CO2 emission vs. GDP growth

Is there a correlation between GDP growth and CO2 emission rates by countries/regions?

Background

Climate change refers to long-term shifts in temperatures and weather patterns. Climate change is a big problem not only for local areas, but for the whole planet on a global scale. As pollutants accumulate in our atmosphere, the so-called greenhouse effect causes global warming, which leads to climate change. Through the last decade humanity realized the importance of climate change and started to react to the levels of greenhouse emissions.

The scope of study

This study examines the influence of economic growth on environmental degradation. It is well known that CO2 is one of the worst pollutants contributing to environmental issues (Houghton, 1996). This project was chosen to understand the relationship between GDP growth and CO2 (one of greenhouse gasses) emissions rate for the different countries and regions with the potential to observe the trend each country has.

We would like to split our analyses into three different stages:

Worldwide level analysis:

At this stage we wish to analyze the general CO2 emission distribution and how it has changed globally throughout history, relationship analyses with GDP growth.

Specific questions:

1. What's the global distribution pattern of CO2 emission? How does it change throughout history?
2. Is there a correlation between GDP growth rate and emission rate?

Critical time point analysis:

At this stage, we intended to investigate the critical time point that may change the global CO2 emission and socio-economic status from 1989-2020.

Specific questions:

1. How does the change of economic status influence the CO2 emission pattern during the depression around 2008?
2. How do countries respond to the Paris Agreement?
3. How do countries respond to the COVID-19?

Country-focus analysis:

At this stage, we intended to study a particular country or group of similar countries and analyze their CO2 emissions and economic patterns throughout history.

Dataset

The primary dataset used in this study is provided by Ourworldindata.org. This dataset is built upon several other datasets and processing steps (BP, EIA, Global Carbon Project, CAIT datasets). Secondary dataset: World Energy Consumption dataset was downloaded from Kaggle. Third dataset: GDP data is gathered from the World Bank. Fourth dataset: World-countries geojason dataset, with country geometries

Results

The hypothesized correlation between GDP rate and CO2 emissions rate is observed for majority of countries.

However, some countries shows negative correlation, like Venezuela, Oman and others. During economic recession in 2008, when GDP dropped down, CO2 emission rate went down as well.

Overall, changes in CO2 emissions in global events were prominent in the Americas, Europe, and Oceania. Relatively, the changes were not significant in Asia and Africa. Comparing to other critical time point, no significant changes were observed in Paris Agreement period.

Among the categories, most countries develop their economies and increase their annual CO2 emissions. This correlation is most pronounced in Guinea, Bangladesh, Cambodia and others. The number of these categories reflects the global dependence of development on fossil energy sources. Meanwhile, many countries have managed to cut CO2 emissions while maintaining GDP growth, such as Cuba and Malta. Countries such as Greece, Italy and Japan have decided to cut CO2 emissions, but at the same time their economies are in recession. And finally, some countries, including Iran, Sudan and Tunisia, unfortunately did not achieve an economic recovery, and CO2 emissions increased.

Prepared by Choonghyun Lee, Shamil Murzin, Yangkang Chen. October 16, 2022

Methodology
Analyzed data

Study was conducted by analyzing four sets of data : CO2 emission data, Energy Consumption data, Growth data and Geographic data.

1. CO2 emission data

The primary dataset used in this study is provided by Ourworldindata.org. This dataset is built upon several other datasets and processing steps (BP, EIA, Global Carbon Project, CAIT datasets).
Features used in this study:
1. Country/Region
2. Country/Region Code
3. Year
4. Annual CO2 emissions in tonnes
Location: https://ourworldindata.org/grapher/annual-co2-emissions-per-country?tab=chart&country=~OWID_WRL

2. Energy consumption data

World Energy Consumption dataset from Kaggle. Shape (17432 rows, 122 columns). Data includes many features (119), study focused on only eight features:
1. Country - Geographic location
2. Year - Year of observation
3. fossil_share_energy - Share of primary energy consumption (fossil fuels)
4. Per_capita_electricity - Electricity consumption per capita (kWh)
5. Population - Total population
6. Renewables_elec_per_capita - Per capita primary energy consumption from renewables (kWh)
7. Renewables_share_elec - Share of electricity consumption (renewables)
8. Renewables_share_energy - Share of primary energy consumption (renewables)
Location:
https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption

3. Economic Growth Data

GDP data is gathered from the World Bank. More than +200 country GDP between the years 1960-2020.
Shape (266 rows, 64 columns)
Features:
1. GDP
2. GDP Growth(Annually)
3. GDP per capita
4. GDP per capita Growth(Annually)
5. GDP PPP
6. GDP PPP per capita
Location:
https://www.kaggle.com/datasets/zgrcemta/world-gdpgdp-gdp-per-capita-and-annual-growths

4. Geographic Data

The fourth and final set of data which was used in this study is countries boundary geometry for visualization purposes
Features: geometry for each country (182 lines)
Location: https://github.com/deldersveld/topojson

5. Methodology

All the data was loaded and formatted in Python. Key steps:
1) checking datasets using pandas:
2) Removing NaN values. Next step included merging all datasets into one Dataframe file through the country code, using pandas pd.merge function. All other analyses, described in this report were done based on the merged dataset in Python, using Numpy, Pandas and Altair libraries.

1. Loading data
2. Data QC
3. Data filtering
4. Feature selection
5. Datasets merging
6. Analyses

Methodology
Data Manipulation

1. Creating Common Merged Dataset

1.1 Filtering

Since there were lots of missing data before 1990, we decided to filter out them using pandas library. We also filtered out unnecessary entities such as small islands or European Unions.

df = df[df['Year'] >= 1989]

Country	Year	...
...	1960	
...	...	
...	2020	

→

Country	Year	...
	1989	
	...	
	2020	

1.2 Melting

To visualize the data efficiently, we needed to convert original data columns for each single year into the rows. To do that, we used pandas dataframe and 'melt' function.
We applied this function to every necessary datasets.

gdp = gdp.melt(id_vars=['Country Name'], var_name='Year', value_name='GDP')

Country	1960	...	2020
...			
...			

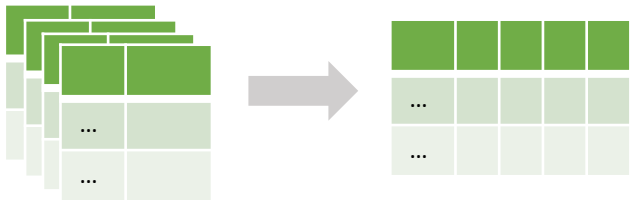
→

Country	Year	GDP
...	1989	
...	...	
...	2020	

1.3 Merging

We needed to combine 4 fundamental datasets to single one based on the key columns like 'Country' and 'Year'.
We used pandas 'merge' function as well as 'merge' function from dataframe depending on the situations.

pd.merge(df1, df2...) or DataFrame.merge(...)



Country Code, CO2 Emissions, GDP, and Energy Consumptions

Merged Dataset

2. Creating Datasets For Purpose

2a. For Worldwide Level Analysis

2a.1 GroupBy

df.groupby(['country','id'])

To group dataframe by country and country id for further correlation calculation

2a.2 corr()

['CO2_rate','GDP_r'].corr().unstack()

To calculate correlation between GDP growth rate and CO2 emission rate in dataset, we used 'corr', and then unstack correlation into individual column

2b. For Critical Time Point Analysis

2b.1 GroupBy

df.groupby('region')[..].rank('dense')

To select top 5 countries in each continent based on the CO2 emission in 2020, we used 'groupby' and 'rank' function.

2b.2 diff() and Applying Lambda

df ['CO2 diff']= df.groupby(['country']).diff() df.apply(lambda x: ...)

To calculate and store CO2 growth rate in dataset, we used 'diff', 'apply' function of DataFrame as well as lambda.

2c. For Country-Focus Analysis

2c.1 df.corr()

df.corr(method='pearson')

Calculate the correlation between metrics and years to represent temporal trends.

2c.2 df.melt()

df = df.melt(id_vars='Year')

Melt dataframe into long form with Year as identifier to show different metrics changed with year.

2c.3 standardize & log

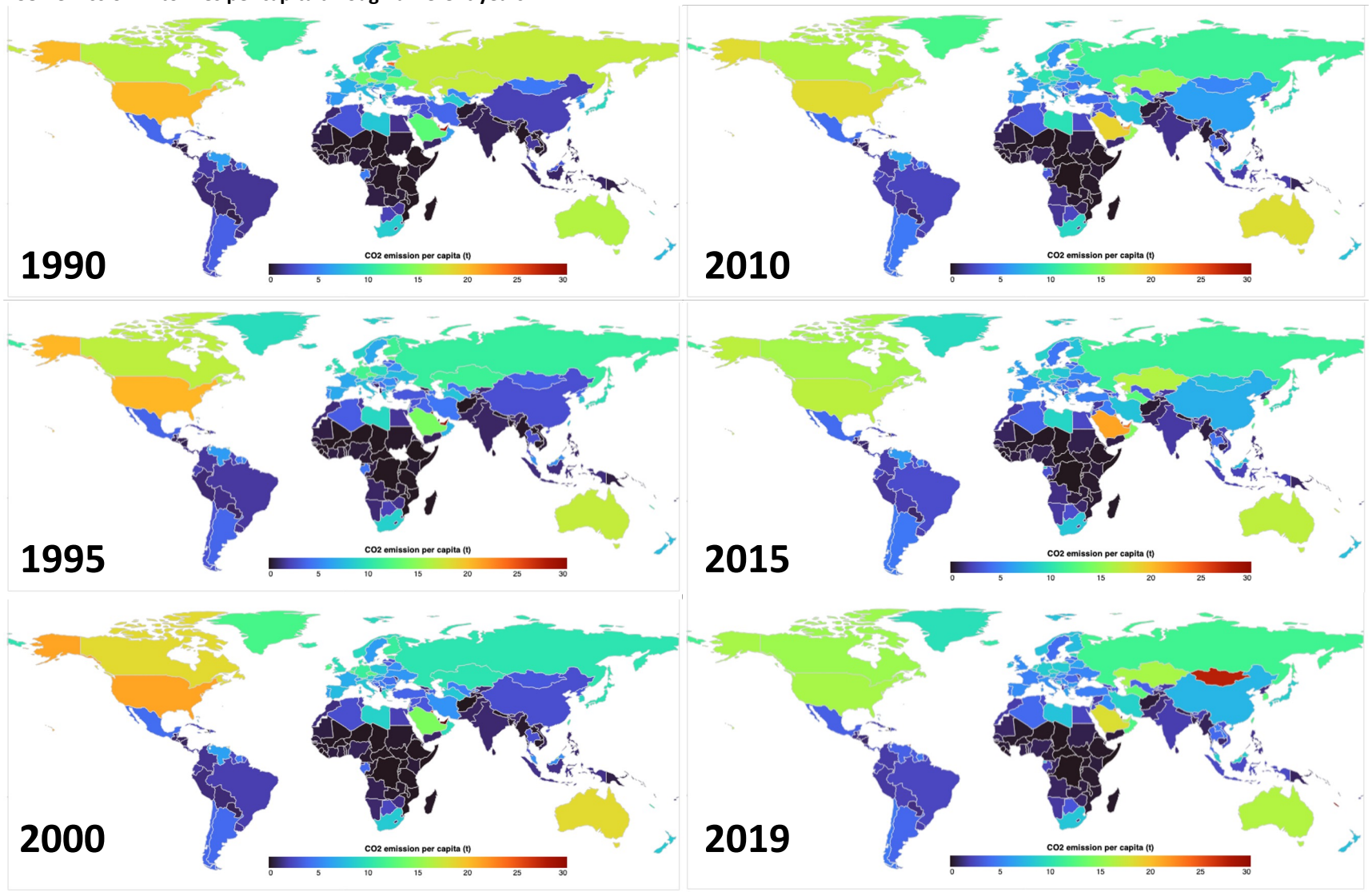
(a - np.nanmean(a))/ np.nanstd(a), np.log()

Standardize and log-transform values to make different metrics comparable.

Worldwide level: CO2 emissions per capita

Countries in C. Africa, S. America and SE Asia produce less CO2 per capita

CO2 emission in tonnes per capita through different years



Understanding how CO2 emissions are distributed around the world

By country, China, the US and India are the countries with highest carbon dioxide emissions, but at the same time they are the most populous countries. Per capita emissions is a better indicator than total emissions because countries have different populations.

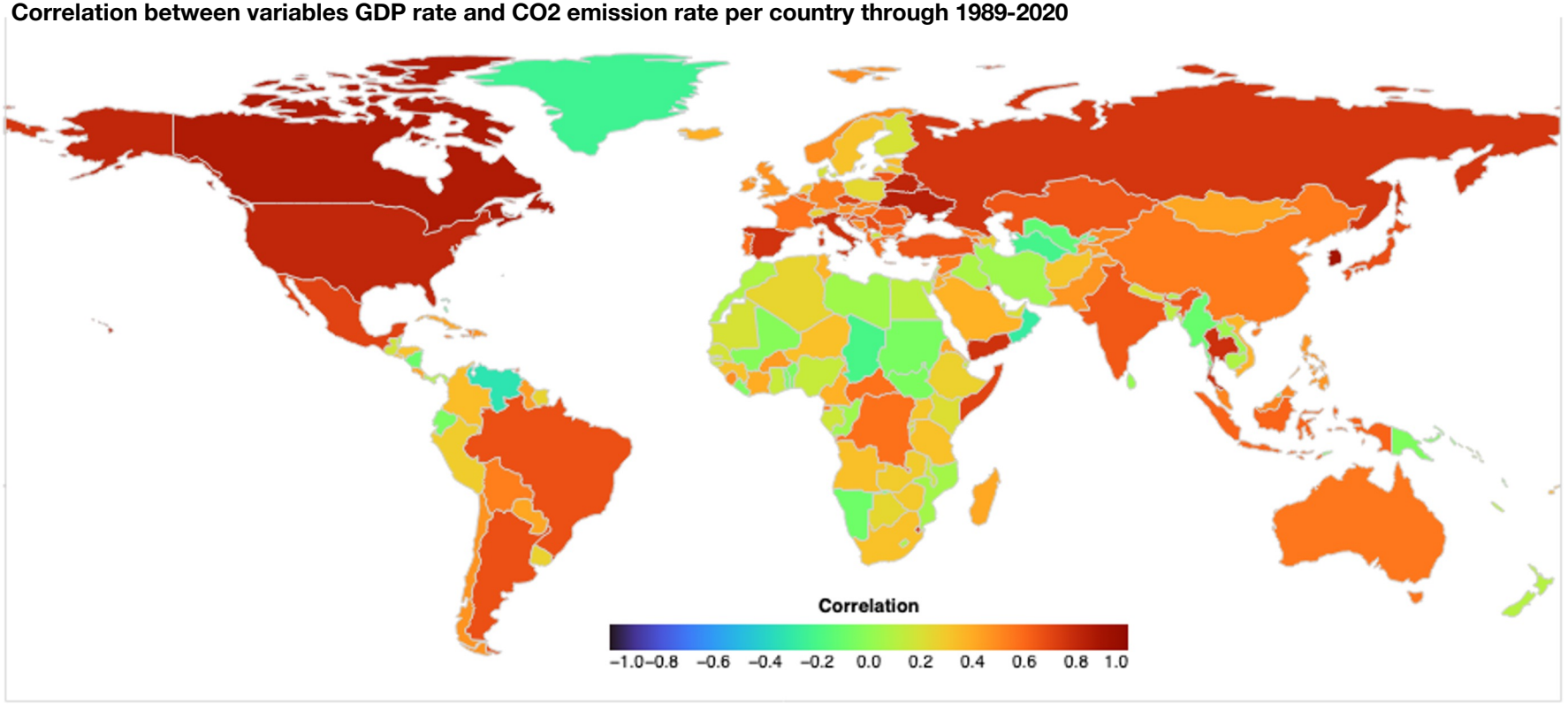
In the visualization on the left, we see the difference in per capita emissions in tons around the world over time (1989-2020). We can observe that in some countries CO2 emissions per capita have decreased over the past 30 years, such as the US and Canada, but in some countries, they have increased, such as China.

The global average in 2019 was 4.88 tons of CO2 per person (t/p). The world's largest per capita CO2 emitters are major oil producing countries, especially for countries with relatively small populations such as countries in the Middle East. Russia is a major oil producer, that also explains the high level of CO2 emissions (11.5 t/p) The lowest per capita emissions are found in Central Africa, where it is less than 1 t/p. The countries of South America emit in the range of 0.9-3.5 t/p. Most populous countries: China - 7.3 t/p, India - 1.9 t/p, USA - 16 t/p.

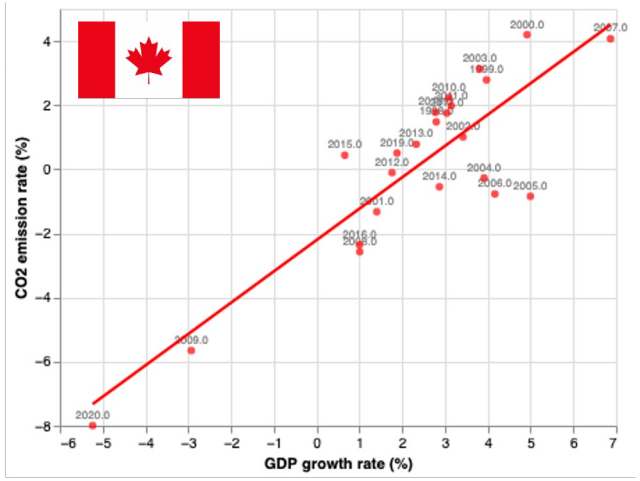
Worldwide level: GDP growth vs. CO2 emission rate correlation

GDP rate positively correlates with CO2 emissions rate for many countries

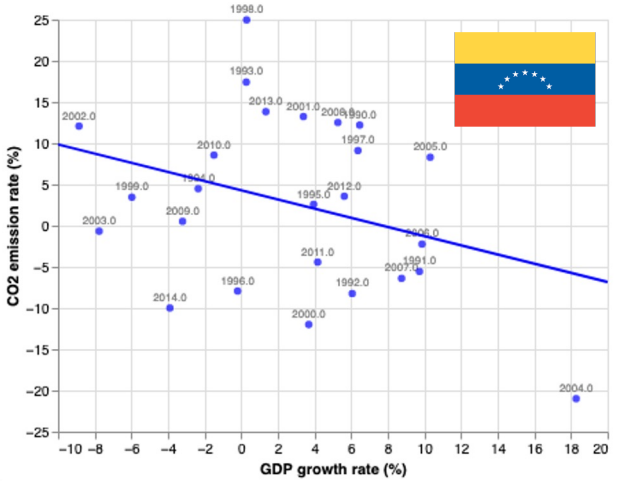
Correlation between variables GDP rate and CO2 emission rate per country through 1989-2020



GDP growth rate vs CO2 emission rate in Canada (1989-2020).



GDP growth rate vs CO2 emission rate in Venezuela (1989-2020).

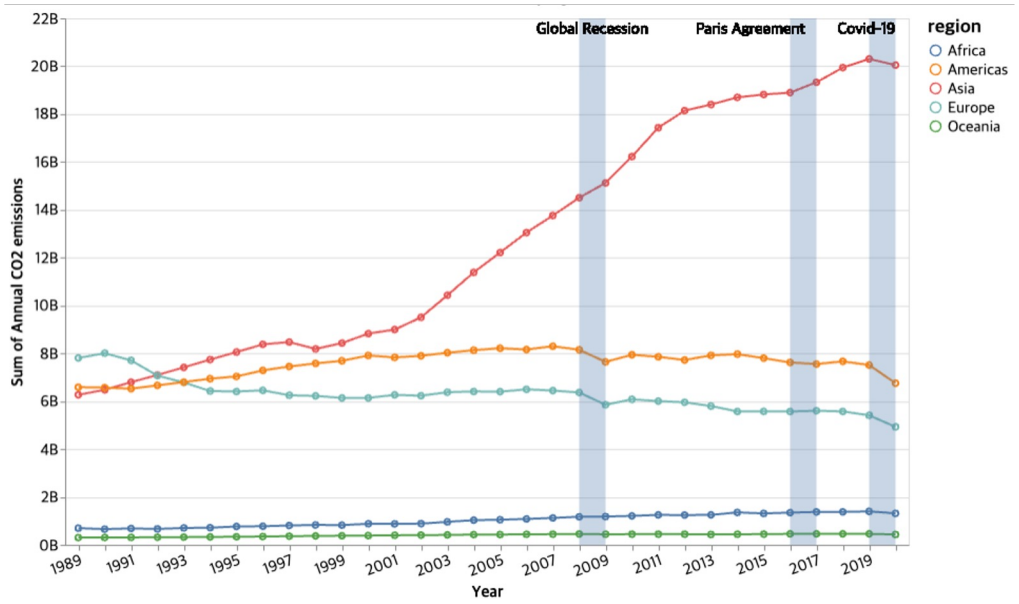


The relationship between GDP growth rate and CO2 emissions rate is tested for all countries for a period 1989-2020 (as data were available for mostly all countries only in this period). To quantify the relationship between these two variables (linear correlation), Pearson's correlation coefficient was used. The coefficient value ranges between -1.0 and 1.0. A coefficient of -1.0 shows a perfect negative correlation and 1.0 a perfect positive correlation. A coefficient of 0.0 on the other hand means that there is no relationship between the two variables. **The regression results show that in most countries there is a strong positive correlation between GDP growth and CO2 emissions, but some countries showed a negative correlation.**.

Critical time point analysis - Recession, Paris Agreement, COVID-19

CO2 emission patterns in global events show different patterns by continent

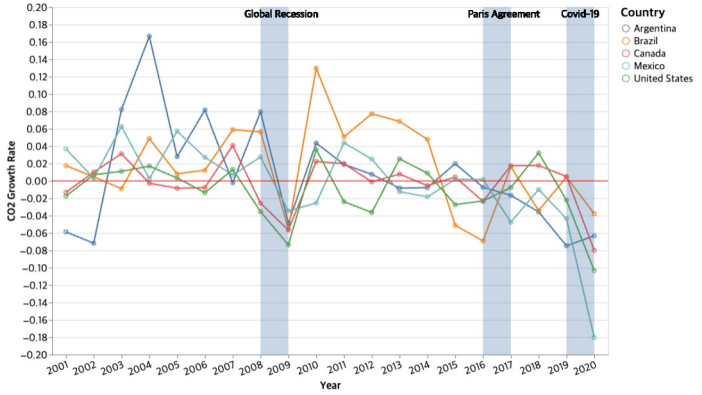
CO2 Emissions Trend by Continent



The trend of CO2 emissions in Americas and Europe seems to have some correlations with all three events (Global recession, Paris Agreement, Covid-19). But in other regions, only Covid-19 seems to have a relatively weak correlation with the CO2 emissions.

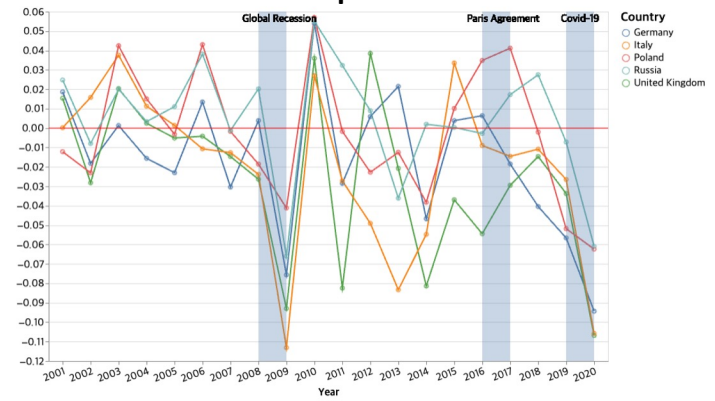
CO2 emissions Growth Rate by Continent 2001~2000, top 5 countries

Americas



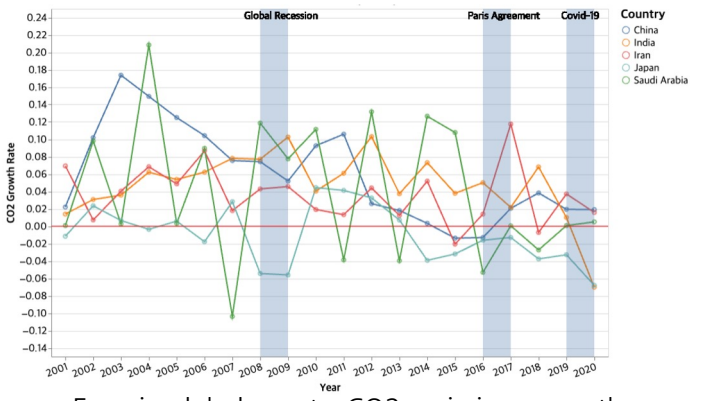
CO2 emissions during all 3 events show decreasing patterns in all 5 countries.

Europe



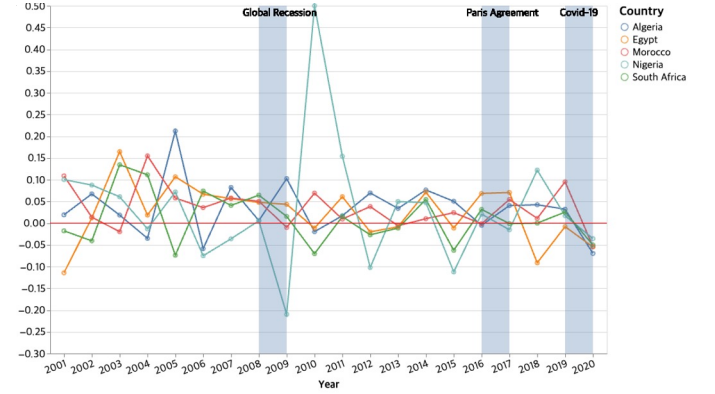
CO2 emissions during Global Recessions and COVID-19 show decreasing patterns in all 5 countries.

Asia



Even in global events, CO2 emissions growth rates are comparatively stable in the countries.

Africa



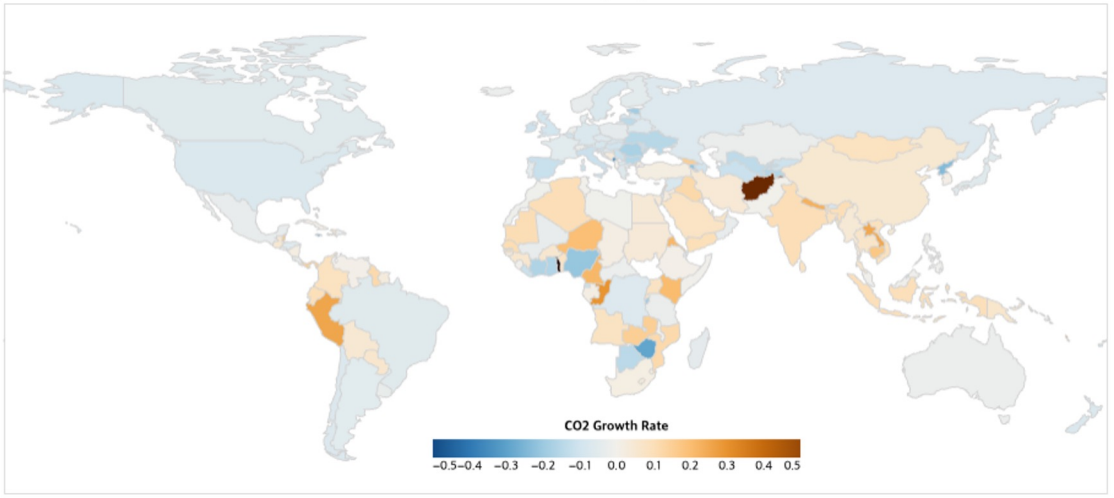
CO2 emissions growth rates are comparatively stable but in Pandemic, the figures show consistent negative trend.

* Top 5 countries in each continent were selected by amount of CO2 emissions in 2020.

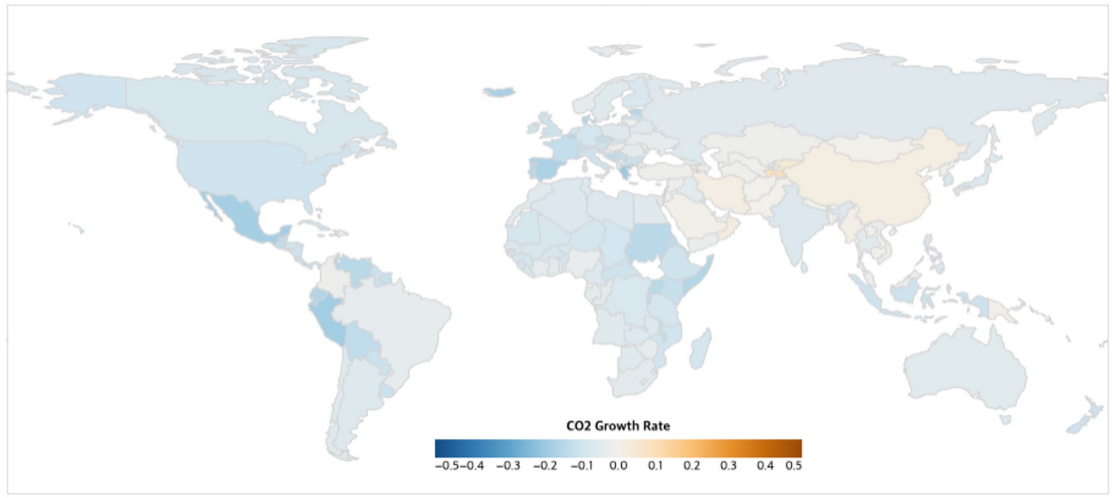
Critical time point analysis - Recession, Paris Agreement, COVID-19

Among three events, COVID-19 has the largest impact on CO2 emission globally

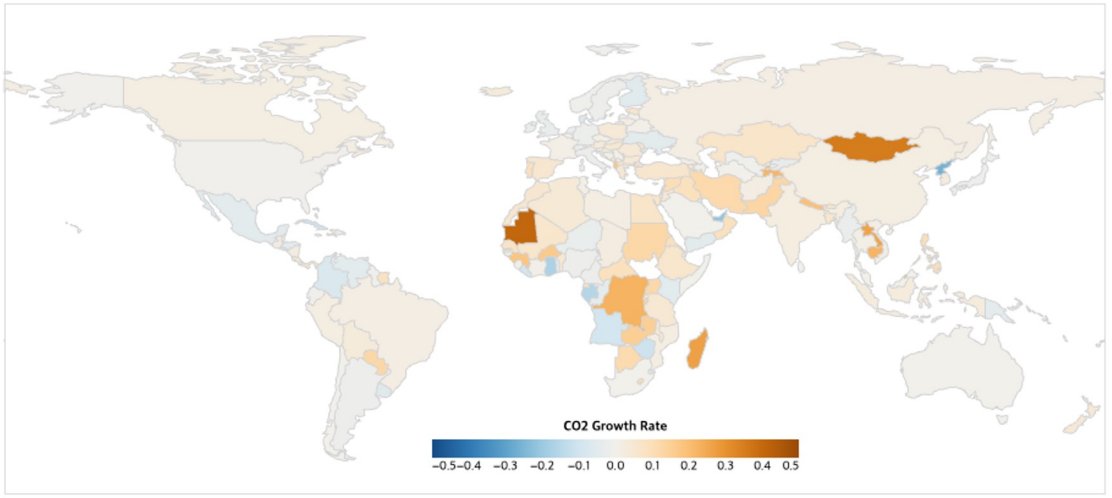
CO2 Emission Growth Rate by Recession (2008~2009)



CO2 Emission Growth Rate by COVID-19 (2019~2020)



CO2 Emission Growth Rate by Paris Agreement (2016~2017)

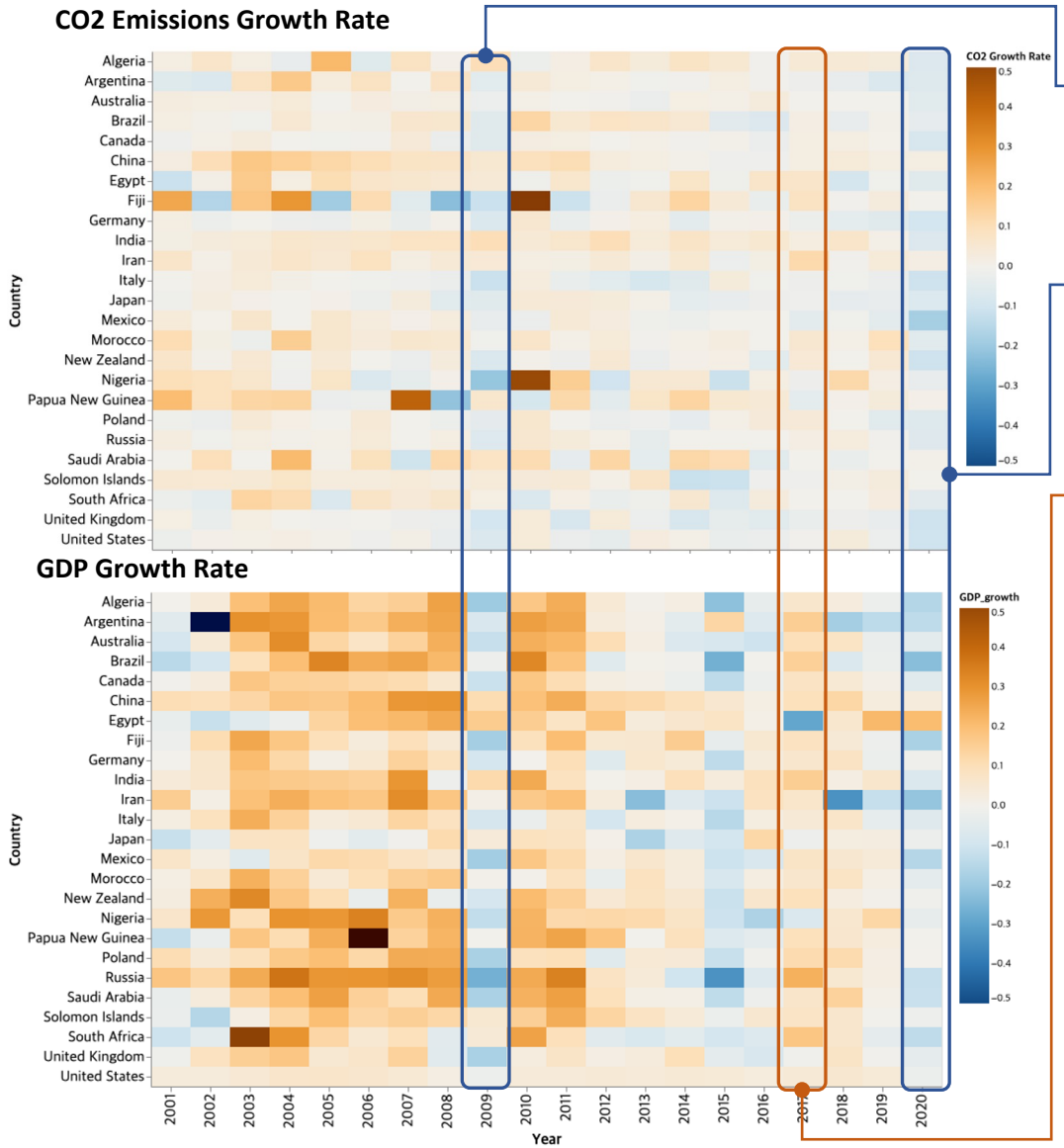


Findings

- In recession period, CO2 emissions declined in most American, Oceanian, and European countries while those in most Asian and African countries increased.
- Even after Paris agreement was accepted, CO2 emissions were still increasing in most countries except some European countries.
- In pandemic period, CO2 emissions were decreased in most countries except some Asian countries.
- However, what if some countries' economy is shrinking or developing regardless of the situation? Can we find any possible confounders?

Critical time point analysis - Recession, Paris Agreement, COVID-19

In Global Recession and Pandemic, Both GDP and CO2 Growth Rates declined



Global Recession ('09) : GDP Growth↓ , CO2 emissions ↓

- Out of a total of 25 countries, annual GDP declined in 19 countries and CO2 emissions in 16 countries declined.
- American and European countries show lower GDP and CO2 emission growth rates.

COVID-19 ('20) : GDP Growth ↓ , CO2 emissions ↓

- Out of a total of 25 countries, annual GDP declined in 23 countries and CO2 emissions in 22 countries declined.
- China is the only country with positive GDP and CO2 emissions growth rates.

Paris Agreement('17) : GDP Growth ↑ , CO2 emissions -

- Out of a total of 25 countries, annual GDP declined in 4 countries and CO2 emissions in 10 countries declined.
- GDP Growth pattern and CO2 emissions show mismatched in 10 countries.

Findings

- GDP Growth Rate and CO2 emissions Growth Rate show similar patterns, decreasing in Global Recession and Pandemic Period while they show different patterns in regard to Paris Agreement.
- Based on our previous observations, GDP growth has a positive correlation with CO2 emissions. But in the Paris Agreement period, it seems to support the counter argument: high growth rate and relatively low CO2 emissions growth rate. Paris Agreement might have a impact on the result. However, more researches are needed to validate the assumption.

* Top 5 countries in each continent were selected by amount of CO2 emissions in 2020.

Country level analysis: A Promising future

GDP growth accompanied by CO2 emission reduction

CO2 emission reduction can be benign to economy

Take US and China as examples, the GDP of both countries grow in monotonicity since 1990. Meanwhile, fossil share of energy presents downtrend and renewable share of energy present uptrend for both countries. However, the annual CO2 emission of US decline sharply since 2007, and reach a low level in 2020, which is even lower than that of 1990. On the other hand, China keeps increasing the annual CO2 emission and reach a historical high in 2020, although the trend seems to slow down a lot.

This interesting comparison indicate that the booming of economy do not necessarily rely on productive process that produce CO2 as by-product. With a high service sector proportion in national economy, US is starting to achieve a green economy.

The result also reveals that the economy of developing country still rely heavily on industry with CO2 emission as by-product, but in the future, another green way is foreseeable.

Summary

- CO2 emission is not necessarily correlated with GDP in country level.
- While renewables take larger fraction in energy consumption, CO2 emission can still increase rapidly.
- The inner correlation between economic and energy consumption need further investigation.

Features

- log Annual CO2 emissions
- log Fossil share energy
- log GDP
- log Renewables share energy

Country level analysis: CO2 & GDP change in the past 10 years

One planet, divergent trajectory

One planet, divergent trajectory

While we all live on the same planet, different countries have different economical and environmental trajectory in the past 10 years.

Among the categories, **most countries develop their economy and increase the annual CO2 emission.** This correlation is most significant in Guinea, Bangladesh and Cambodia, etc. The amount of this categories reflect a **global dependent on fossil energy for development.** Meanwhile, many countries managed to reduce their CO2 emission while maintaining GDP growth, like Cuba and Malta.

CO2 emission reduction can be positive or negatively correlated with economy

Countries like Greece, Italy and Japan choose to reduce their CO2 emission, **but their economy is in recession simultaneously**. And finally, some countries including Iran, Sudan and Tunisia, unfortunately, did not boom the economy while gain an increase in their CO2 emission.

While most of the countries Increase their renewables share in energy, some did not, indicating **different environmental-friendly development strategies**.
This result remind us that Global Warming management is an complex issue which influence countries in different ways. Expert need to consider specific circumstances when making decision.

Summary

The hypothesized correlation between GDP rate and CO2 emissions rate is observed for majority of countries. When production increases in intensity, by-products and pollution increase as a result. However, some countries shows negative correlation, like Venezuela, Oman and others. **Understanding of the negative correlation needs further investigation.** During economic recession in 2008, when GDP dropped down, CO2 emission rate went down as well.

Overall, changes in CO2 emissions due to global events were prominent in the Americas, Europe, and Oceania. Relatively, changes in CO2 emissions were not significant in Asia and Africa.
Recession and pandemics had a significant impact on both GDP and CO2 emissions, but the Paris Agreement had no significant effect on GDP and only had a minor effect on CO2 emissions in Europe and the Americas.

Among the categories, most countries develop their economies and increase their annual CO2 emissions. This correlation is most pronounced in Guinea, Bangladesh, Cambodia, etc. The number of these categories reflects the global dependence of development on fossil energy sources. Meanwhile, many countries have managed to cut CO2 emissions while maintaining GDP growth, such as Cuba and Malta. Countries such as Greece, Italy and Japan have decided to cut CO2 emissions, but at the same time their economies are in recession. And finally, some countries, including Iran, Sudan and Tunisia, unfortunately did not achieve an economic recovery, and CO2 emissions increased. While most of the countries Increase their renewables share in energy, some did not, indicating different environmental-friendly development strategies.

Statement of Work and References

All

- Drafting proposal
- Data exploration
- Data manipulation
- Data cleaning
- Report preparation

Shamil Murzin

- Data availability analyses
- Data merging
- Correlation map
- Correlation plots country

Yangkang Chen

- Data manipulation
- Line charts, maps, word cloud
- Country-focus analysis
- Code pruning

Choonghyun Lee

- Critical time point analysis
- Data merging
- Visualizing line charts, world maps, and heatmaps
- Ethical issue consideration

[1] Houghton, E. (1996). Climate Change 1995: The Science of Climate Change: Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change, 2. Cambridge: Cambridge University Press.