école normale supérieure paris—saclay

Analysis of time series

Arima vs. arimax – which approach is better to analyze andforecast macroeconomic time series?

Alexandre Anderini Antoine Doizé

I/ Theoretical review

II/ Reproducing the results of the article

III/Contesting the article conclusion

April 2, 2021

I/ Theoretical review

$$B\big((x_t)_{t\in R}\big) = (x_{t-1})_{t\in R_*^+}$$

ARIMA :

$$\begin{aligned} y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \Phi_1 y_{t-s} + \Phi_2 y_{t-2s} + \cdots + \Phi_p y_{t-P*s} \\ + a_t - \theta_1 a_{t-1} - \cdots - \theta_p a_{t-q} + \Theta_1 a_{t-s} - \Theta_2 a_{t-2s} - \cdots - \Theta_Q a_{t-Q*s} \end{aligned} \tag{1}$$

We generalize to $ARIMA\big(p, d, q\big)\big(P, D, Q\big)$ of x_t if we apply ARIMA to $\big(1-B\big)^d\big(1-B^s\big)^D\big(x_t\big)$

ARIMAX :

$$\begin{aligned} y_t = C + \nu_0 x_t + \nu_1 x_{t-1} + \nu_2 x_{t-2} + \cdots + \nu_K x_{t-K} \\ + a_t - \theta_1 a_{t-1} - \cdots - \theta_p a_{t-q} + \Theta_1 a_{t-s} - \Theta_2 a_{t-2s} - \cdots - \Theta_Q a_{t-Q*s} \end{aligned} \tag{2}$$

Predict GDP using either ARIMA or (ARIMAX and UR)

1. different approaches
2. what to expect ?

II/ 1.GDP and UR curves

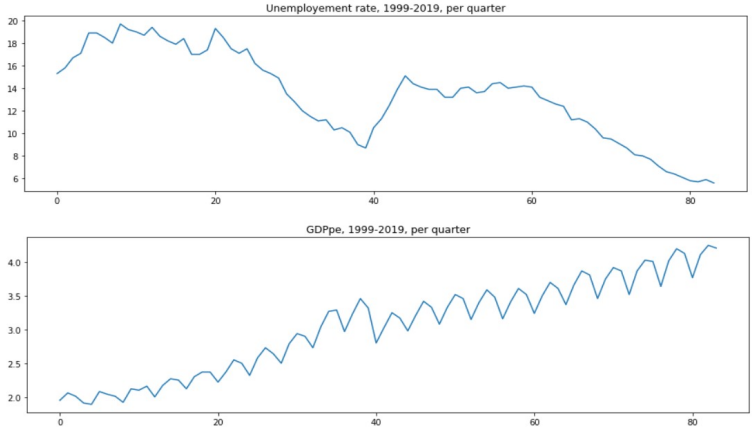
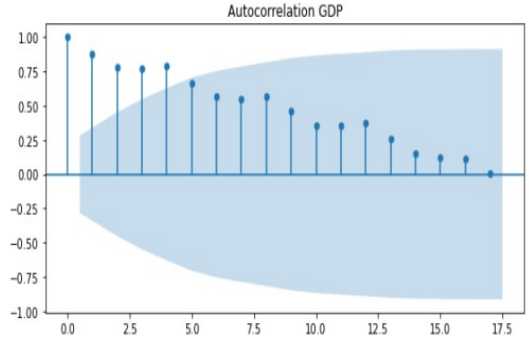


Figure: GDP and UR curves

II/ 2. GDP time analysis

ADF Statistic: -0.41222388753647227
p-value: 0.9080377328658321



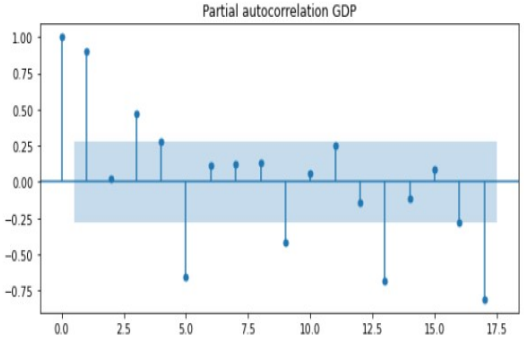
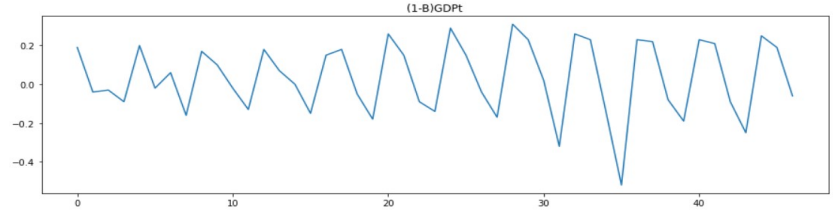
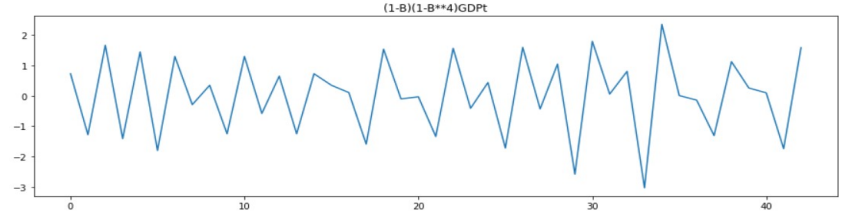


Figure: GDP autocorrelation

II/ 3. Differentiated GDP



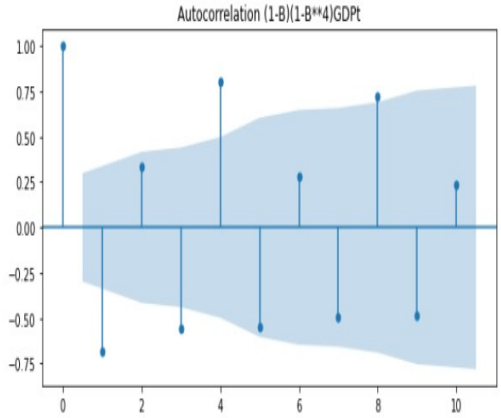
ADF Statistic: -3.2482944501445097
p-value: 0.017353957952695436



ADF Statistic: -4.580890578688619
p-value: 0.00014006087010199163

Figure: GDP differenciated

II/ 4. Differentiated GDP autocorrelation



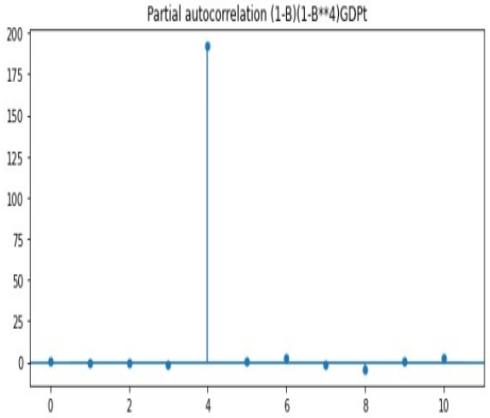


Figure: GDP differenciated

II/ 5. Sarima model results (1)

SARIMAX Results

Dep. Variable:	y	No. Observations:	48
Model:	SARIMAX(0, 1, 0)x(1, 1, 0, 4)	Log Likelihood	48.164
Date:	Fri, 02 Apr 2021	AIC	-92.327
Time:	05:02:51	BIC	-88.805
Sample:	0	HQIC	-91.029
	- 48		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.S.L4	-0.2969	0.094	-3.159	0.002	-0.481	-0.113
sigma2	0.0062	0.001	7.384	0.000	0.005	0.008

Ljung-Box (L1) (Q):	0.33	Jarque-Bera (JB):	21.84
Prob(Q):	0.57	Prob(JB):	0.00
Heteroskedasticity (H):	3.71	Skew:	0.04
Prob(H) (two-sided):	0.02	Kurtosis:	6.49

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

Figure: Sarima results

II/ 5. Sarima model results (2)

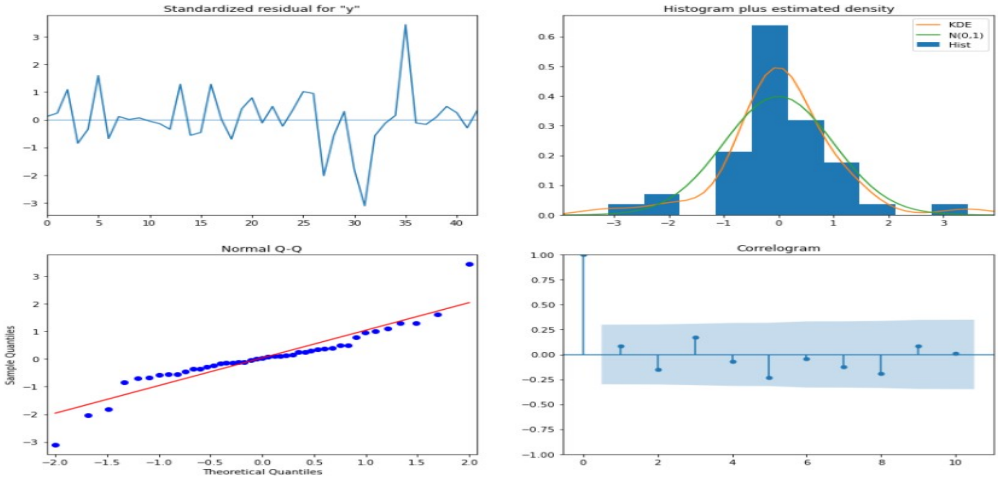


Figure: Sarima diagnostic

II/ 5. Sarima model results (3)

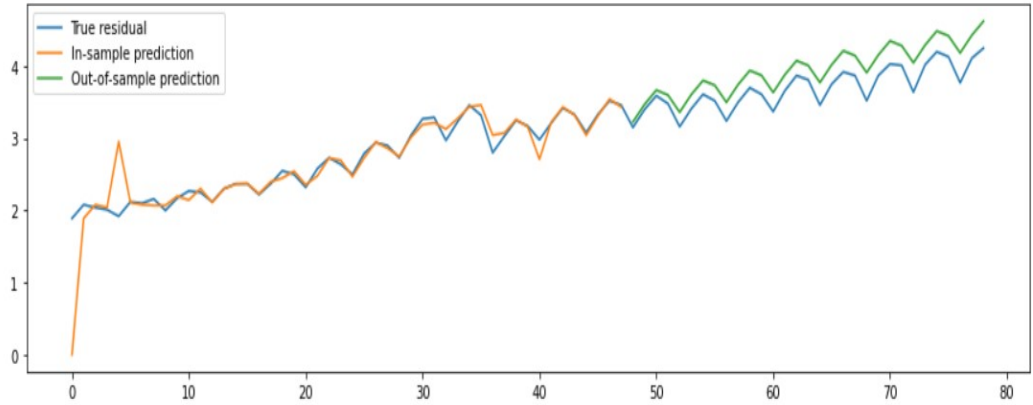


Figure: Sarima model predictions

II/ 6. Unemployment rate training plot

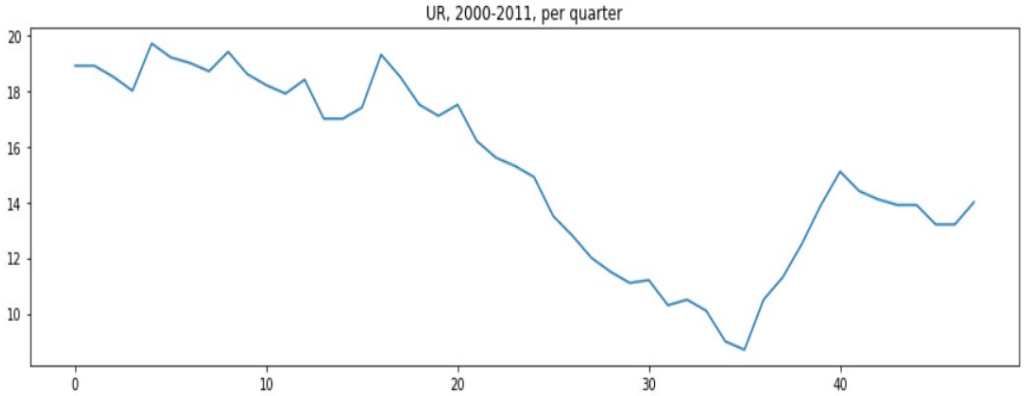
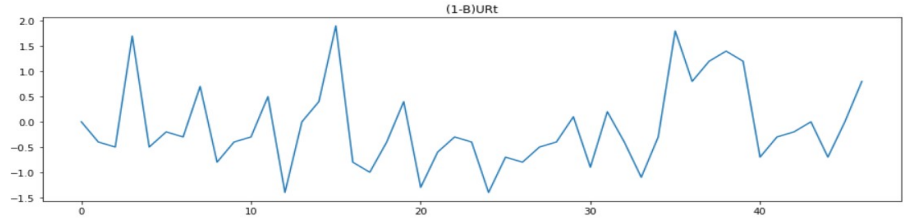
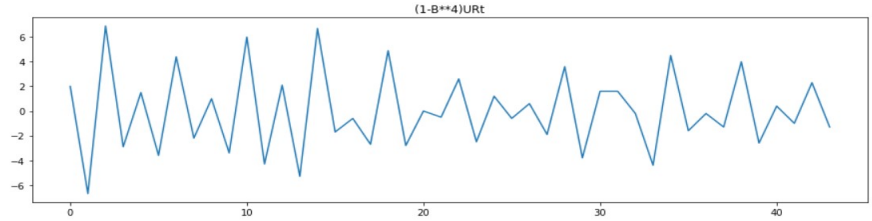


Figure: Unemployment rate training plot

II/ 7. UR training plot differentiated



ADF Statistic: -2.158903626105008
p-value: 0.2214975748881618



ADF Statistic: -3.0509339143968766
p-value: 0.030403892039647832

Figure: Differenciated unemployment rate training samples

II/ 8. UR time correlations

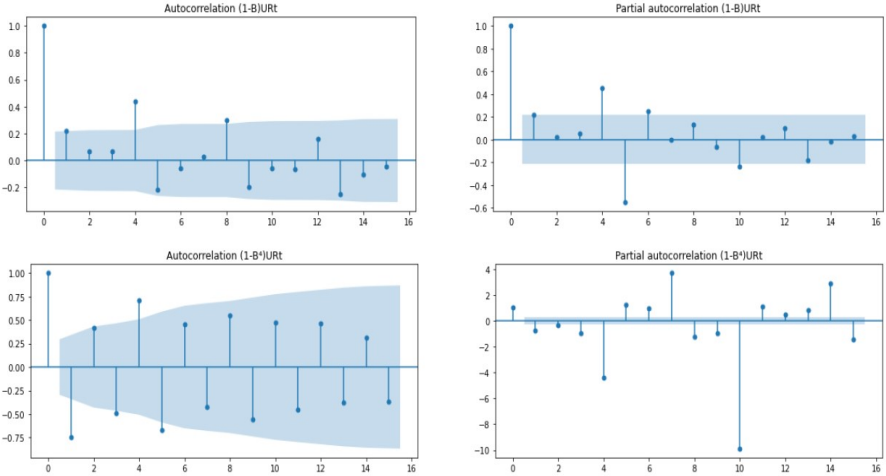


Figure: UR time autocorrelation

II/ 8. UR predictions with sarima

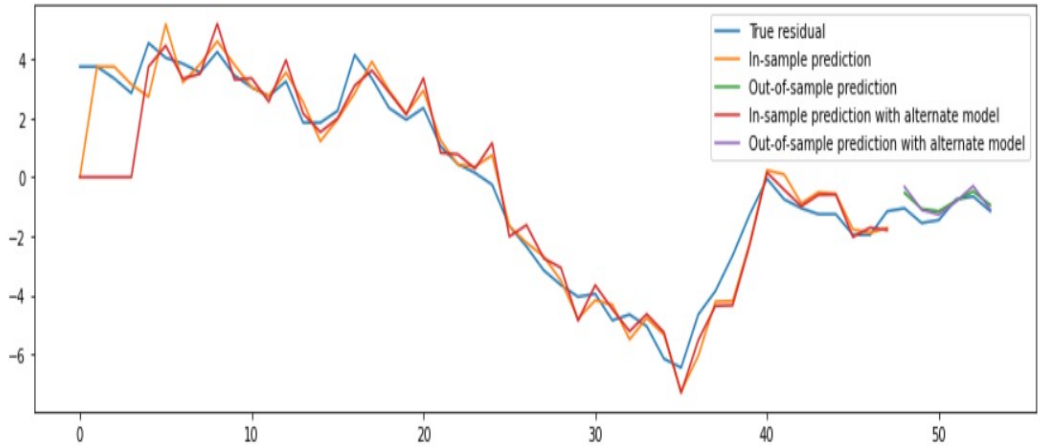


Figure: Sarima model predictions

II/ 8. GDP prediction with sarimax

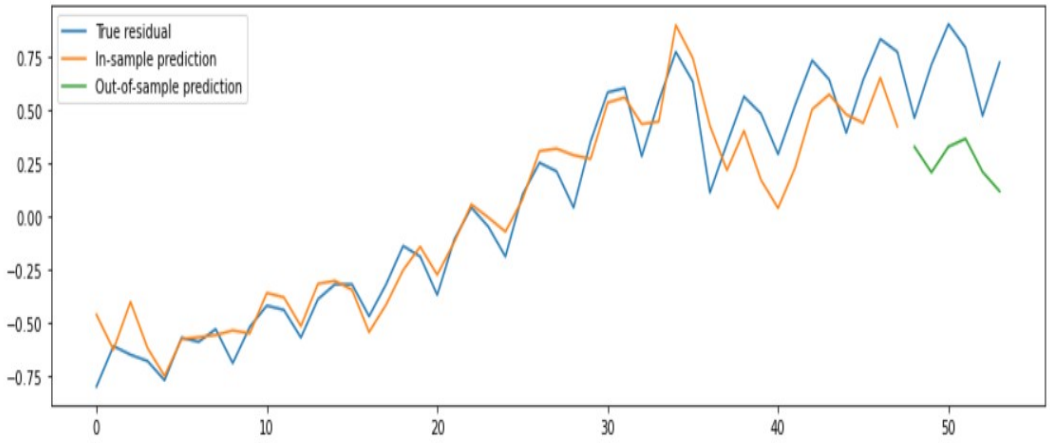


Figure: Sarima model predictions

II/ 9. Article conclusion

ARIMA $\geq$ ARIMAX ?

Better question : why that model do not yield great results

Two underlying questions are to be asked : can UR be pertinently predicted through its past and can it be used pertinently to predict GDP.

III/ 1. Varima predictions

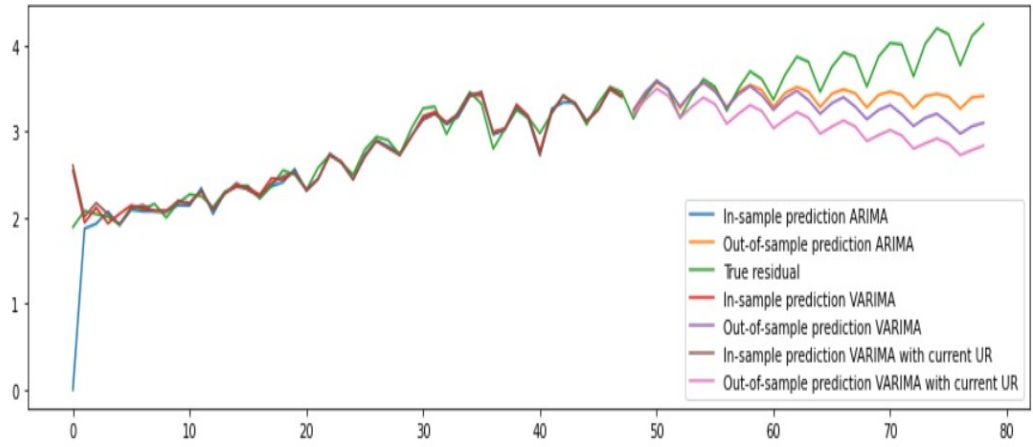


Figure: Varima model predictions

III/ 2. LSTM UR training plot

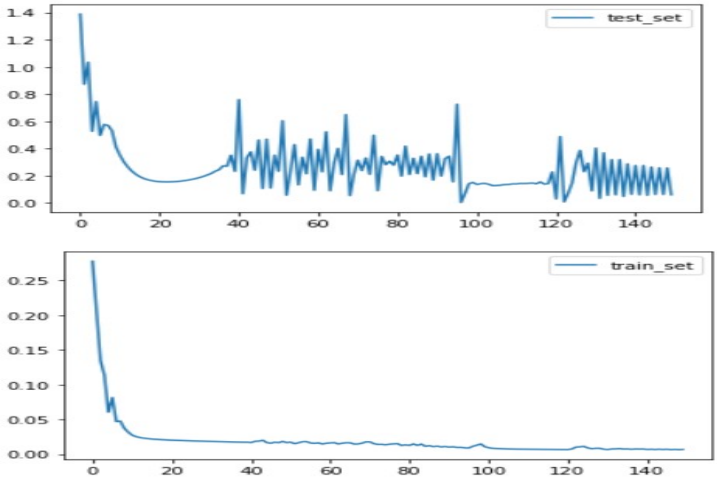


Figure: Varima model predictions

III/ 3. LSTM GDP training plot

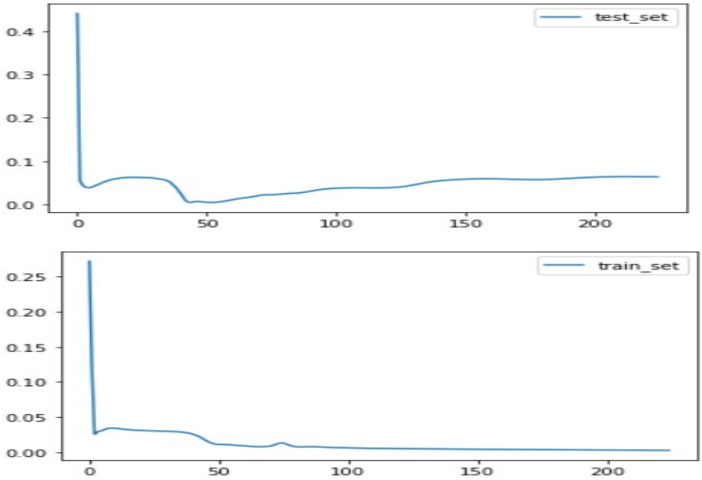


Figure: Varima model predictions

III/ 4. LSTM GDP from UR training plot

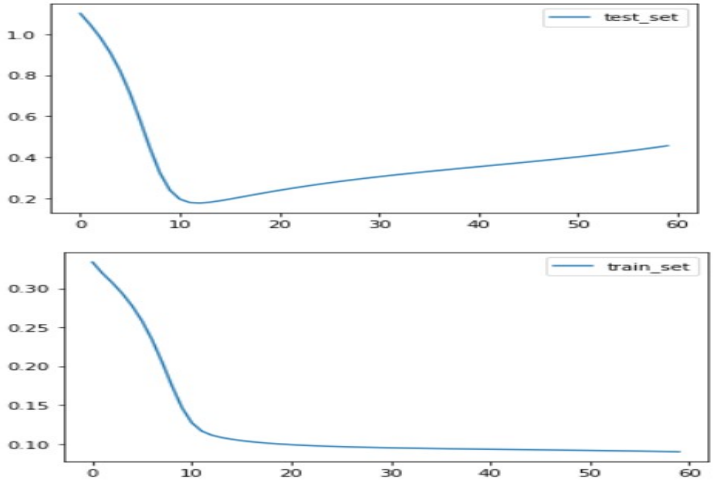


Figure: Varima model predictions

III/ 5. Our conclusion

The right approach is not :

Among ARIMA and ARIMAX which model is better

but :

Among the two tasks, which is easier?

Thank You

for Listening.