# DRIVERS OF GLOBAL $CO_2$ EMISSIONS

Costa, F. & Gouveia, J. & Rivera, P. & Rocha, N.

June 2019

## INTRODUCTION

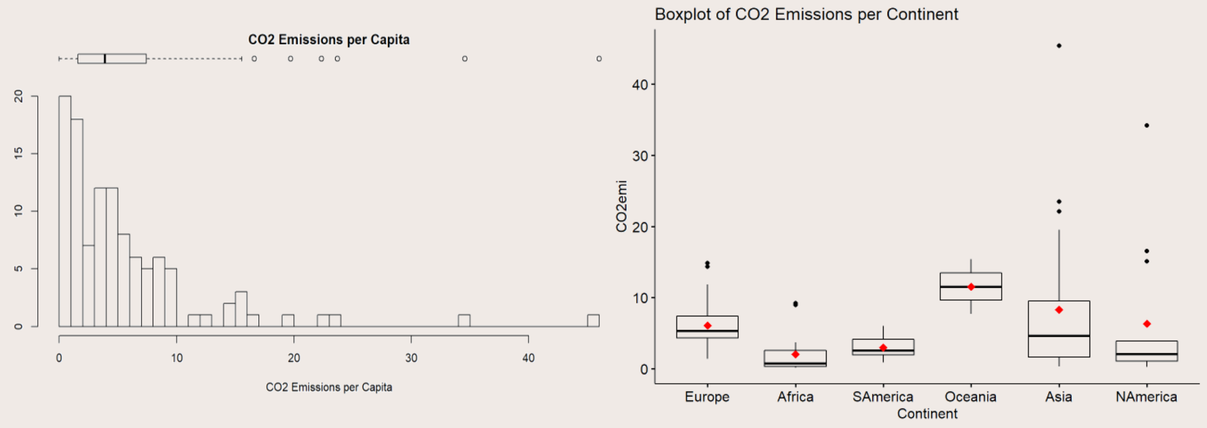
Scientists have recently detected the highest levels of planet-warming $CO_2$ in Earth's atmosphere since records began, sounding a new alarm over the rise of man-made greenhouse gas emissions. Who's there to blame? Fossil fuels? Farming practices? Cattle raising? Industrialization? Regular human activities? One thing we know for sure, climate change is real!

## DESCRIPTIVE ANALYSIS

The following variables (from 2014) served as the basis to carry out the subsequent study.

| | | | |
|---|---|---|---|
| ***Continent*** | Name of the continent | ***Upop*** | Urban pop. In thousands |
| ***CO2emi*** | $CO_2$ emissions in metric tons per capita | ***GDPcap*** | GDP per capita in US$ |
| ***Pest*** | Use in cropland in kg/há | ***ElConsCap*** | Electricity consumption per capita in kWh |
| ***CattleProd*** | Livestock count in thousands | ***ElRenProd*** | Electricity production from renewable sources in millions of kWh |
| ***Rpop*** | Rural pop. In thousands | | |

Focusing on the dependent variable, the histogram below contains all the values of $CO_2$ emissions from 112 countries worldwide.




There is evidence of high positive correlations between **CattleProd** and **Rpop** (+0.68), **CattleProd** and **Upop** (+0.72), **Rpop** and **Upop** (+0.85) and finally between **ElRenProd** and **Upop** (+0.74).

## REGRESSION ANALYSIS

One of the main objectives of this study is to identify and measure the effects of several variables on $CO_2$ emissions per capita in a given country.

By the analysis of the $CO_2$ boxplot, it is possible to notice the existence of outliers, which is also true for the independent variables. Since the regression analysis is highly susceptible, we decided to use a transformation process (Cap/Floor technique).

Also, we chose the forward selection procedure to build a model that explains the variation in $CO_2$ emissions. By starting with an empty model the sequential sum of squares was computed in order to chose the first variable to be considered.

The linear regression assumptions were tested by conducting several analysis to the residuals such as linearity, independence, normality (Shapiro-Wilk) and equal variances (Breusch-Pagan).

The process was continuously repeated until no variable was worth to be added to the model.

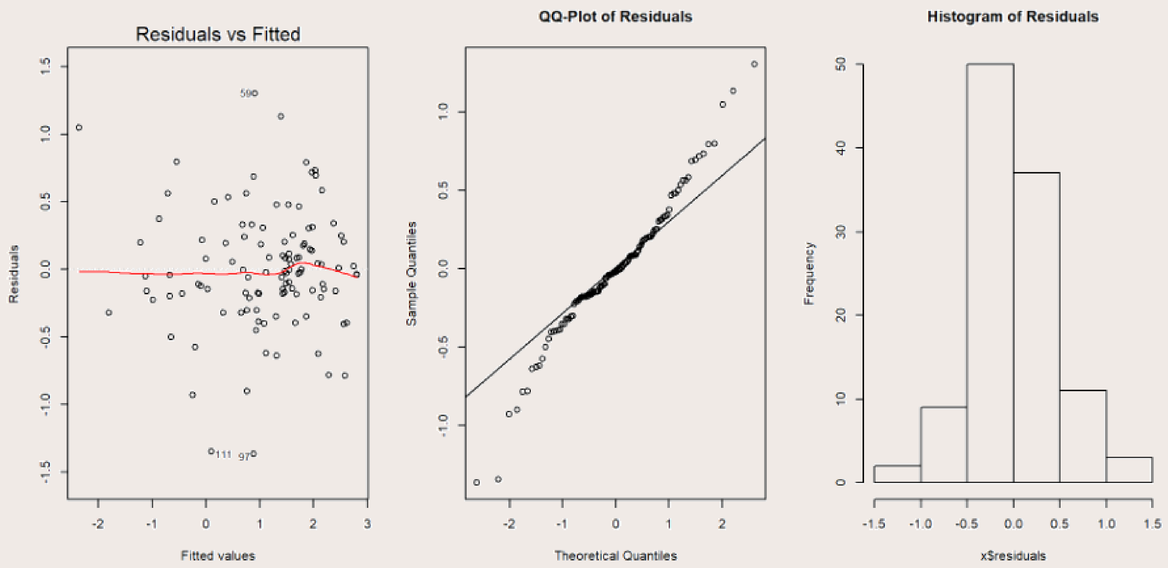
$$CO2emi = \beta_0 + \beta_1 ElConsCap + \varepsilon$$

$$\log(CO2emi) = \beta_0 + \beta_1 \log(ElConsCap) + \varepsilon$$

$$\log(CO2emi) = \beta_0 + \beta_1 \log ElConsCap + \beta_2 Asia + \varepsilon$$

$$\log(CO2emi) = \beta_0 + \beta_1 \log(ElConsCap) + \beta_2 Asia + \beta_3 Upop + \varepsilon$$

$$\log(CO2emi) = \beta_0 + \beta_1 \log(ElConsCap) + \beta_2 Asia + \beta_3 Upop + \beta_4 ElRenProd + \varepsilon$$




```
Shapiro-Wilk normality test
W = 0.97217, p-value = 0.01926

studentized Breusch-Pagan test
BP = 6.5867, df = 4, p-value = 0.1594
```

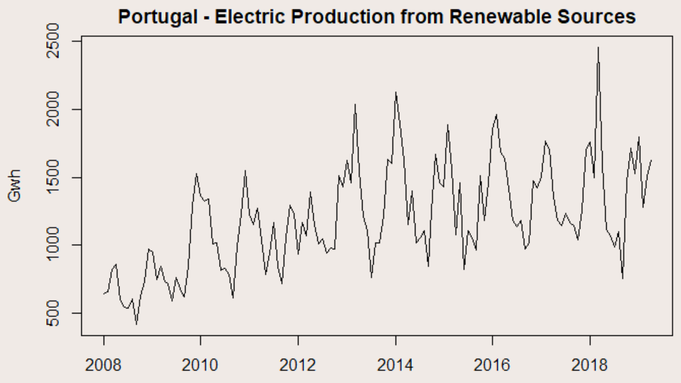
From the plots above, it is possible to conclude that all the above-mentioned assumptions hold except for normality (p = 0,0193). However, the sample size is large enough to assume that the central theorem holds, and therefore the final model is not biased nor inefficient.

Once the final model is set, the interpretation of the coefficients should be as follows.

- **log(ElConsCap)** - a 1% increase in the consumption of electricity per capita results in a 0,84% increase in $CO_2$ emissions per capita, on average, ceteris paribus.
- **Asia** - if the country is Asian, there is a positive 18,01% change in $CO_2$ emissions per capita, on average, ceteris paribus.
- **Upop** - an increase of one thousand people in the urban population results in a positive 0,00045% change in $CO_2$ emissions per capita, on average, ceteris paribus.
- **ElRenProd** - an increase of one million kWh produced from renewable sources results in a 0,00058% decrease in $CO_2$ emissions per capita, on average, ceteris paribus.

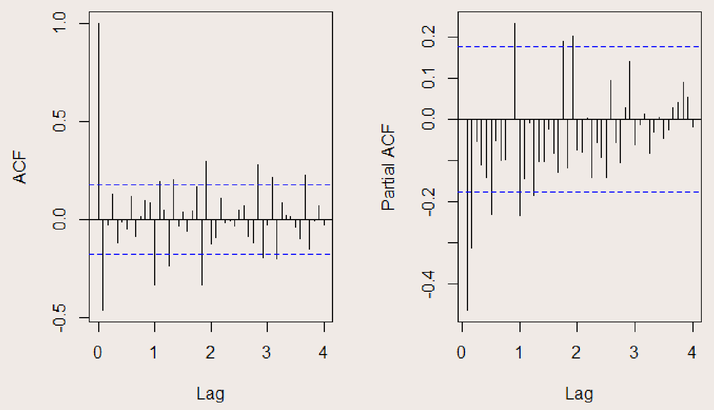
## TIME SERIES

One of the main conclusions of the multiple regression analysis is that the production of electricity from renewable sources is negatively related with $CO_2$ emissions and, consequently, with greenhouse gas emissions, which now leads to the second part of this study - building a time series model of the monthly electricity production from renewable sources in Portugal.




Applying the Box-Jenkins methodology, after the analysis of the time series' stationarity, the presence of a trend and seasonality in the series was found (Summer months register lower values, on average). In addition, the inexistence of a constant variance (Box-Cox lambda=0.015) suggested the application of the logarithm as well as two differencing transformations (one for the non-seasonal part and another for the seasonal one).

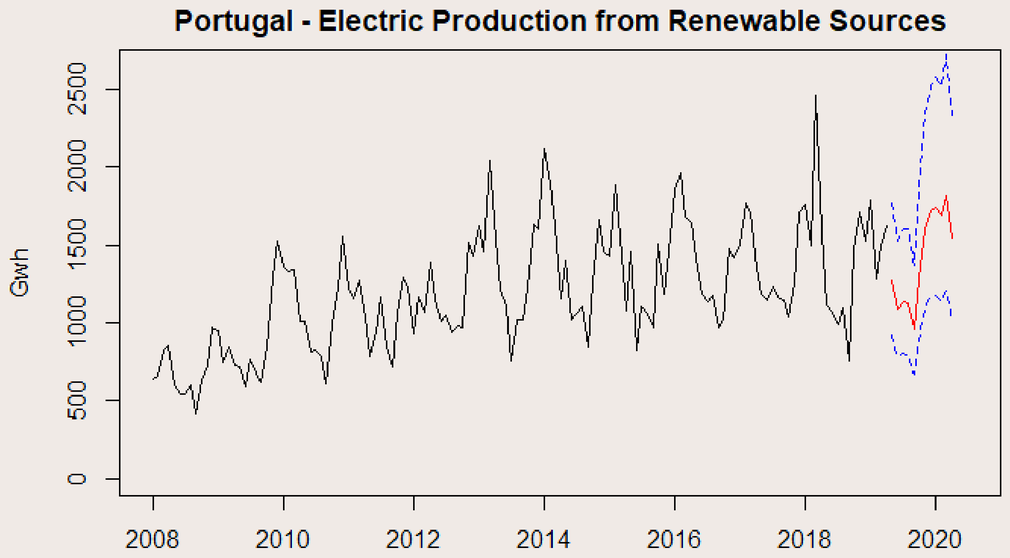
Soon after, the Autocorrelation and the Partial Autocorrelation functions were plotted to identify the order of MA and AR terms in the ARIMA model.




Regarding the non-seasonal part, there was a significant spike in the ACF, indicating MA(1). In addition, the PACF shows two significant spikes, suggesting AR(2).

As for the seasonal one, there is a significant spike on lag 12 (MA(1)), with the same occurring in the PACF (AR(1)). By testing different models, it was possible to identify $(0,1,1)(0,1,1)_{12}$ as the best model, according to its AIC.

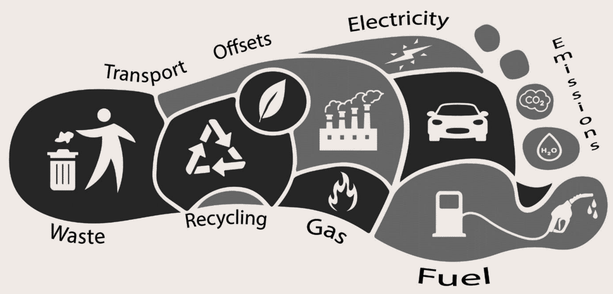
After performing the Portmanteau test and after evidencing that the residuals were independent (p=0.76), it was possible to forecast the values of electricity production from renewable sources until May 2020 as shown in the figure below.




## CONCLUSION

All things considered one can conclude that the factors discussed above help explain the levels of $CO_2$ emissions of a given country. Although this study focused essentially on agriculture and demographic variables, the economic situation might also play an essential role on greenhouse gas emissions, which would be interesting to analyze in future studies.

In the specific case of Portugal, the recent growth of electricity production from renewable sources represents a clear urge to tackle an issue that has contributed to a climate crisis, a problem that jeopardizes not only the future of humanity but all the planet's ecosystem as well. As our carbon footprint gets bigger, the planet demands urgent action. Let's not wait until it's too late!




## REFERENCES


- https://data.worldbank.org
- http://www.fao.org/faostat/en/#data/EP/visualize
- https://ourworldindata.org/meat-and-seafood-production-consumption
- http://www.centrodeinformacao.ren.pt/PT/InformacaoExploracao/Paginas/EstatisticaMensal.asp
- Larose, D & Larose, C - Data Mining and Predictive Analytics (2015)
- Shumway, R & Stoffer, D - Time Series Analysis and Its Applications (2011)
- Newbold, P & Thorne, B & Carlson, W - Statistics for Business and Economics (2013)
- Makridakis, S. & Wheelwright, S.C. & Hyndman, R.J. - Forecasting:Methods and Applications (1998)