edu/bitstream/handle/11651/673/151476.pdf?sequence=1>  
<https://estima.com/ratshelp/index.html?causalitytesting.html>  
<https://kevinkotze.github.io/ts-6-tut/>  
<https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python/>  
<https://www.mathworks.com/help/econ/varm.gctest.html>  
<https://stats.stackexchange.com/questions/194453/interpreting-accuracy-results-for-an-arima-model-fit>  
[https://guiasjuridicas.wolterskluwer.es/Content/Documento.aspx?params=H4sIAAAAAAEAMtMSbF1jTAAASMTczNLtbLUouLM\\_DxblwMDSONDQ3OQQGZapUt-ckhlQaptWmJOcSoAFtLiADUAAAA=WKE#I4](https://guiasjuridicas.wolterskluwer.es/Content/Documento.aspx?params=H4sIAAAAAAEAMtMSbF1jTAAASMTczNLtbLUouLM_DxblwMDSONDQ3OQQGZapUt-ckhlQaptWmJOcSoAFtLiADUAAAA=WKE#I4)  
[https://www.upo.es/econ/hidalgo/wp-content/uploads/2014/09/tema\\_var.pdf](https://www.upo.es/econ/hidalgo/wp-content/uploads/2014/09/tema_var.pdf)  
<https://bookdown.org/amanas/traficomadrid/m%C3%A9todos-basados-en-deep-learning.html#redes-autorregresivas>  
<https://bookdown.org/amanas/traficomadrid/m%C3%A9todos-basados-en-deep-learning.html#lstm-univariado>  
<https://cran.r-project.org/web/packages/zFactor/vignettes/statistics.html>

## • VI) Anexo:

Se junta el código en R y el Dataset en el siguiente enlace:

<https://github.com/fero1987/Consigna-Final>

### Anexo – Medidas de precisión:

Accuracy Measuring Methods	Formulation
MPE	$((\text{Actual} - \text{Forecast}) / \text{Actual}) \times 100$ $MPE = \frac{100}{N} \sum_{i=1}^N \frac{A_i - F_i}{A_i} \quad (3)$
MAPE	$MAPE = (\text{Absolute Value}(\text{Actual} - \text{Forecast}) / \text{Actual}) \times 100$ $MAPE = \frac{1}{n} \sum_{t=1}^n \frac{ e_t }{ Y_t } \quad (4)$
RMSE	$RMSE = \text{Sqrt}(\text{Avg}(\text{Power}(\text{Actual} - \text{Forecast})))$ $RMSE = \sqrt{MSE} \quad (5)$
MAE	$MAE = \text{Avg}(\text{Abs}(\text{Actual} - \text{Forecast}))$ $MAE = \frac{1}{n} \sum_{t=1}^n  e_t  \quad (6)$
ME	$ME = \frac{1}{N} \sum_{i=1}^N (A_i - F_i) \quad (7)$
MSE	$MSE = \frac{1}{n} \sum_{t=1}^n e_t^2 \quad (8)$

Error (e) = Actual – Forecast or  $e_t = A_t - F_t$

Para mayor información y detalle, consultar la siguiente fuente:

<https://repository.cesa.edu.co/bitstream/handle/10726/1577/MFC00491.pdf?sequence=1&isAllowed=y>