Explorative data analysis of electricity production, is it possible to decouple it from fossil fuel?

Final paper for the course Introduction to Sustainable Development and Global Governance

Davide Rosatelli

Academic year 2022/2023

Word count: 2898 (3948 including graphs and tables)

Complete analysis: https://github.com/DavideRosatelli/Energy-data-exploration

1. Introduction

One of the central themes in the field of sustainability is undoubtedly that of energy sources. Mankind, and the world it has built, is in constant and ever-increasing need of energy, but it has only recently realised that the sources of energy that have enabled their huge growth are coming to an end, and are deeply affecting the planet's ecosystem.

The seventh of the 17 Sustainable Development Goals (SDGs) of United Nations is "*Ensure access to affordable, reliable, sustainable and modern energy for all*", and every country in the world, in different ways, is moving towards this goal. Having enough energy in the world to satisfy every people is the great challenge for the next future, especially if this energy must be sustainable.

The transition that the world needs to have sustainable energy everywhere is extremely complex and costly, which is why it must be monitored under each aspect. In this paper, data from around the world from the last three decades are considered and analysed to draw insights about this transition; the focus of this work is on electricity production, since as a second energy carrier, can provide information about the primary source from which its produced.

2. Historical background

Studying climate change and all the processes related to it through historical data is crucial, to be able to give the right context to each event, understand the factors that influenced it, and apply this knowledge to the future.

History teaches us that man has always exploited and modified the environment in which he lived to benefit from it, but the change that really made man's impact on nature exponential from that point on, was the industrial revolution, and consequently the use of coal.

Since about 1750, firstly in Europe and then everywhere, the world has witnessed great growth in terms of population, life expectancy and GDP per capita. In the twentieth and twenty-first century this situation has exaggerated, leading the various countries to obtain the so-called modern economic growth, defined as: "*long-term rise in the capacity to supply increasingly diverse economic goods to its population*". All this has been possible thanks to an ever-increasing use of energy, almost entirely derived from fossil fuels. With the increase in population and the improvement of technology, it has been necessary over time to make the ways of generating energy more and more efficient, but soon this will no longer be sufficient, given the fact that fossil resources will run out and we can no longer ignore their environmental impact.

The production of electricity from fossil fuels generates a high quantity of CO_2, an emission related to greenhouse effects and generally indicator of environmental degradation; however, for the growth of a country it is essential to have an adequate amount of energy, and being large energy producers for states is generally synonymous with wealth. Energy production is therefore interesting to be studied together with GDP, to verify

if in these data we can find evidence of the fact that countries, especially the richer ones, are switching to using sustainable sources, and if this affects their economic growth. The theory of the Environmental Kuznets Curve argues that the factors of environmental degradation tend to decrease in rich countries once a critical point is reached; it is then reasonable to test this theory on a variable that represents economic growth and environmental damage at the same time.

3. Research questions

In this work, the main objective is to extract, from a large and complex dataset, some information on whether a transition towards the sole use of sustainable sources is taking place, and how rapid and homogenous this transition is worldwide. To this end, the various amounts of sources to generate energy used over the past 30 years and in certain countries will be analysed to answer the question:

Considering electricity production, do the trends over time and over different countries suggest that is it possible to decouple it from fossil fuel?

In addition to the analysis of these variables alone, it is necessary to investigate what influences them most, and in this analysis the focus is on the GDP per capita of the countries. As already mentioned, electricity production is closely linked to a country's economic growth; it shall therefore be verified from the data if their correlation is linear and positive or if it can be guessed that the use of fossil fuels is a variable that will decrease given its environmental impact. The second research question is then:

Are these trends related with GDP? Is it possible to notice evidence of the EKC theory considering the production from fossil fuels?

4. Methods and dataset

The analysis for this paper was carried out using Python, the full notebook can be found here: "*https://github.com/DavideRosatelli/Energy-data-exploration*". The packages Pandas and NumPy were used for data manipulation, while Matplotlib and Plotly were used for data visualisation.

The dataset used for this work is the energy dataset by "*Our world in data*", which can be downloaded as a csv file from Kaggle. The data are updated regularly, and were created by several reliable sources, the main ones being BP Statistical Review of World Energy, SHIFT Data Portal, and EMBER - Global Electricity Dashboard.

The dataset is large and very detailed, containing over 17,000 rows, each having values for a specific year (between 1900 and 2019) and per country or geographical area. Each row contains values for 122 columns, which include general data such as GDP and population, and information about each type of resource to generate energy. For each resource, and for each country and year, we have the total production of the resource,

the electricity produced from them, the change from one year to the next, the share of total energy, and more. The complete description of the dataset and variables can be found on Kaggle, while for this work I used mainly those data concerning the production of electricity.

In the first steps of the code, after importing the data and packages, I noticed that there were many missing values, probably due to the difficulty of finding verified data for past years and for each country. The list of variables I worked with, and from which I extracted other values, is the following:

['iso_code', 'country', 'year', 'biofuel_electricity', 'hydro_electricity', 'nuclear_electricity', 'solar_electricity', 'wind_electricity', 'other_renewable_electricity', 'coal_electricity', 'gas_electricity', 'oil_electricity', 'population', 'gdp']

Regarding the year range, to deal again with the missing values problem, only data from 1990 to 2018 were considered, since the latest period is the most important for this analysis.

5. Data visualization and results

5.1 Total electricity production by source over time

For the first step of this analysis, I ignored the geographical areas, grouping by years the variables indicating the electricity production by the different sources, to obtain a data-frame showing for each year between 1990 and 2018 the total electricity generated from the specific sources.

In the first stacked bar plot (Fig.1), it is possible to observe the total production of electricity over the years, which as expected is continuously increasing. In 1990 the Terawatts per Hour produced were just over 35000, while in 2018 this number exceeded 80000; this growth is totally constant, except for a single year of recession in 2009, probably related to the economic crisis. It can also be seen that only in recent years many energy sources have begun to gain importance, whereas in the early 1990s almost all the electricity came from fossil fuels, nuclear power plants and water.

The next two graphs show the same type of relationship, but containing only energy produced from fossil fuels (Fig. 2) and only energy from sustainable sources (Fig. 3), respectively. Again, both productions are increasing, but is it possible to note that the proportion of growth is higher in sustainable sources: it goes from about 7500 to more than 20000 TwH, so it almost tripled over these years, while fossil energy, which still accounts for most of the energy produced, starts from about 22000 to reach in 2018 50000 TwH, so it just doubled. What is also interesting to notice is the change of preferred source over years, coal being the most used at all, followed by gas and water that are continuisly gaining shares over years, while other sources like sun and wind became relevant only from the early 2000s.

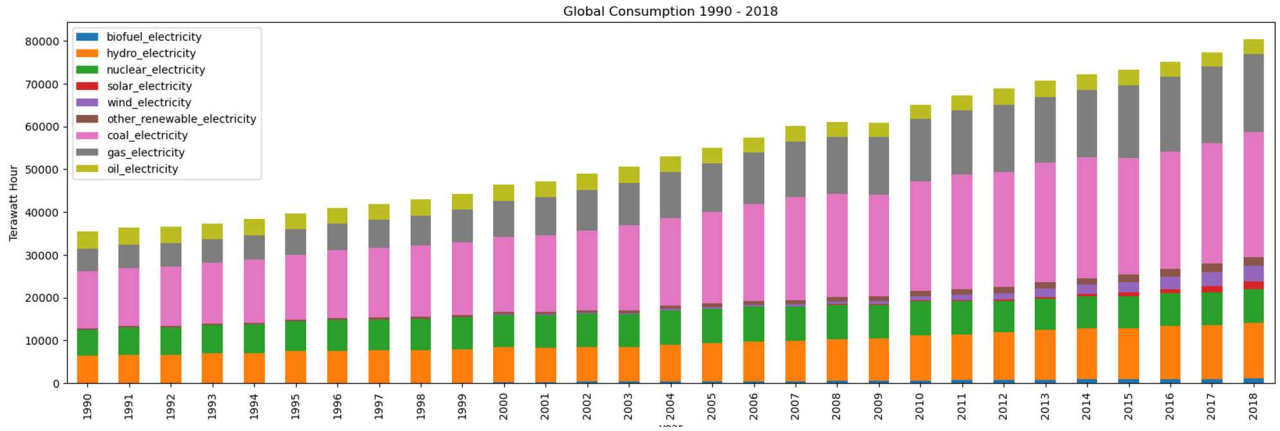


Figure 1

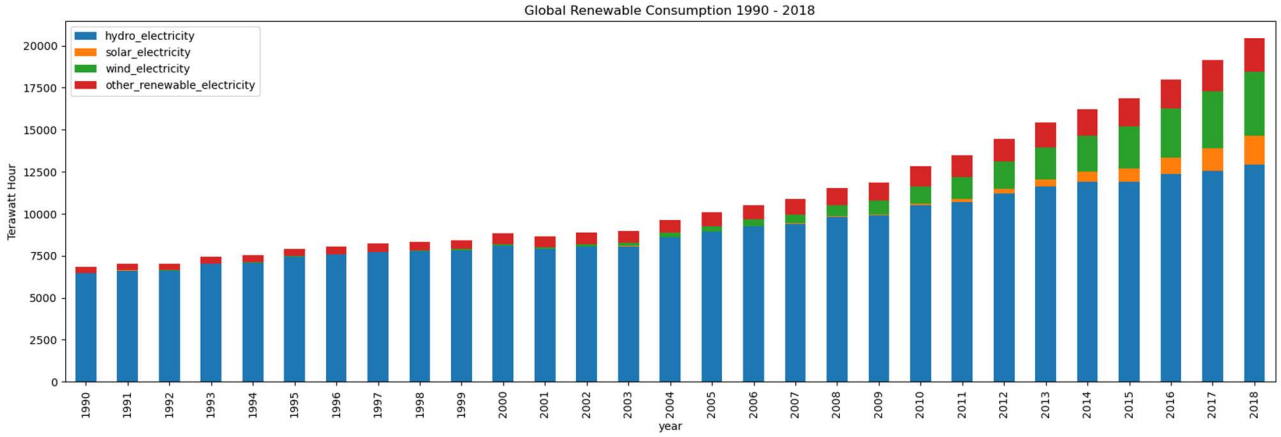


Figure 2

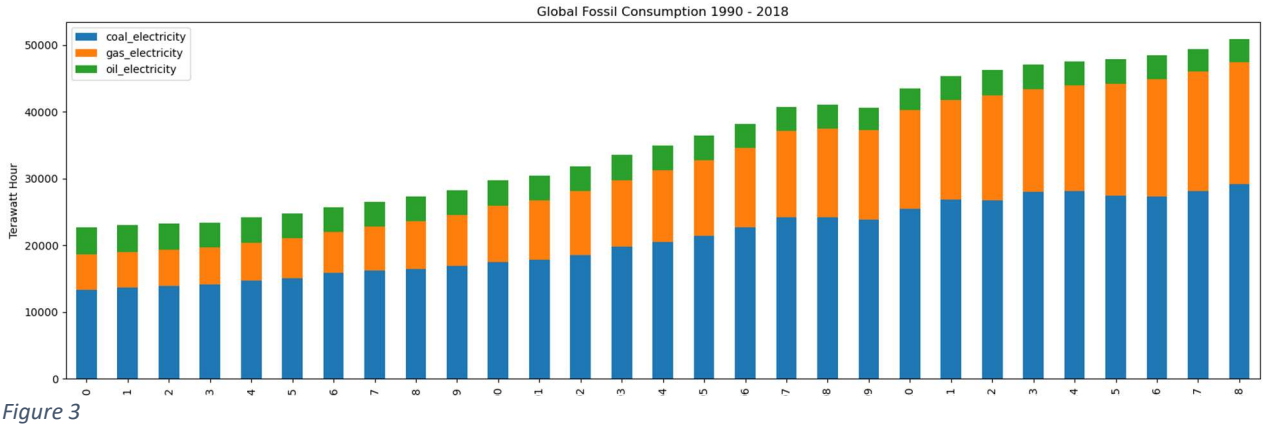


Figure 3

5.2 Sources of electricity production over years

In the next step of the analysis, I analysed again the cumulative production of electricity during the period 1990-2018, plotting a time series for each of the main resources (Appendix A).

Looking at these time series for fossil sources, is immediately noticeable that coal and gas are increasing in a linear way, while oil follows no clear trend. This can be explained by the fact that the "oil" data refers to fuel oil, a residual product of petroleum refining, used to generate electricity in small power plants, but always less used than the directly mined and cheaper products. Gas and coal have always been used to generate energy; coal is the oldest method, while in recent years the production of electricity from natural gas is becoming the most used, because compared to coal, burning the same mass produces more energy with a lower percentage of greenhouse gas emissions.

The time series for the three principal sustainable sources of energy, namely hydro, wind and solar, are all increasing. The production from water follows a linear trend, similar to gas and coal, since among the sustainable resources it is the one used from the oldest time, and which, as the graph shows, over the years has been improved and has maintained its importance. The production of electricity from sun and wind follows an exponential shape, this suggests that these resources are new, but that they are spreading and are growing rapidly, and this growth is obviously derived from the need to find alternatives to fossil fuels.

5.3 Total electricity production by source in some countries

For the next part of the analysis, some countries were analysed more in detail, to observe through the data the differences in energy production across the world. In order to have a sample that includes very different states, but gives an idea of what the world situation may be, the 10 countries chosen are the following: Egypt, United Arab Emirates, Colombia, Italy, Germany, United States, Japan, India, Russia and China.

In the bar plot (Fig. 4) the differences in energy production can be seen, first in the amount produced, which is incredibly higher than average in the USA and China. Compared to the previous analysis, here the composition of the resources used can be seen to vary widely from country to country, but even here one can glimpse the trends of the most used sources. This first graph shows the total energy production, but to compare these countries with each other, it is necessary to look at the production data per capita, which are shown in bar plots for each of the main production methods (Appendix B).

The graphs show that in these countries the main sources from which energy is generated change a lot, but even when we talk about production per capita, the most used resources are coal, gas and water, and in general worldwide there is a greater use of fossil fuels than of sustainable sources. USA are the biggest producer of energy from coal, while among these countries, the UAE far exceeds all the others in terms of production from gas; Italy and Japan appear to be large producers from oil, but looking at the numbers, we are reminded that this figure is much lower than the other two fossil sources. Among these countries, the ones that seem to make

good use of sustainable sources compared to others are Russia for water, Germany and USA for wind, and again Germany, followed by Italy and Japan, for sun. In general, the energy source preferred by different countries is not the same, and is certainly influenced by many factors that are not measured here.

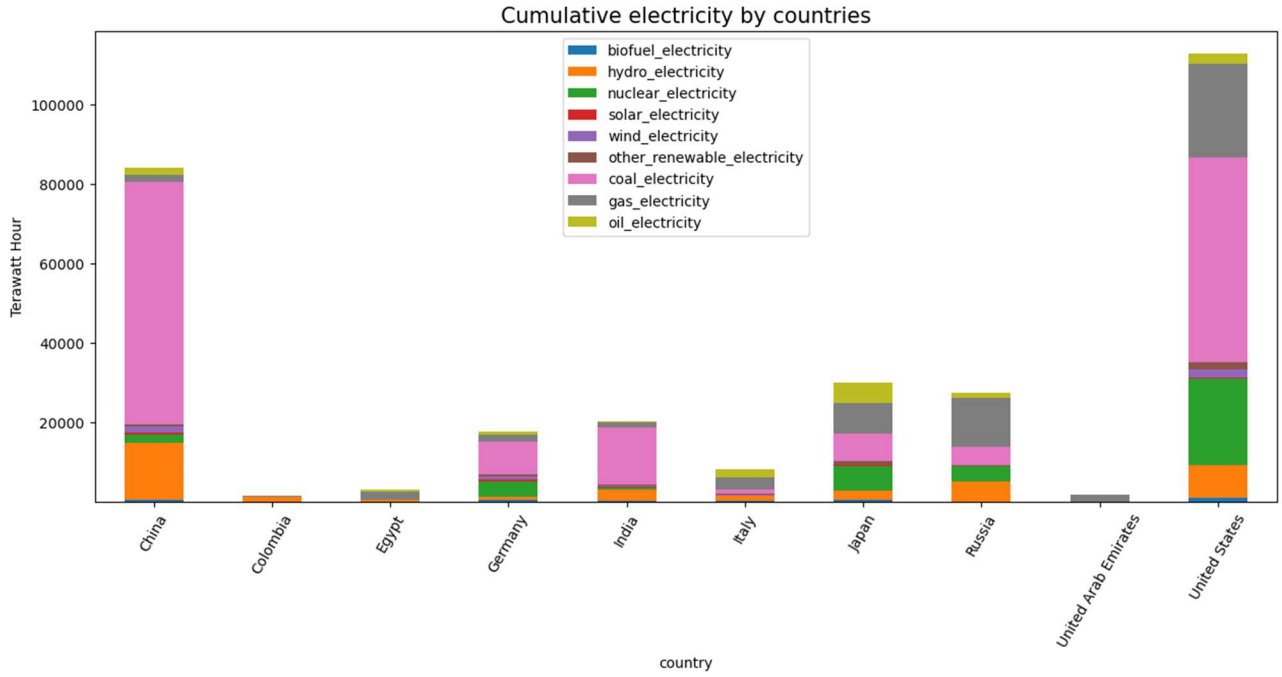


Figure 4

5.4 Comparing GDP and energy production

In the last part of this analysis, data on energy production per capita were compared with the GDP per capita of the countries. Table 1 shows, for the same countries considered in the previous part, production from fossil sources, production from renewable sources and GDP, all in per capita terms; the data refers to the year 2015, since it was the last of the few years with available data about GDP. Here, the countries with the highest GDP per capita are in order the United Arab Emirates, the United States and Germany; among these, only the USA

country	sus_electricity	fossil_electricity	gdp
China	9.509201e-07	2.968906e-06	11657.265035
Colombia	1.013426e-06	5.653290e-07	12604.953599
Egypt	1.691421e-07	1.597006e-06	11574.700085
Germany	1.690978e-06	4.423686e-06	47073.495788
India	1.309520e-07	6.967222e-07	5396.320691
Italy	1.375466e-06	2.840305e-06	34831.126812
Japan	9.933508e-07	6.324710e-06	35629.175294
Russia	1.150450e-06	4.528034e-06	22967.894498
United Arab Emirates	3.217100e-08	1.289550e-05	68984.130411
United States	1.476387e-06	8.543041e-06	52667.989570

Table 1

and Germany have a high value of sustainable energy production, while the UAE is the state that uses the most fossil resources. Other countries on the list that make extensive use of renewables include Colombia, Italy and Russia, and this variable does not necessarily seem to be correlated with a high GDP.

In the last two scatter plots, all countries were again considered, and their GDP per capita and energy production per capita were compared, first from only sustainable sources (Fig. 5) and then from fossil sources (Fig. 6)

In the sustainable electricity graph, we can see that most countries have a level not exceeding 10μ, and there are two outliers with values above 25μ and 40μ respectively. There is a large cluster at the bottom left, which shows that poor countries are not yet large users of renewables, but in general we do not see a strong correlation between the two variables.

The second scatter plot is particularly interesting to analyse, because if the production of fossil energy can be considered as a factor of environmental degradation, the graph could be interpreted as a Kuznets curve. However, the trend in the graph is not well defined and certainly does not follow the shape of a downward curve; anyway, compared to the previous one there is certainly a more linear relationship between the two variables. Again, there is a cluster in the bottom left and several outliers, underlining the inequality among the countries of the world.

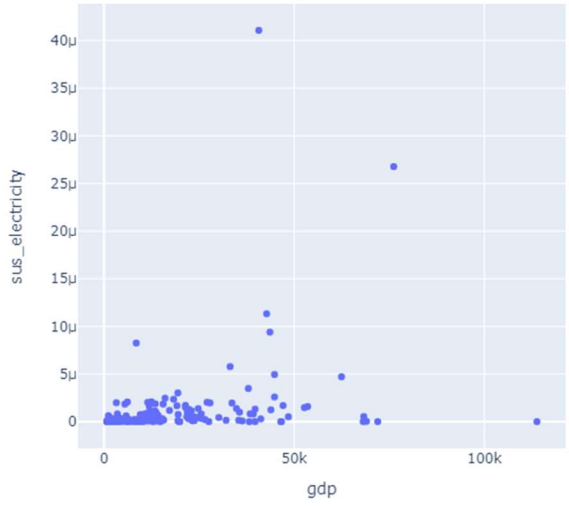


Figure 5

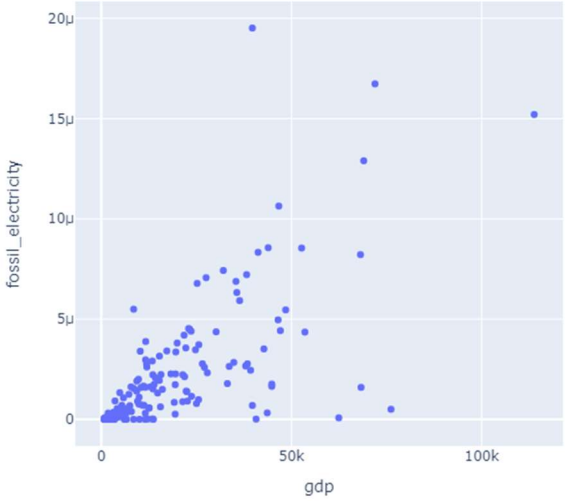


Figure 6

6. Conclusions

By reviewing the insights extracted from the graphs, it is possible to partly answer the research questions posed above.

In the period 1990-2018, electricity production from sustainable sources accounts for a small percentage of the total. We can certainly see a growing trend for all renewables, but in 2018 not even hydro production, the most developed, reaches the numbers of the fossil sources.

However, it must be considered that this data does not include the last 5 years, and that if the rate of transition has been the same, the situation should be improved now, as the growth in the use of clean resources is faster than traditional ones, and the current research takes the environmental impact into account. This is confirmed by the time series of energy production from the various sources, in which we can see the linear trend of the older sources, mainly coal, gas and water, while the newer and sustainable forms of production, such as from wind and sun, have an exponential growth rate.

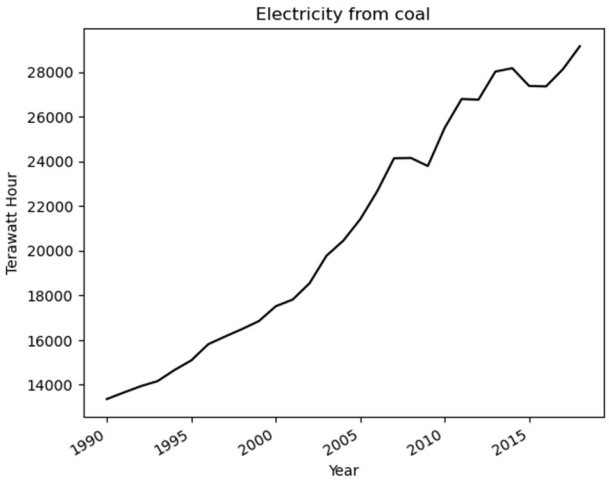
The second part of the analysis, in which few specific countries were considered, shows how in different parts of the world the most used resources vary greatly. Here it is also showed that fossil fuels are used in every country, while sustainable fuels are widespread in only a few, first among these Germany.

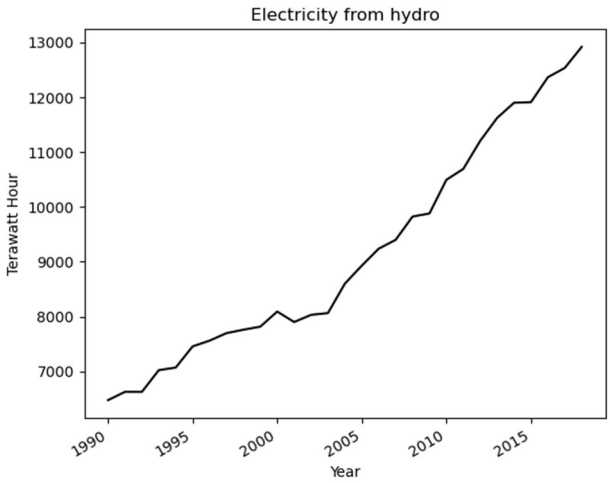
Looking at the GDP of these countries, one can see that their wealth is not necessarily linked to their sustainable energy development, as for instance in the United Arab Emirates. This relationship with GDP is confirmed in the scatter plots showing all countries: in the year to which these data refer, is it possible to see a linear relationship between GDP and energy production, but this relationship is stronger for fossil sources, while for sustainable ones the average production level is lower, and some isolated countries that turn out to be large clean energy producers are not necessarily the richest.

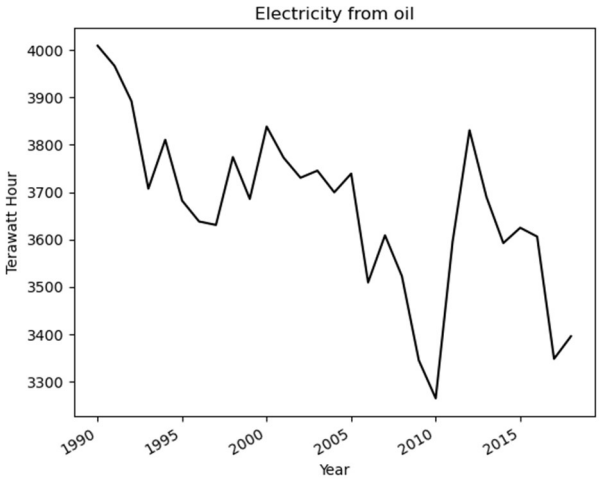
At first glance, the relationship between fossil production and GDP per capita does not appear to be related with the Kuznets curve theory: if this was the case, the graph should have a downward curve shape, in other words, the richer countries should have already abandoned this more harmful type of production. However, it can be assumed that over the years the graph will take this shape, but for this to happen research and investment in sustainable sources must continue to grow, and when countries will be able to obtain an adequate amount of energy only from these sources, they will be ready to abandon fossil sources.

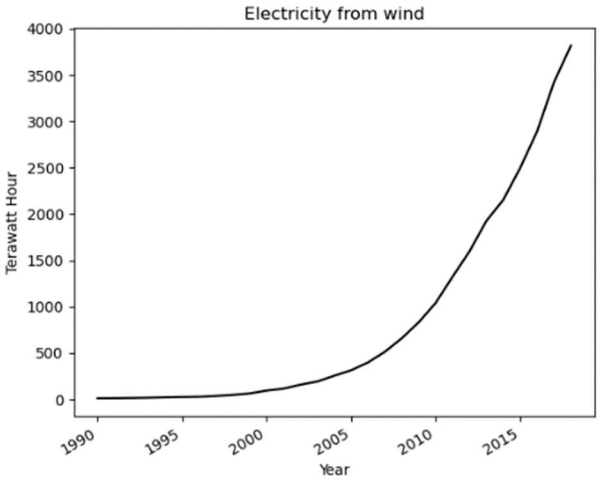
What can be concluded from this analysis is that the world has become aware of the problem, and it is certainly going through a transition, but this transition is still in its early stages, and it will be a long process to go through. Detaching completely from fossil resources is not possible right now, but the data show that if we continue in this direction, one day it could be achievable. Such a great goal will require the cooperation of the whole world, but at the moment, not all countries seem ready to give up part of their economic growth for the sake of the planet.

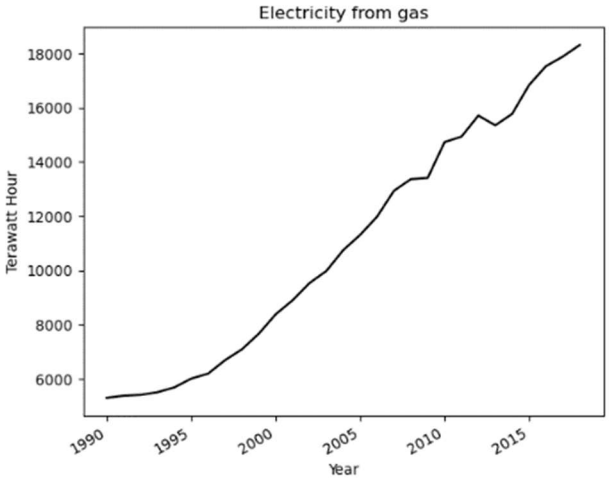
Appendix A.

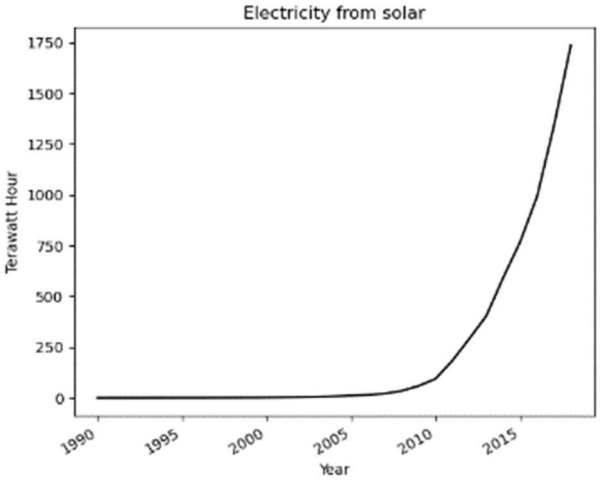

Electricity from coal
Terawatt Hour
Year


Electricity from hydro
Terawatt Hour
Year

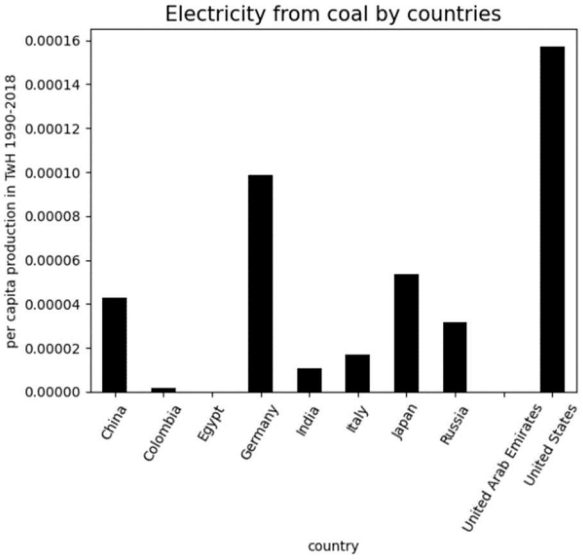

Electricity from oil
Terawatt Hour
Year

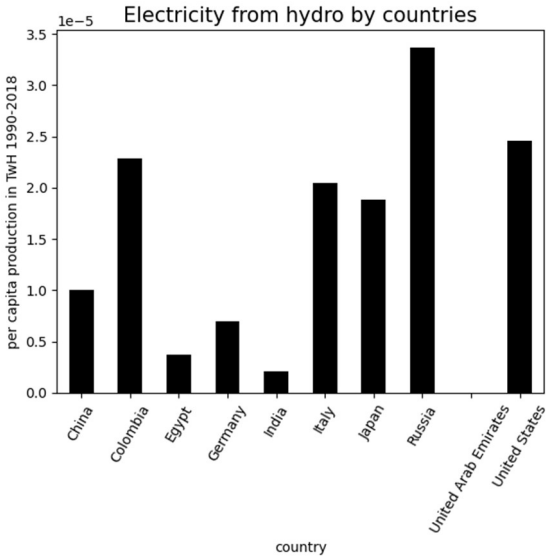

Electricity from wind
Terawatt Hour
Year

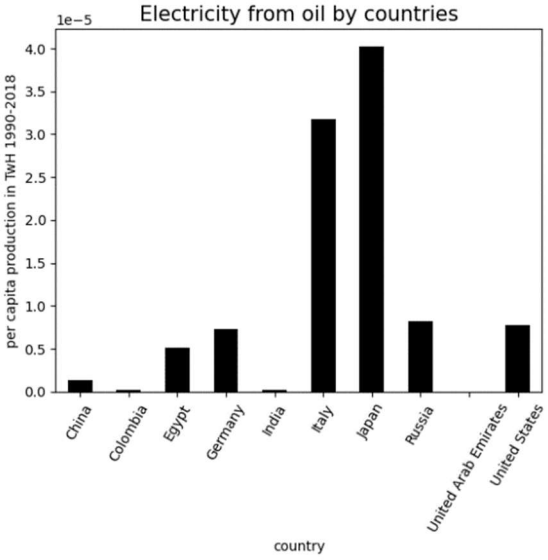

Electricity from gas
Terawatt Hour
Year

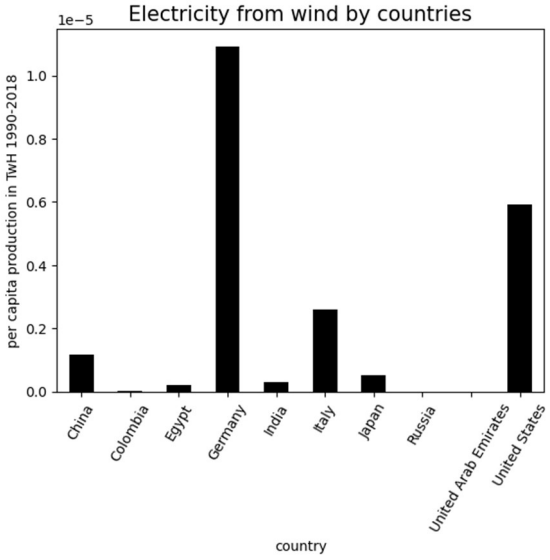

Electricity from solar
Terawatt Hour
Year

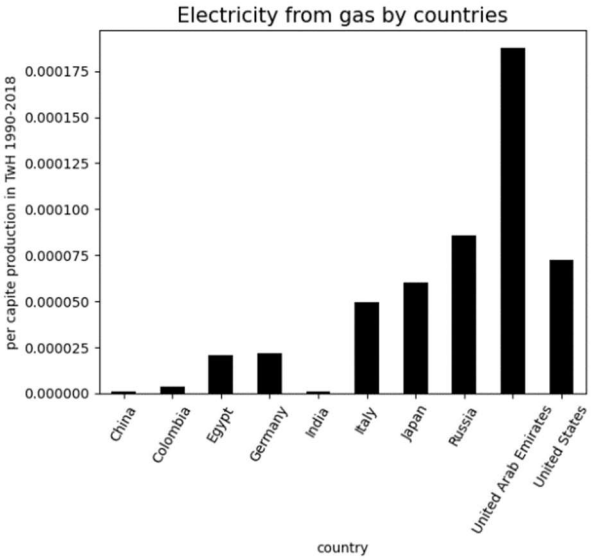
Appendix B.

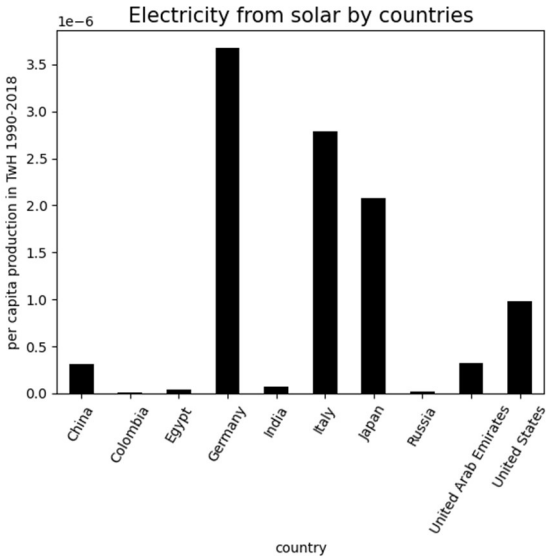

Electricity from coal by countries
per capita production in TwH 1990-2018
0.00016
0.00014
0.00012
0.00010
0.00008
0.00006
0.00004
0.00002
0.00000
China
Colombia
Egypt
Germany
India
Italy
Japan
Russia
United Arab Emirates
United States
country


Electricity from hydro by countries
1e−5
per capita production in TwH 1990-2018
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
China
Colombia
Egypt
Germany
India
Italy
Japan
Russia
United Arab Emirates
United States
country


Electricity from oil by countries
1e−5
per capita production in TwH 1990-2018
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
China
Colombia
Egypt
Germany
India
Italy
Japan
Russia
United Arab Emirates
United States
country


Electricity from wind by countries
1e−5
per capita production in TwH 1990-2018
1.0
0.8
0.6
0.4
0.2
0.0
China
Colombia
Egypt
Germany
India
Italy
Japan
Russia
United Arab Emirates
United States
country


Electricity from gas by countries
per capite production in TwH 1990-2018
0.000175
0.000150
0.000125
0.000100
0.000075
0.000050
0.000025
0.000000
China
Colombia
Egypt
Germany
India
Italy
Japan
Russia
United Arab Emirates
United States
country


Electricity from solar by countries
1e−6
per capita production in TwH 1990-2018
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
China
Colombia
Egypt
Germany
India
Italy
Japan
Russia
United Arab Emirates
United States
country

References

Wrigley (2010) “Energy and the English Industrial Revolution”, Cambridge University Press

Tubi et al. (2021) “Can we learn from the past? Towards better analogies and historical inference in society-environmental change research”, Global Environmental Change

McNeill (2015) “Energy, population, and environmental change since 1750: entering the Anthropocene”, University of Cambridge Centre of International Studies

Fernihough and ORourke (2020) “Coal and the European Industrial Revolution”, The Economic Journal

Ross (2012) “The Oil Curse”, Princeton University Press

Stern (2004) “The Rise and Fall of the Environmental Kuznets Curve”, World Development

Our world in data, “Energy”, https://ourworldindata.org/energy

Kaggle, “World Energy Consumption”, https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption

The Global Goals, “7 Affordable and clean energy”, https://www.globalgoals.org/goals/7-affordable-and-clean-energy/

Let’s talk science, “Generating Electricity: Fossil Fuels”, https://letstalkscience.ca/educational-resources/backgrounders/generating-electricity-fossil-fuels

Just energy, “How Is Oil Needed for Gas and Electricity?”, https://justenergy.com/blog/oil-needed-gas-and-electricity/